SIMULATING SEISMIC SIGNAL DETECTORS

Yu. V. Fedorenko[1, 2] and Eystein S. Husebye[1]

1. Institute of Solid Earth Physics, University of Bergen, Norway
2. Institute of North Ecology Problems, Kola Science Center Russian Academy of Sciences

Sponsored by U.S. Department of Defense
Defense Threat Reduction Agency
Grant DSWA01-98-C-0159

ABSTRACT

x

The initial, critical step in Comprehensive Nuclear-Test-Ban Treaty (CTBT) International Monitoring System (IMS) seismic monitoring is that of detecting signals in ambient noise at many stations deployed globally. A commonly used detector is the STA/LTA, which is tied to an almost continuous comparison of 'rectified' or RMS trace power in two windows of short (STA) and long (LTA) durations. Moreover, for enhanced performance, the records are pre-filtered in suites of frequency bands for noise suppression prior to the detection process itself. As well-established observationally the noise vary on a diurnal basis; nearly monochromatic during nights and week ends since microseisms dominate while more white during day time due to the influx of high-frequency cultural noise from traffic, running machinery etc. For a fixed STA/LTA threshold, the false-alarm rate is higher during night time. This empirical example serves to illustrate a major problem, highly relevant in the context of CTBT/IMS monitoring, namely how to quantify performances of signal detectors of various designs? We have explored this problem through synthetic but realistic modeling.

Firstly, artificial noise traces are generated via the Gaussian probability density function (p.d.f) for all three sensor (3C) components for which noise is presumed uncorrelated. Processing in terms of noise suppression is achieved through wavelet transform discarding FIR and Butterworth kinds of bandpass filtering (Fig. 1). Test statistics are of the STA/LTA and Kolmogorov-Smirnov (K-S) types defined for four different types of 3C sensor "power-rectified amplitude" combinations. Using the Monte Carlo Method we addressed critical detector design issues through numerical simulations. We started with estimations of confidence intervals for detector performance statistics for given false alarm rates. STA/LTA proved better than K-S types test statistics in terms of variances of critical false alarm levels. The STA/LTA detector is also best if noise is free of spikes and/or other artificial interferences. It can recognize signals with signal-to-noise ratio down to 1.0 given moderate false alarm and missing signal rates of 0.001 and 0.1 respectively. However. inserting spikes with both flat time and amplitude p.d.f.s gave that the STA/LTA detector performances degrades while K-S type detectors are insensitive to spikes. In particular STA/LTA power detectors perform poorly. Likewise, K-S detectors unlike STA/LTA ones have relatively poor performances for "sharp" signals like those generated by WKBJ synthetics but excellent performances for "flat" signals like Pn, volcanic tremors and the like. We have also considered single sensor versus 3C combinations; the latter have better performances for large to moderate angles of incidence ($i > 20°$). Finally, this study is complementary to (Hysebye and Fedorenko, 1999, ibid) because its relevance to seismic stations designed in order to solve real-time phase-picking and event recognizing problem.

Key Words: Signal detectors, STA/LTA, Kolmogorov-Smirnov, test statistics, spike degrading, 3C enhancement.

OBJECTIVE

The objective of this research is to simulate realistically signal detector designs and the corresponding performances for typical operational scenarios for 3-component (3C) stations. The noise is presumed Gaussian while signals are generated by WKBJ synthetics and relative component amplitudes made a function of incident angles. Four different test statistics are considered i) | Z |, ii) Z^2, iii) (Z^2; $N^2 + E^2$) and iv) $Z^2 + N^2 + E^2$ where Z, N and E are respective seismometer components. Two signal detectors are currently tested, namely the popular STA/LTA and the non-parametric Kolmogorov-Smirnov (K-S) detectors in one- and two-dimensional sample space. Preliminary synthetic results are as follow; rectified power STA/LTA detectors have better performances than K-S type detectors. However, for spike prone recordings STA/LTA detector performances degrade in particular for power test statistics while the K-S detectors are insensitive to such irregularities. Most interesting and promising results so far is that including horizontal components in the test statistics gives enhanced signal detection performances for signal incident angles exceeding 20° relative to Z-components on a stand alone basis. Despite moderate testing on real data one result stands out; so-called unexplainable missed signal detections are probably due to blinding of the STA/LTA detector caused by spikes. This is not an entirely uncommon phenomenon in many local network operations where signals are missed by the digital detector but clearly seen visually in the analog seismograms.

RESEARCH ACCOMPLISHED

The advent of digital recording in seismology also introduced the problem of detecting relatively weak signals in ambient noise; strong signals do not pose any problem in this context. The most popular and widely used signal detector is the so-called STA/LTA or variants hereof which is a power ratio measured over two separate windows. The Short Term Average (STA) is measured over windows being 1 - 3 sec while the Long Term Average (LTA) is measured over long windows exceeding minutes. For Gaussian noise the STA/LTA detector is approximately optimum explaining the general popularity of this simple detector in array and network monitoring (Ruud *et al*., 1993). There has also been some "competition' in designing the best signal detector; best in the sense of detecting the largest number of events within a given time period for a given data set. Such contests have never proved conclusive since close to the STA/LTA threshold the number of false alarms increases steeply while the probability of earthquake occurrence only increase linearly according to the earthquake recurrence relation. Besides there is no simple way of confirming the presence of marginal signals from other stations. Also in a CTBTO/IMS context of global seismic monitoring detector design is of considerable interest since only about 50-70 percent of reported signal detections are associated with time-space matching signal detections from other stations as required for acceptable epicenter locations.

A peculiarity of network operations is that most stations are equipped with 3-component seismometers while only vertical recordings are used for detection purposes. In this paper we first discuss briefly the logic of signal detector design and then present different signal diagnostics including joint usages of both vertical and horizontal recordings. The detector performances for various designs and noise conditions would be simulated in a realistic manner and in this context we discuss the redundancy of non-associated signal detections.

Basic logic of signal detector design

From the point of view of statistical hypotheses testing any signal detector perform the same task. It compares current recordings with the previous ones forming the equivalent of the STA/LTA ratio. Depending on the specific value of this ratio, larger or smaller than a prespecified threshold value Y,, the choice is one of two alternative hypotheses. If smaller than Y,, then Ho or the null hypothesis is true that is current data is noise. The alternative or Hi hypothesis is true for exceedance of Y,, that is current data is signal or more correctly (signal + noise). To choose between Ho and HI statistically, we need to translate these two hypotheses from qualitative into quantitative terms. Usually this is done by specifying a probability distribution for the parameters describing the data (often denoted "test statistics"). For example, the STA/LTA detector uses the STA/LTA ratio as test statistics. If this ratio exceeds the threshold level, H_o is rejected that is, H_1 is accepted meaning that a signal is detected.

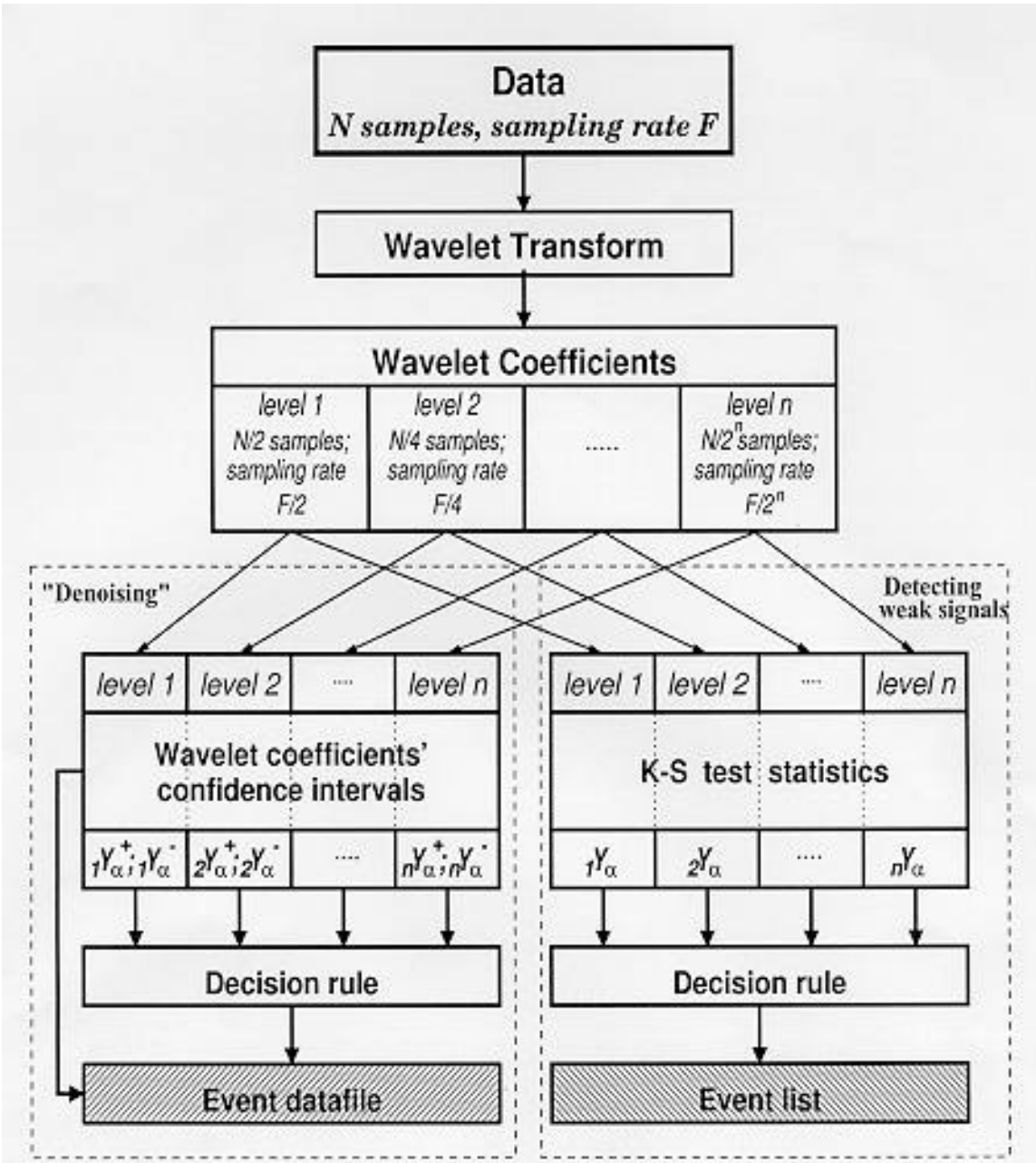

Fig. 1. Dataflow chart. N data samples, sampling rate F are subject to wavelet transform that results in wavelet co-efficients of n hierarchy levels, $n = \log_2 N$. Sampling rate at each level k is $F_k = F/2^k$. "De-noising" algorithm processed wavelet co-efficients at each level for each component independently obtaining estimates for upper and lower critical values ${}^{i}_{k}Y_a^{+}$ and ${}^{i}_{k}Y_a^{-}$ (confidence interval boundaries). Here $i = 1,2,3$ is component number, $k = 1, \ldots, n$ is wavelet co-efficients' hierarchy level, α is confidence level. If the number of wavelet co-efficients that fell outside of confidence interval $\left[{}^{i}_{k}Y_a^{+}, {}^{i}_{k}Y_a^{-}\right]$ exceeds some threshold value, to say 30 to 50%, data are written to the hard disk together with ${}^{i}_{k}Y_a^{+}$ and ${}^{i}_{k}Y_a^{-}$. Latter parameters will be used later for "de-noising" of seismic traces as shown in (Husebye and Fedorenko, ibid). Algorithm responsible for detecting of weak signals has been described in details above.

In general, decision rules may lead to two types of errors: H_0 when H_0 is true and H_0 accepted when H_0 is false. These are called "Type I" and "Type II" errors, respectively. In our case Type I errors are called "false alarm" and Type II errors are called "missing signal," which is easier to understand. There is no way to know for sure whether our decision is incorrect, but it is possible to evaluate the probability of making an incorrect decision. The probability of a "false alarm" or Type I error is referred to as the test's level of significance and is denoted by α. The probability of "missing signal or Type II error is denoted by β. If we know or presume the probability density function (p.d.f.) for the test statistics we can compute the "false alarm" and "missing signal" critical values Y,, and Y,3 which in turn completely determine the decision-making process. Since we aim at designing detection algorithms not using any assumption on the p.d.f. of the test statistics, we use data in the reference time window to obtain the critical value Y_α. How to assess Y_α, is demonstrated below using synthetic signal and noise data. Firstly, we assume that that current data are in short windows (STA equivalent, denoted below as STW) of 2- to 5-sec length. Duration of reference (LTA equivalent, LTW) window is arbitrary and chosen to be some 30 to 60 min. The LTW choice is long enough to reduce test statistics variance and short enough to adapt to changes in the noise structure. Note, that seismic noise exhibits diurnal variations; nearly monochromatic during nights due to dominance of microseism while more white during daytimes due to human activities like fast running machinery etc. Anyway, using a sliding procedure each STW is compared with LTW and the resulting test statistics Y_i, i = 1,.., N are sorted and stored in descending order. Obviously Y_α, = Y_k, where $k = \alpha N$ and N is large. The estimator Y_α, is a function of the random data so Y_α, itself is a random variable and its behavior is described by the presumed p.d.f. Its quality depends on whether or not the estimator's p.d.f. is suitably centered over its expected value. In other words, Y_α, must be unbiased. Another important estimator feature is precision implying that its p.d.f. dispersion should be as small as possible. The precision measure is used for choosing between several estimators.

To study the estimator Y_α, we assume that i) Y_α, has Gaussian distribution with σ = 1; and ii) Y_α has Kolmogorov-Smirnov distance p.d.f. for N_1, = 32 and N_2 = 4096 where N_1, and N_2 are numbers of samples in STW and LTW, respectively. Exploring case i) is of practical interest because many distributions converge asymptotically to Gaussian. Case ii) is important because, as to be shown later, Kolmogorov-Smirnov "distanc" D has several advantages versus other detector test statistics. For instance, extensive Monte Carlo integrations have shown that the distribution of D is nearly identical for quite different distributions (Press et al., 1993), so the Kolmogorov-Smirnov test does not require the assumption that the noise is normally distributed. Monte Carlo techniques were used to evaluate Y, mean values and confidence intervals for α = 0.1, 0.01 and 0.001. Successive Gaussian random deviates are produced by routines from the Numerical Recipes library (Press et al., 1993), while K-S random deviates are obtained by the "rejection" method. Here the K-S distribution for D is approximately $p_D(x) = dQ_D/dx$, where $Q_D(x) = 2\sum_{j=1}^{a}(-1)^{j-1} e^{-2j_2x_2}$ then $p_D(x) = -8x\sum_{j=1}^{a}(-1)^{j-1j^2e-2j^2x^2}$ (Press et al., 1993) . When x is distributed according to the $p_D(x)$, distance $D = x(\sqrt{N} + 0.12 + 0.11/\sqrt{N})$ for one-dimensional K-S test (which was applied to $|Z|$, Z^2 and $Z^2 + N^2 + E^2$); however $D = x\sqrt{N}\,[1 + \sqrt{1- r^2}\ (0.25 - 0.75/\sqrt{N})]$ for two-dimensional K-S test which was applied to $[Z^2; N^2 + E^2]$) and r2 = $(r_1^2 + r_2^2)$ / , r_1 and r_2 are correlation coefficients between Z^2 and $N^2 + E^2$ in STW and LTW. Fast algorithm ($n$log$n$) for calculating the two-dimensional K-S test distributed under Gnu GPL and written by Andrew Cooke has been used.

The results in Figs. 2 and 3 imply that the above method for estimating critical "false alarm" value Y, can be used successfully in real-time detectors. In order to obtain stable results about 5000 test statistics samples should be stored which in turn corresponds to 2.5 - 6 hours of recording and hence is representative of daily noise variations. In addition, the $\rho_D(\chi)$ virtue of being almost independent of the actual noise distribution so asymptotic approximation of the p.d.f. may be used in cases of a small α.

Simulating signal detector performance.

As mentioned above, using 3-component (3C) seismic recordings in network operations it may be worthwhile to use all 3 components in the detection process. There are many possible ways to construct such detector test statistics but our preference are for the 4 combinations i) Z; ii) Z^2; iii) (Z^2 ; N^2+ E^2) and iv) Z^2 ; N^2+ E^2 where Z, N, and E are respective seismometer components. Below Z, N, and E are also used

as component indices. Anyway, our signal $s(t)$ and noise $n(t)$ models have zero mean Gaussian p.d.f. ρ, (χ) and ρ_n, (χ)

$$p_s(x) = \exp(-x^2/s_s^2)/\sqrt{2\pi s_s}; \quad p_n(x) = \exp(-x^2/2s_n^2)/\sqrt{2\pi s_n}$$

where σ_s^2 and σ_n^2 are signal and noise variances and signal-to-noise ratio (SNR) is σ_s^2/σ_n^2. Vertical n_z, North n_N, and East n_E noise components are independent; $< n_N^2 + n_E^2 + n_Z^2 > = <n^2>$ and $< n_N^2 >=< n_E^2 >=< n_Z^2 >$ where angular brackets imply averaging.

CONCLUSIONS AND FUTURE PLANS

Our extensive experiments in simulating 3-component seismic signal detections in realistic operational environments are instructive for their designs and workings. Somewhat surprising is how easily poor data like spiky records can ruin performances for STA/LTA type of detectors. The non-parametric Kolmogorov-Smirnov detector is insensitive to such irregularities since rank testing is used for declaring signal presence. The other significant result obtained so far is that utilizing also the horizontal components in the signal detection process performances will improve even at teleseismic ranges i > 20°. The above results reflect extensive synthetic experiments for in our opinion realistic noise and signal conditions. Practical experiences confirm these results. For example, the often puzzling missing signal detections when the signal *per se* is clearly seen in analog records are obviously due to the 'blinding' of the detector due to spike occurrences. This is a persistent problem in some national seismograph networks but far less so for CTBT/IMS stations due to the high technical standard of the latter. Regarding horizontal components our results here reflect the common observation that at local ranges the strongest signal amplitudes are often on the EW- or NS-component. Signal shape is also important in detector design where the STA/LTA detector favor sharp onsets while the K-S detectors are relative insensitive to such signals and spikes as well. On the other hand, signals with gradually growth like Pn are sometimes missed by STA/LTA detectors and the later detected Sg/Lg-phase will then initially be identified erroneously as Pn. This point to likely advantages by operating signal detectors of different design properties to better match variations in signal shapes and durations for different distant ranges and tectonic environments.

Our future plans aim at running both STA/LTA and Kolmogorov-Smirnov detectors in parallel for local 3C stations and naturally for various kinds of test statistics. We will also investigate detector differentiations (phase identifications) between P- and S-arrivals at local distance ranges which is not much of a problem using seismic array data through inherit spatial resolution here. Up to now most of our practical tests have been tied to local (Norwegian) network data while in a CTBT context IMS network data should be used not at least in view of the uniform high technical quality of IMS stations.

REFERENCES

Husebye, S. E. and Yu. V. Fedorenko, 1999. First Breaks - Automatic Pickings of P- and S-offsets, ibid.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, 1992. Numerical Recipes in C: The Art of Scientific Computing, 2nd Edition, Cambridge University Press, MA, USA.

Ruud, B. 0. and E. S. Husebye, 1992. A New Three-Component Detector and Automatic Single-Station Bulletin Production. Bull. Seism. Soc. Am., **82**, 221-237.

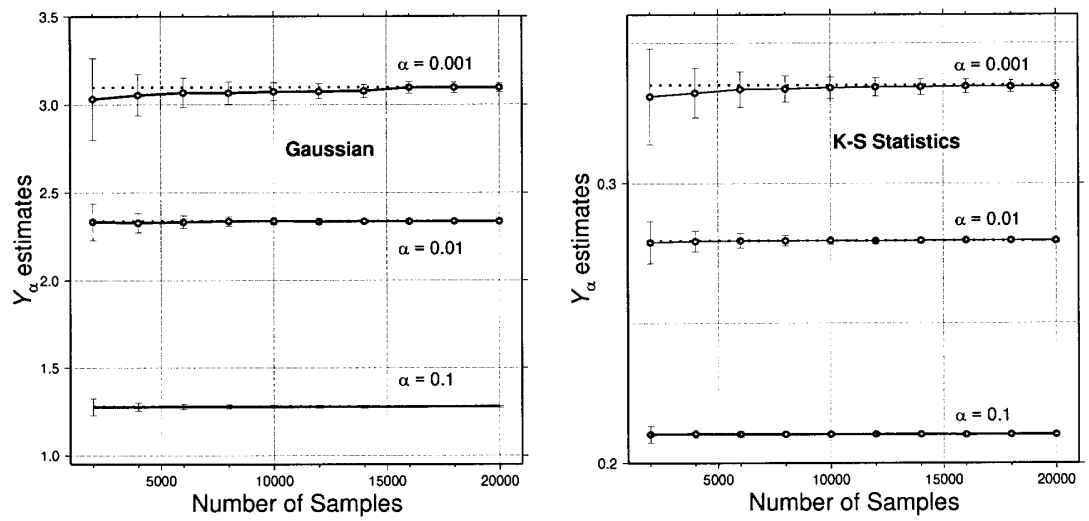

Fig. 2. Y_α estimator mean value versus N (number of samples). Error bars indicate 90% confidence intervals. Dot lines show theoretical values for Y_α. Estimators are consistent but have a small biases for $\alpha = 0.001$ and $N < 10000$.

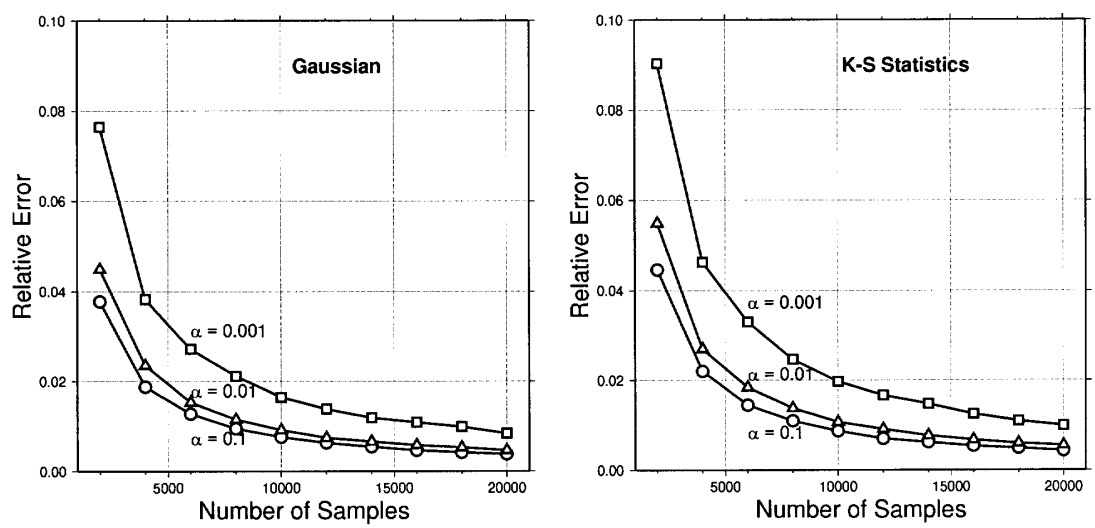

Fig. 3. Relative errors of estimator Y_α versus N (number of samples). Error values calculated as $\delta Y_\alpha = \Delta Y_\alpha / (\overline{Y}_\alpha - Y_{\alpha=0.5})$ where ΔY is 90% confidence interval, $\overline{Y}_\alpha$ is mean value for given α's and $Y_{\alpha=0.5}$ is median value. The Gaussian p.d.f. median value $Y_{\alpha=0.5} = 0$ while K-S p.d.f. median $Y_{\alpha=0.5} = 0.1422$ for $N_1 = 32$ and $N_2 = 4096$. The K-S estimator has smaller confidence intervals ΔY_α but its relative error δY_α is larger than that of the Gaussian.

Seismic signals are modeled as upgoing P-wave with apparent incidence angle ϕ and azimuth θ;

$$[s_N,\, s_E,\, s_Z]^T = [\sin\phi\cos\theta,\, \sin\phi\sin\theta,\, \cos\phi]^T s$$

where $\tan\phi = F\tan i$, i is "true" angle of incidence and F is :

$$F = \frac{2\cos i\cos j}{1 - 2\beta_0^2\sin^2 i/\alpha_0}, \quad \frac{\sin j}{\beta_0} = \frac{\sin i}{\alpha_0}$$

where α_0 and β_0 are surface P- and S-wave velocities. The STW contains $M_s = 32 = N_1$ samples of signal+noise, $S_i = s_i + n_i$, $s_i = s(t_i)$, $t_i = (i-1)/f_{sampl}$, $i = 1, .., M_s$, f_{sampl} is sampling rate while the LTW contains $M_n = 4096 = N_2$ samples of pure noise n_i, $i = 1, .., M_n$. Note that the azimuth parameter is excluded from the test statistics above.

Spikes occurrences in seismic data are not entirely unusual and in particular seismic stations in towns are occasionally spike prone. In order to investigate the influence of spikes on signal detector performance

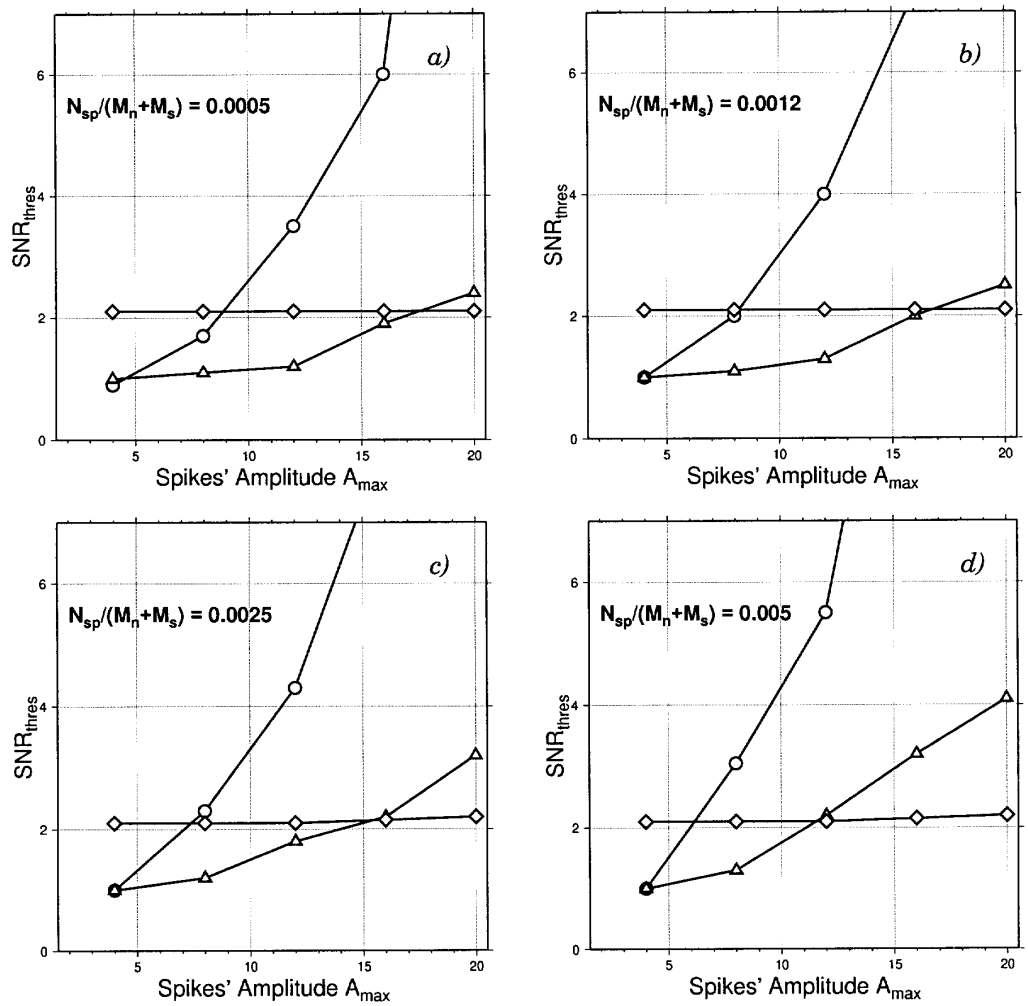

Fig. 4. Spike $A_{\max}$ influence to detector performance for different rates $N_{sp}/(M_n + M_s)$ that is the relative number of spike occurrences. Circles indicate SNR_{thres} for STA/LTA detector test statistics (eq. 2), triangles - for "rectified" STA/LTA (eq. 1), diamonds - Kolmogorov-Smirnov D-statistics (eq. 3). STA/LTA that used power ratio is very spike dependent, on the other hand "rectified" STA/LTA is superior to K-S for small spikes' amplitudes and rates but it is also spike dependent. The K-S D-statistics has poorest performance when noise is pure Gaussian, but on the other hand it is not spike dependent. The latter is not unexpected because D-statistics uses ranks of samples while STA/LTA detectors use amplitudes of samples.

we proceeded in the following manner. Firstly, spikes are modeled as N_{sp} independent single samples uniformly distributed in time with p.d.f. $p^t_{sp}(x) = 1/(M_n + M_s)$, $N_{sp} \ll M_n + M_s$. Likewise the amplitude of spikes also supposed to be uniformly distributed with p.d.f. $p^A_{sp}(x) = 1/A^{\max}_{sp}$. To summarize, the N_{sp} and $A^{\max}_{sp}$ parameters define spikes appearances in the modeled seismic records.

Three types of detector test statistics were studied at this stage, all of them include the 3 seismometer components;

$$R_1 = \frac{\sum_{j=1}^{M_s}(|S_{Nj}| + |S_{Ej}| + |S_{Zj}|)}{\sum_{j=1}^{M_n}(|n_{Nj}| + |n_{Ej}| + |n_{Zj}|)} \quad (1)$$

$$R_2 = \frac{\sum_{j=1}^{M_s}(S_{Nj}^2 + S_{Ej}^2 + S_{Zj}^2)}{\sum_{j=1}^{M_n}(n_{Nj}^2 + n_{Ej}^2 + n_{Zj}^2)} \quad (2)$$

$$D_2 = D_2([S_{Zj}^2,\ S_{Nj}^2 + S_{Ej}^2];\ [n_{Zk}^2,\ n_{Nk}^2 + n_{Ek}^2]),\ j = 1,..,\ M_s;\ k = 1,..,\ M_n \quad (3)$$

where R_1 is the "rectified" STA/LTA detector, R_2 is the conventional STA/LTA detector and D_2 is two-sample Kolmogorov-Smirnov test statistics. In our modeling we set $\beta = 0.1$ which means that 10% of seismic signals with SNR_{thres} will be missed. Fig. 4 shows the results of modeling.

Kolmogorov-Smirnov detectors

Above we have demonstrated that STA/LTA is the best detector test statistics while detector input is free of spikes and noise p.d.f. close to Gaussian. We concentrate further more on K-S detectors because they are not conventional in seismology and otherwise exhibit some useful features. For K-S detectors we summarize from above:

- by direct numerical simulation we have shown that detector performances for all 4 combinations of 3C seismometers are azimuth independent. This is not obvious because of non-linear mixture of noise and signal but demonstrated by simulations to be valid.
- sensor combination iv) being subject to one-sample K-S test shows poorer performances than iii) being subject to two-samples K-S test. In the latter case vertical Z^2 and horizontal $(N^2 + E^2)$ form two-dimensional distribution while $(Z^2 + N^2 + E^2)$ is a one-dimensional random variable;
- test statistics (3) and statistics

$$D_1^i = D_1(S_{Zj}^2;\ n_{Zk}^2),\ j = 1,..,\ M_s;\ k = 1,..,\ M_n \quad (4)$$

were investigated with respect to P-wave angle of incidence. Results obtained (Fig. 5) clearly show

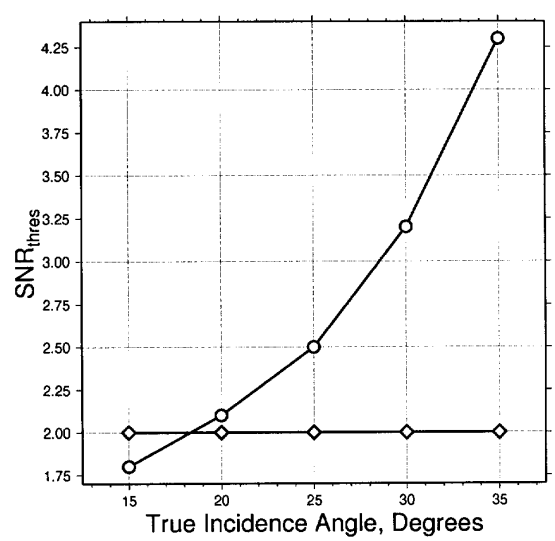

Fig. 5. SNR dependence on the "true" incidence angle. Diamonds represent $D_2(Z^2,\ N^2 + E^2)$, circles - $D_1(Z^2)$. For small incidence angles $D_1(Z^2)$ has better performance but much poorer for large incidence angles.