Robust automatic P-phase picking: An on-line implementation in the analysis of broad-band seismogram recordings

Reinoud Sleeman and Torild van Eck

Royal Netherlands Meteorological Institute (KNMI)
Wilhelminalaan 10, 3732 GK, De Bilt - The Netherlands

tel. +31.30.2206343 fax. +31.30.2201364
email : sleeman@knmi.nl , vaneck@knmi.nl

February 28, 1999

Published in: Physics of the Earth and Planetary Interiors, 113 (1999) 265-275

Abstract

The onset of a seismic signal is determined through joint AR modeling of the noise and the seismic signal, and the application of the Akaike Information Criterion (AIC) using the onset time as parameter. This so-called AR-AIC phase picker has been tested successfully and implemented on the Z-component of the broad-band station HGN to provide automatic P-phase picks for a rapid warning system. The AR-AIC picker is shown to provide accurate and robust automatic picks on a large experimental data-base. Out of 1109 P-phase onsets with signal-to-noise ratio (SNR) above 1 from local, regional and teleseismic earthquakes, our implementation detects 71% and gives a mean difference with manual picks of 0.1 sec. An optimal version of the well established picker of Baer and Kradolfer (1987) detects less than 41% and gives a mean difference with manual picks of 0.3 sec using the same data-set.

Keywords: phase-picker, AR model, AIC, seismic warning system

1 Introduction

Seismic warning systems are dependent on automatic, quasi real-time, procedures for detecting, onset picking and identifying of signal phases in seismogram recordings. For systems which release alert messages quickly after an earthquake, the accuracy of the phase picks is essential for automatic and reliable hypocenter determinations. Today, such systems use arrival times from many stations operated by different seismological institutions, using different picker algorithms. In this context, it is important to evaluate different phase pickers in terms of accuracy of the onset times, using identical data-sets. This paper presents a performance test of a phase picker which has been implemented at the Dutch broad-band station HGN in its role as as contributor of automatic phase onset times to the seismic warning system at the European-Mediterranean Seismological Center (EMSC).

A variety of methods has been developed in the past to detect and pick seismic phases and estimate their onset time. Detectors are considered to be processes for detecting the presence of seismic phases, whereas pickers are considered to be processes for estimating accurately the onset time of such phases. An extensive review of single component seismic detectors and pickers based on the comparison of short term average (STA) and long term average (LTA) of the trace is given by Berger and Sax (1980). This review, however, was not carried out uniformly on the same or even similar data-sets. Allen (1978, 1982) used the trace amplitude and the time derivative of the trace to generate a characteristic function which must be compared to some threshold value. Baer and Kradolfer (1987) adopted this idea to construct a modified amplitude envelope function. Furthermore, they incorporated a dynamic threshold value. Morita and Hamaguchi (1984) used a statistical adaptive algorithm to estimate the onset time using 1-component data. The same approach using 3-component data was taken by Pisarenko et al. (1987), Takanami and Kitagawa (1991), Tarvainen (1992) and Kvaerna (1994). Ruud and Husebye (1992) combined signal polarization and STA/LTA to devise a 3-component phase identifier, mainly designed for P waves. Detectors and pickers, specifically designed for S waves and Rayleigh waves, have been suggested by Cichowicz (1993) and Chael (1997) respectively.

In this study we follow Morita and Hamaguchi (1984) who modeled a time series as a multiple AR process, using the Akaike Information Criterion (AIC). Hereafter, we will refer to this method as the AR-AIC method, as in GSE/JAPAN/40 (1992). The method is applied on a large data-set of P-phases and we compare the automatic picks with the manual picks. Also, we compare the AR-AIC method with the picker proposed by Baer and Kradolfer (1987), hereafter called the BK picker, using the same data-set for this purpose.

2 Multiple AR modeling approach

The work of Morita and Hamaguchi (1984) is mainly based on Kitagawa and Akaike (1978), in which a time series is divided into locally stationary segments each modeled as an AR process. The technique assumes that we have (i) a time series $x_n = \{x_1, ..., x_N\}$ which includes the onset of a seismic signal, and (ii) a first estimate of the onset time. The intervals before and after the onset time are assumed to be two different stationary time series. In this paper we work with P-phases only, but the technique can be applied to other phases as well, as long as an estimate of signal onset time is available.

In each interval i=1,2, the one preceding and the one including the phase onset, we define a model window in which we fit the data to an autoregressive model of order M with coefficients a^i_m (m=1,...,M) :

$$x_t = \sum_{m=1}^{M} a^i_m x_{t-m} + e^i_t \tag{1}$$

with t=1,...,M for interval 1 and t=N-M+1,..,N for interval 2. The model divides the time series within a model window into a deterministic and a non-deterministic part. The non-deterministic time series e^i_n, or noise, is supposed to be Gaussian, with mean $E\{e^i_n\} = 0$, variance $E\{(e^i_n)^2\} = \sigma^2_i$ and uncorrelated with the deterministic part of the time series:

$E\{e_n^i x_{t-m}^i\} = 0.$

We use AR coefficients a_m^i to extract the non-deterministic part of the time series in intervals [M+1,K] and [K+1,N-M] using [1], where K is the division point. As we assume the non-deterministic parts to be Gaussian, we can express the approximate likelihood function $\mathcal{L}$ for the two non-deterministic time series in intervals [M+1,K] and [K+1,N-M] as:

$$\mathcal{L}(x; K, M, \Theta_1, \Theta_2) = \prod_{i=1}^{2} \left(\frac{1}{\sigma_i^2 2\pi}\right)^{\frac{n_i}{2}} exp\left(-\frac{1}{2\sigma_i^2} \sum_{j=p_i}^{q_i} \left(x_j - \sum_{m=1}^{M} a_m^i x_{j-m}\right)^2\right) \quad (2)$$

where $\Theta_i = \Theta(a_1^i, ..., a_M^i, \sigma_i^2)$ represents the model parameters for interval i (σ_i^2 is depending on K), and p_1=M+1, p_2=K+1, q_1=K, q_2=N-M, n_1=K-M, n_2=N-M-K.

Taking the logarithm of [2] and searching for the maximum likelihood estimation of the model parameters we get:

$$\frac{\partial log(\mathcal{L}(x; K, M, \Theta_i))}{\partial \Theta_i} = 0 \quad (3)$$

which has the solution:

$$\sigma_{i,max}^2 = \frac{1}{n_i} \sum_{j=p_i}^{q_i} \left(x_j - \sum_{m=1}^{M} a_m^i x_{j-m}\right)^2 \quad (4)$$

The maximum of the logarithmic likelihood function for the two models as function of K becomes:

$$log(\mathcal{L}(x; K, M, \Theta_1, \Theta_2)) = -\frac{1}{2}(K-M)log(\sigma_{1,max}^2) - \frac{1}{2}(N-M-K)log(\sigma_{2,max}^2) + C_1 \quad (5)$$

where C_1 is a constant.
This expression is the 'core' of Akaike's Information Criterion (AIC), which is defined as: AIC = - 2 log (maximized likelihood function) + 2 P, where P is the number of independently estimated parameters and equals AR order M in the model. Originally this function was designed to determine the order of the AR process in [1]. The first term indicates the badness of the model fit and the second term the unreliability (Akaike, 1974). In our application we have fixed the order to M, therefore this function is a measure for the model fit. Point K where the joint likelihood function [5] is maximized, or AIC is minimized, determines the optimal separation of the two stationary time series. This division point leads to the best fit for both models in the least squares sense, and is interpreted as the phase onset.

The AIC of the 2-interval model is given as function of merging point K:

$$AIC(K) = (K-M)log(\sigma_{1,max}^2) + (N-M-K)log(\sigma_{2,max}^2) + C_2 \quad (6)$$

where C_2 is a constant.

3 Implementation

In the framework of the EC funded project 'A rapid warning system for earthquakes in the European-Mediterranean region' the Royal Netherlands Meteorological Institute (KNMI) has implemented the AR-AIC processing scheme for real-time analysis of the vertical component of station HGN, as to provide in near real-time automatic P-phase onset picks for transmission to the EMSC data center near Paris, France. The goal of this project is to release rapid and accurate locations for earthquakes greater than magnitude 5.0 occurring in the European-Mediterranean region, through cooperation between a number of national and regional networks in this region. In order for the KNMI to do so, we have interfaced the AR-AIC picker to our Automatic Data Request Manager (AutoDRM; Kradolfer, 1993) in such a way that automatic onset times are provided for an AutoDRM request. In case of an earthquake trigger EMSC sends a request to our AutoDRM to extract the P-phase arrival time from a small time window around the expected P-phase arrival time. This request initiates a two-step procedure: a rough estimation of the onset time through the STA/LTA detector, followed by an accurate onset time estimation through the picker. This scenario presumes a reliable detector. Extensive experience has shown that the presently implemented STA/LTA detector displays a good performance, so we focussed our attention to an accurate implementation of the phase picker.

Station HGN is equipped with STS-1 seismometers (3-components) operating in the very broad-band mode and with 40 samples/second. A bandpass filter (0.6 - 6.0 Hz) is applied on the vertical component, followed by a STA/LTA detector (STA = 1 second, LTA = 30 second and threshold 8). In this configuration we obtained 3517 detections in 1996 (Figure 1). Out of these 2807 (80%) were confirmed by the analyst as relevant seismic phases. 32% of the detections were identified as initial P-phases, 16% as secondary P-phases (PP etc) and 32% as other phases. False alarms (detections without phase identification) and missed detections (identified phases without being detected) are not part of this study.

The AR-AIC picker operates on high-pass filtered (2.0 - 10.0 Hz) traces. It uses a noise and a signal window of length 4 seconds and searches for the minimum in [6] in a time window of length 12 seconds starting 8 seconds before the initial STA/LTA detection. Presently, the AR-AIC picker is running on a routine basis at station HGN since September 1997 and its performance is monitored in terms of differences between the automatic picks and manual picks.

4 Performance tests

The presented picker algorithm was tested on STA/LTA detections at station HGN in 1996, and limited to those which were confirmed manually as first P-phase onsets. Next, we compared the AR-AIC picker performance with those of the BK picker.

Manual picks versus automatic picks
The picker performance has been studied using a random search through the possible parameter settings and testing this on a large data-set for which we have assumedly correct picks, i.e. the above mentioned first (1109) P-phase onsets.
Firstly we determined the time window, relative to the STA/LTA detection, in which to search for the phase onset, i.e. the search-window. Figure 2a compares the HGN detections with the manual picks as a function of signal-to-noise ratio (SNR) for EMSC reported events in 1996. The maximum time difference between the STA/LTA detections and manual readings is about 4.5 seconds. Figure 2b shows the detector performance for all detections in 1996 identified as initial P-phases (32% of all detections). As expected they display, with one exception, positive delays with a maximum of 5 seconds. Consequently, we estimated a search-window of 5 seconds as reasonable.

Secondly, we determined which parameters to consider within the test. The most important ones of the AR-AIC picker are: (i) The AR model order, (ii) the parameters defining the

AR model windows, and (iii) the signal enhancement prior to applying the picker algorithm. Preliminary performance tests showed a preference in our application for AR model of order 8. This has later been corroborated using optimal parameter settings as discussed below. The AR model windows were defined by three parameters; W, O1 and O2 (Figure 3). With respect to signal enhancement we considered a number of alternative bandpass filters or filter combinations (Table 1).

Thirdly, we defined an empirical distribution enabling us to compare the estimated picks of one test with those of another test and assumed the manual picks to be the 'true' ones. The parameter of concern is the time difference between the estimated and 'true' pick. The weighted number of differences are plotted within a moving time window (Figure 4). The weight of each pick is the SNR, where the signal and noise samples were chosen as indicated in Figure 3. Signals with SNR < 0 Db have been disregarded as being insignificant or non-real events. This weighting procedure corresponds to our notion that correct picking of 'clear' onsets is more important than correct picking of 'less distinct' onsets, as clearly illustrated in Figure 4. This experimental distribution provides an estimate of the mean value and the 90% confidence interval. A representative summary of our tests is presented in Tables 2 and 3a, which show a preference for a high frequency bandpass filter (2 - 10 Hz) and the following window settings: W = 4 seconds, O1 = 4 seconds and O2 = 0 seconds.

The manual picks, defined as those picked by the analyst in his/her routine analysis, have been assumed as being the correct ones. In order to check the validity of this assumption, we scrutinized the 'best' test of Table 3a in more detail, i.e. especially those picks with SNR > 3 and time differences larger than 0.5 seconds between automatic and manual picks. Surprisingly, about 35 onsets could be considered as mis-picks by the analyst, with deviations ranging from 0.1 to 4.0 seconds, the larger errors confined to low SNR onsets only. The example in Figure 5a clearly shows Pn, however, Pg had been identified as first P-phase onset by the analyst. A probable explanation for these mis-picks is the 'incorrect' application of bandpass filters in the routine analysis procedure. The same tests were applied to the data using the scrutinized and corrected manual picks as reference data-base (revised data-base) and presented in Table 3b.

However, we also found correct manual picks that were picked incorrectly by the automatic picker. Figure 5b demonstrates, for example, a P-phase onset with a clear low frequency onset. Obviously, the selected high frequency bandpass filter suppresses this and enhances the high frequency onset. As yet, we have not been able to find a parameter/filter combination that succeeds in optimizing picking both low and high frequency phase onsets.

The robustness of the AR-AIC picker has also been tested using different combinations of data-sets, initial criteria and additional conditions (Table 3c). Replacing the original data-set with the revised data-base results in only small differences. Likewise, replacing STA/LTA detections with 'true' picks results in only negligible differences. In applying the automatic picker to submit picks to automated location procedures a 90% confidence interval of 1.43 seconds is too high. An alternative approach is then to eliminate weak, unreliable onsets. This we obtained by adding an additional detection criterion: A detection T0 is declared if STA/LTA > 8, however a reliable detection T0 is declared if STA/LTA > 15 within 5 seconds of T0. This additional condition does indeed improve performance with regard to the spread (90% confidence interval reduces to 1.27 seconds), however at the price of a reduced number of picks from 71% to 39%.

Finally, we investigated once more our AR model order assumption, using the preferred parameter setting as obtained above. From Table 3d we find that higher order AR models tend to provide better mean onset times. Order 8 results in a minimum in the 90% confidence interval.

AR-AIC picker versus BK picker
The BK picker, which is successfully applied in the Swiss National Network, has been implemented with a similar purpose as our AR-AIC implementation. Therefore a comparison seems natural. Details of the procedure can be found in Baer and Kradolfer (1987), in which they essentially modify the amplitude envelope function of the time series, as described by Allen (1978, 1982), to calculate a characteristic function CF to be compared to a dynamic threshold value. The envelope function time series e_t is described by:

$$e_i^2 = x_i^2 + \dot{x}_i^2 \left(\frac{\sum_{j=0}^{i} x_j^2}{\sum_{j=0}^{i} \dot{x}_j^2} \right) \tag{7}$$

where $\dot{x}$ denotes the time derivative of x. They introduced the characteristic function CF:

$$CF_i = \frac{e_i^4 - \overline{e_i^4}}{\sigma^2} \tag{8}$$

in which σ^2 denotes the variance of e_t^4 and the bar indicates the average.

They implemented the picker on an optimum bandpass filter, based on stacked amplitude spectra from several traces. We implemented BK on a single trace. As the BK picker is essentially a modified amplitude envelope function it is mainly applicable to band-limited signals. Consequently, we tested the BK picker using bandpass filters 2 and 5 in Table 1. Three parameters must be defined in this algorithm (Figure 6). A threshold or detection level DL, and two time intervals (T1, T2) related to the duration of exceeding this threshold. CF must exceed DL for at least T1 samples, but is given a tolerance window of T2 samples in which it may drop below DL.

The performances of the two pickers are compared using the revised data-base of P-phases in station HGN. For this we have aimed at an optimal tuned BK picker. Table 4 summarizes the performance tests for the BK picker. From this we see that varying the parameters both influences the accuracy and the number of the picks. However, we did not succeed in any test to recover more than 41% of all P-phases. This is in contrast to the AR-AIC picker where we picked more than 71% of the P-phases. AR-AIC performs better than BK for this data-set in terms of the 90% confidence interval and the number of picks. Also BK in general picks the onset times too late, as reflected by the mean deviation of about 0.3 seconds for BK and 0.1 seconds for AR-AIC.
In conclusion we find that the proposed AR-AIC implementation performs better than BK for the purpose of fast automatic phase picking in single traces.

5 Discussion and conclusions

We have presented the AR-AIC picker as a tool for single station automatic P-phase picking. The picker has been tuned and implemented within the framework of the EMSC rapid

determination of epicenters. In our implementation more than 70% of the manual picks were automatically estimated within a 90% confidence interval of 1.6 seconds and a mean of about 0.1 seconds when compared with the manual picks. Individual visual comparisons of the automatic picks with the manual onsets showed that a number of manual picks were actually improved by the automatic ones, indicating that AR-AIC is a very useful tool to improve the accuracy in manual picking.
AR-AIC has been compared to the modified and improved picker by Allen (1978, 1982) as proposed and implemented successfully by Baer and Kradolfer (1987). For this we have made a random search in the parameter space to find an optimal setting with respect to our data-set. From this comparison we believe that AR-AIC is most appropriate for our purpose, using one channel only.
We consider AR-AIC as an iterative improvement of picking the phase onset. It relies on a good performance of the detector with respect to detecting a P-phase, but is not very sensitive to the accuracy of the detector. In our case the detector behaves satisfactory, while AR-AIC is sufficiently robust to handle inaccurate detection times.
However, to quantify the reliability of the onset remains a difficult problem. SNR is normally used to quantify the quality of a signal and can be improved in a single component seismic trace by signal enhancement using, for example, bandpass filters or Wiener filters (Douglas, 1997). In our application we use SNR as a measure to quantify the accuracy of the automatic pick. Ideally, the assigned measure is large for high accuracy onset estimates and low for unreliable estimates. A problem, however, is that an inaccurate onset time leads to incorrect windows, not representing the true noise and phase, to calculate SNR. We have chosen to calculate SNR in fixed windows relative to the detector onset time (Figure 3) which is satisfactory for our implementation.
In summary, the AR-AIC picker appears to be a useful tool for manual analysis and assigning automatically onset times to detected seismic signals for the purpose of rapid epicenter calculations. Provided, however, that we apply high frequency filtering to the broad-band data. Moreover the AR-AIC picker displays a robust behaviour as relative large errors in the detection times do not influence the automatic picks significantly. Quantifying the reliability of the onsets in terms of SNR remains difficult, although from a practical point of view using the SNR as a measure for the quality is sufficient in our implementation. The current AR-AIC picker is implemented and operating at HGN, and used regularly by the EMSC, while its performance is monitored.

Acknowledgements

We wish to thank JMA for making available the software to us which initiated this study. This work was performed at the Royal Netherlands Meteorological Institute and was supported by the European Commission, Directorate General XII for Science, Research and Development under contract ENV4-CT96-0282.

References

Akaike, H., 1974. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. Ann. Inst. Stat. Math., 26:363-387.

Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces. Bull. Seism. Soc. Am., 68:1521-1531.

Allen, R.V., 1982. Automatic phase pickers: their present use and future prospects. Bull.

Seism. Soc. Am., 72:S225-S242.

Baer, M. and Kradolfer, U., 1987. An automatic phase picker for local and teleseismic events. Bull. Seism. Soc. Am., 77:1437-1445.

Berger, J. and Sax, R.L., 1980. Seismic detectors: the state of the art. Technical report, SSR-R-80-4588.

Chael, E.P., 1997. An automated Rayleigh-wave detection algorithm. Bull. Seism. Soc. Am., 87:157-163.

Cichowicz, A., 1993. An automatic S-phase picker. Bull. Seism. Soc. Am., 83:180-189.

Douglas, A., 1997. Bandpass filtering to reduce noise on seismograms: is there a better way ?. Bull. Seism. Soc. Am., 87:770-777.

GSE/JAPAN/40, 1992. A fully automated method for determining the arrival times of seismic waves and its application to an on-line processing system. Paper tabled in the 34th GSE session in Geneva, July 1992.

Kitagawa, G. and Akaike, H., 1978. A procedure for the modeling of non-stationary time series. Ann. Inst. Stat. Math., 30, Part B:351-363.

Kradolfer, U., 1993. Automating the exchange of earthquake information. EOS, Transactions, AGU, V74, N39:442.

Kvaerna, T., 1994. Accurate determination of phase arrival times using autoregressive likelihood estimation. Annali di Geofisica, Vol. XXXVII, N.3:287-300.

Morita, Y. and Hamaguchi, H., 1984. Automatic detection of onset time of seismic waves and its confidence interval using the autoregressive model fitting. Zisin, 37:281-293.

Pisarenko, V.F., Kushnir, A.F. and Savin, I.V., 1987. Statistical adaptive algorithms for estimation of onset moments of seismic phases. Phys. Earth Planet. Inter., 47:4-10.

Ruud, B.O. and Husebye, E.S., 1992. A new three-component detector and automatic single station bulletin production. Bull. Seism. Soc. Am., 82:221-237.

Takanami, T. and Kitagawa G., 1991. Estimation of the arrival times of seismic waves by multivariate time series model. Ann. Inst. Stat. Math., 43:407-433.

Tarvainen, M., 1992. Automatic seismogram analysis: statistical phase picking and locating methods using one-station three-component data. Bull. Seism. Soc. Am., 82:860-869.

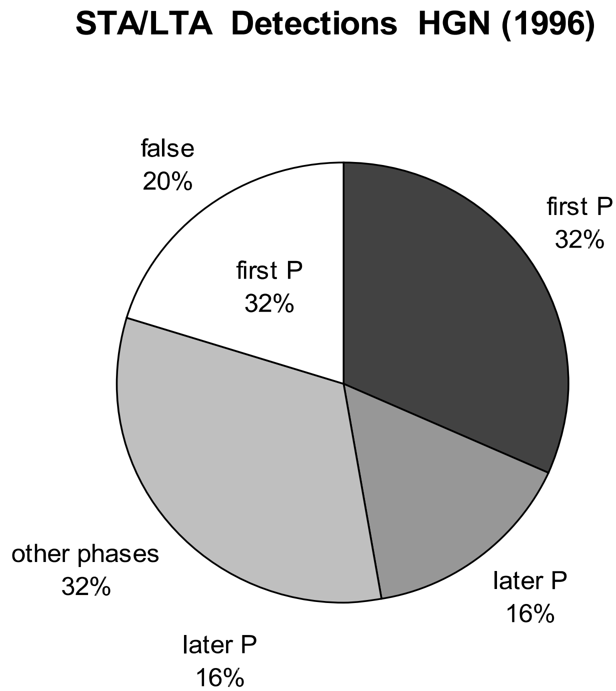


Figure 1: *Identification of detections from the STA/LTA detector on the vertical component of station HGN during 1996. A total of 3517 detections were obtained of which 80% were confirmed by the analyst, and 20% were false detections. 32% of the detections were identified as initial P-phases, 16% as subsequent P-phases and 33% as other phases.*

Table 1

Filters used for signal enhancement

Filter	Description
1	None: use raw data
2	Bandpass filter which maximizes SNR: 0.02 - 0.2 Hz 0.1 - 1.0 Hz 0.5 - 5.0 Hz 1.0 - 10.0 Hz 2.0 - 10.0 Hz
3	Bandpass filter: 0.5 - 5.0 Hz
4	Bandpass filter: 1.0 - 10.0 Hz
5	Bandpass filter: 2.0 - 10.0 Hz

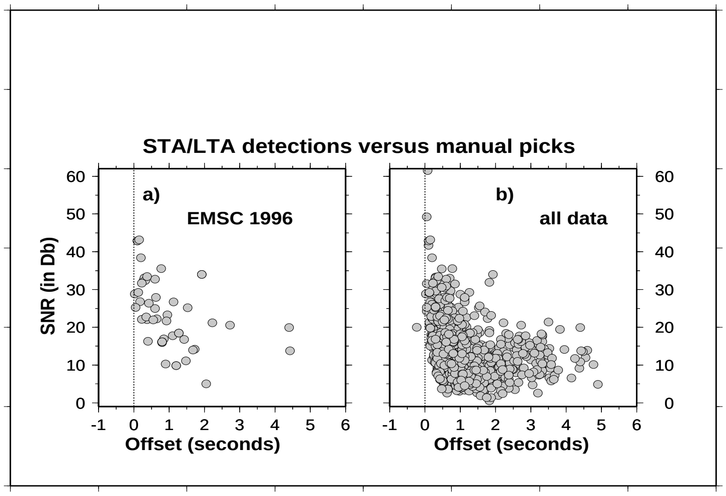


Figure 2: *STA/LTA detector performance at HGN as compared with the manual picks as a function of signal to noise ratio (SNR) for (a) EMSC reported events in 1996 and (b) all identified initial P-phases in 1996.*

Table 2

Results of AR-AIC method in combination with filters in Table 1

W = 5; O1 = 4; O2 = 0

Filter	90% Confidence interval	Mean	# Picks
5	1.729	0.099	783
4	2.170	0.080	860
1	2.948	-0.047	645
2	3.199	0.068	796
3	3.265	0.074	867

Table 3a

AR-AIC test results using original manual picks and filter 5

W	O1	O2	90% Confidence interval	Mean	# Picks
4	4	0	1.646	0.099	783
4	4	1	1.658	0.096	783
4	5	0	1.728	0.098	783
5	4	0	1.729	0.099	783
3	4	1	1.733	0.098	783
5	5	0	1.741	0.098	783
4	5	1	1.768	0.091	783
5	3	0	1.775	0.110	783
5	3	1	1.790	0.105	783
2	4	1	1.796	0.105	783
5	4	1	1.799	0.090	783
3	4	0	1.815	0.103	783
5	5	1	1.823	0.089	783
4	3	0	1.839	0.115	783
2	4	0	1.869	0.105	783
3	5	0	1.874	0.096	783
3	3	0	1.896	0.119	783
4	3	1	1.953	0.118	783
3	5	1	1.967	0.089	783
3	3	1	1.969	0.124	783
2	3	0	2.040	0.123	783
2	5	1	2.098	0.096	783
2	5	0	2.148	0.095	783
2	3	1	2.155	0.134	783
5	2	0	2.538	0.141	783
4	2	0	2.656	0.151	783
3	2	0	2.763	0.167	783
5	2	1	2.784	0.138	783
2	2	0	2.845	0.162	783
4	2	1	2.963	0.143	783
3	2	1	3.006	0.166	783
2	2	1	3.108	0.172	783

Table 3b

AR-AIC test results using revised data-base and filter 5

W	O1	O2	90% Confidence interval	Mean	# Picks
4	4	0	1.4303	0.0992	783
4	4	1	1.4412	0.0942	783
5	4	0	1.5080	0.0989	783
3	4	1	1.5390	0.0977	783
4	5	0	1.5420	0.0986	783
5	5	0	1.5503	0.0974	783
5	4	1	1.5582	0.0895	783
3	5	0	1.5662	0.0966	783
4	5	1	1.5725	0.0913	783
5	3	0	1.5877	0.1098	783
3	4	0	1.6020	0.1027	783
5	3	1	1.6020	0.1043	783
5	5	1	1.6071	0.0886	783
3	5	1	1.6309	0.0890	783
4	3	0	1.6421	0.1138	783
2	4	1	1.6626	0.1073	783
2	4	0	1.6706	0.1048	783
4	3	1	1.7345	0.1162	783
3	3	0	1.7510	0.1198	783
2	5	1	1.7835	0.0939	783
3	3	1	1.7932	0.1232	783
2	3	0	1.8326	0.1231	783
2	3	1	1.9546	0.1337	783
2	5	0	1.9633	0.0974	783
5	2	0	2.3985	0.1412	783
4	2	0	2.5111	0.1525	783
3	2	0	2.6193	0.1686	783
5	2	1	2.6265	0.1387	783
2	2	0	2.7014	0.1604	783
4	2	1	2.8126	0.1452	783
3	2	1	2.8516	0.1668	783
2	2	1	2.9993	0.1720	783

Table 3c

Results of the 'best' test in Table 3b using different detections and reference phases

Filter 5; W = 4; O1 = 4; O2 = 0

90% Confidence interval	Mean	# Picks	Detection	Reference phases
1.646	0.099	783	STA-LTA	Original data-base
1.430	0.099	783	STA-LTA	Revised data-base
1.406	0.079	805	Manual picks	Revised data-base
1.272	0.130	427	STA-LTA + additional criterion	Revised data-base

Table 3d

Influence of AR order (2-10)

using the 'best' test settings of table 3c

90% Confidence interval	Mean	AR order
1.272	0.130	8
1.320	0.159	6
1.348	0.113	10
1.372	0.125	9
1.423	0.162	4
1.434	0.179	5
1.443	0.149	7
1.470	0.216	2
1.811	0.223	3

Table 4a

BK test results using test data-base and filter 2

DL	T1	T2	90% Confidence interval	Mean	# Picks
60	20	8	1.584	0.296	108
60	5	2	1.819	0.355	146
60	10	4	1.975	0.347	126
8	20	8	1.992	0.330	448
8	10	4	1.996	0.330	449
30	20	8	2.042	0.378	252
15	5	2	2.108	0.420	444
30	5	2	2.109	0.453	337
8	5	2	2.140	0.300	449
30	10	4	2.160	0.412	276
15	10	4	2.233	0.431	441
15	20	8	2.702	0.412	427
4	10	2	3.253	0.215	451
4	4	2	5.239	-0.114	451
4	2	1	5.572	-0.667	451

Table 4b

BK test results using test data-base and filter 5

DL	T1	T2	90% Confidence interval	Mean	# Picks
30	20	8	1.548	0.279	148
60	5	2	1.560	0.319	136
60	20	8	1.674	0.281	97
60	10	4	1.918	0.330	116
15	20	8	2.085	0.269	209
30	10	4	2.086	0.301	164
30	5	2	2.200	0.334	192
15	10	4	2.271	0.284	225
15	5	2	2.485	0.315	262
8	20	8	2.498	0.250	279
8	10	4	2.622	0.273	296
8	5	2	2.663	0.266	351
4	10	2	3.518	0.234	372
4	2	1	5.568	-1.868	426
4	4	2	5.689	-0.125	425

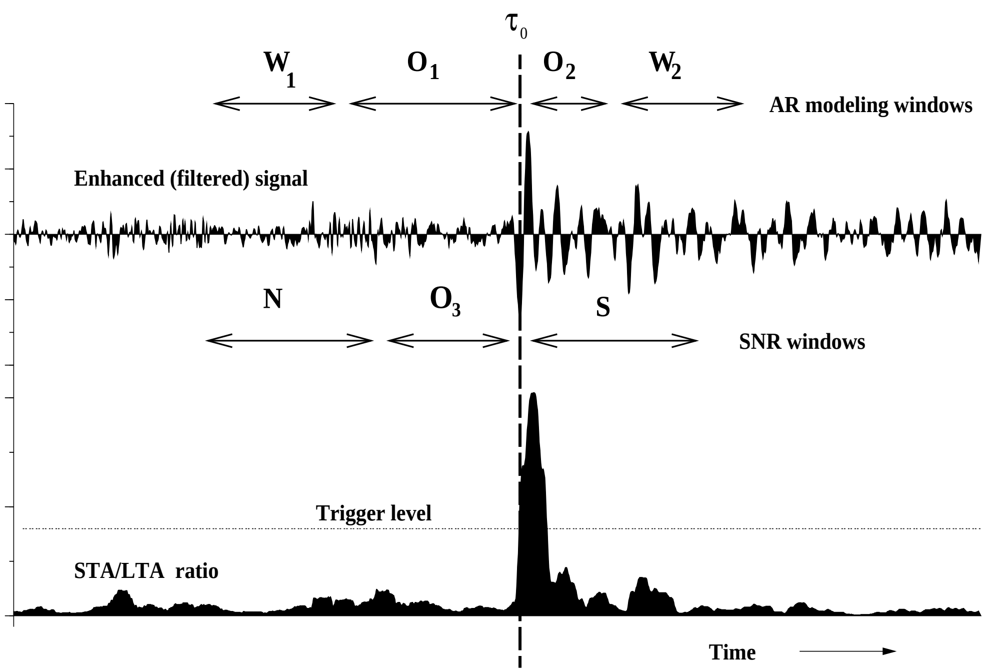


Figure 3: *The window definitions as used in modeling within the AR-AIC picker. The upper trace shows the enhanced (filtered) signal on which the AR-AIC picker is operating, the lower trace shows the STA/LTA ratio and its detection pick. Note the delay of the detection of about 1/4 period with respect to the 'true' onset. Definition of parameters W = W1 = W2, O1 and O2 as used in the AR modeling: W1 indicates the window used for modeling the noise AR model, W2 indicates the window used for modeling the seismic signal AR model, O1 and O2 are the relative time offsets from the initial detection. Noise window length N at offset O3 from the detection, and signal window S are used for SNR calculation.*

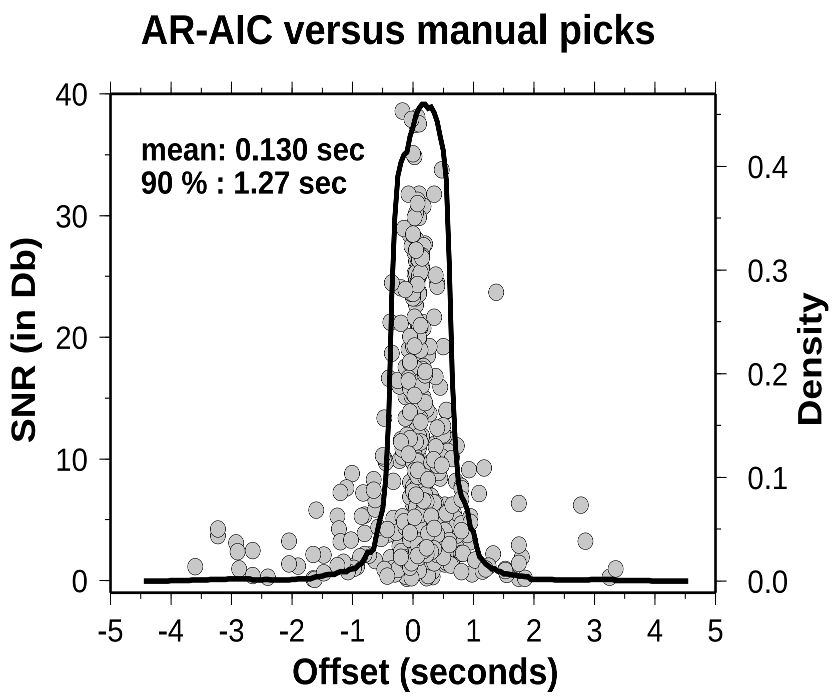
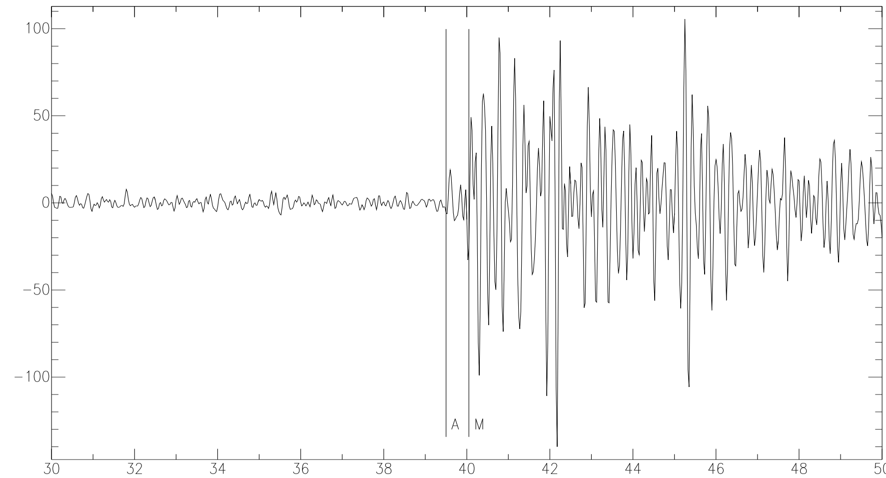


Figure 4: *An example of the differences (offset) between the automatic picks (circles) and the 'true' picks and its corresponding empirical distribution (thick line). Only picks with SNR > 0 Db have been considered. The mean and 90% confidence interval are indicated in the graph.*



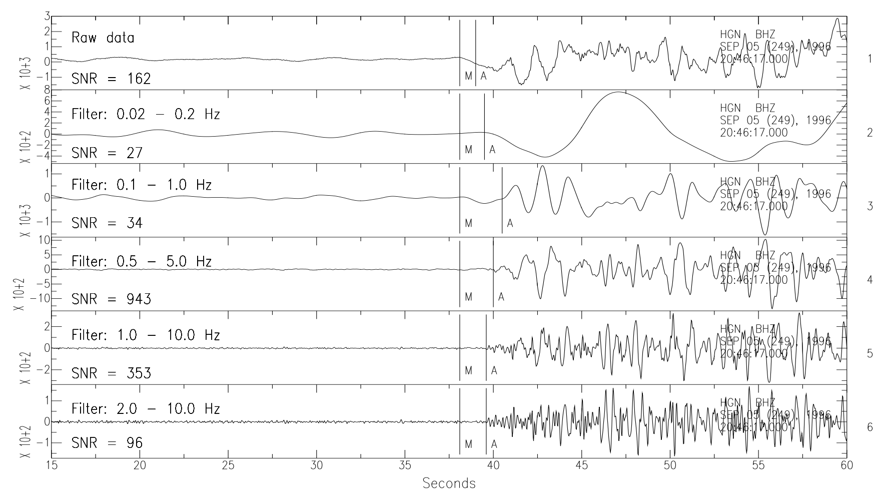


Figure 5: *Examples of the performance of the automatic picker as compared to the manual picks. (a) Improved pick using the AR-AIC picker. The trace is filtered between 2 and 10 Hz and indicates the initial identified P-phase (Pg) by the analyst (M) as well as the automatic pick (A) of Pn. About 1% of the manual picks of the identified initial P-phases could be improved using AR-AIC. (b) Poor performance of the AR-AIC picker on a low frequency onset as picked by the analyst (M). The automatic pick (A) could not be improved using different bandpass filters. This example also demonstrates the weakness of using SNR in filter 2 (Table 1).*

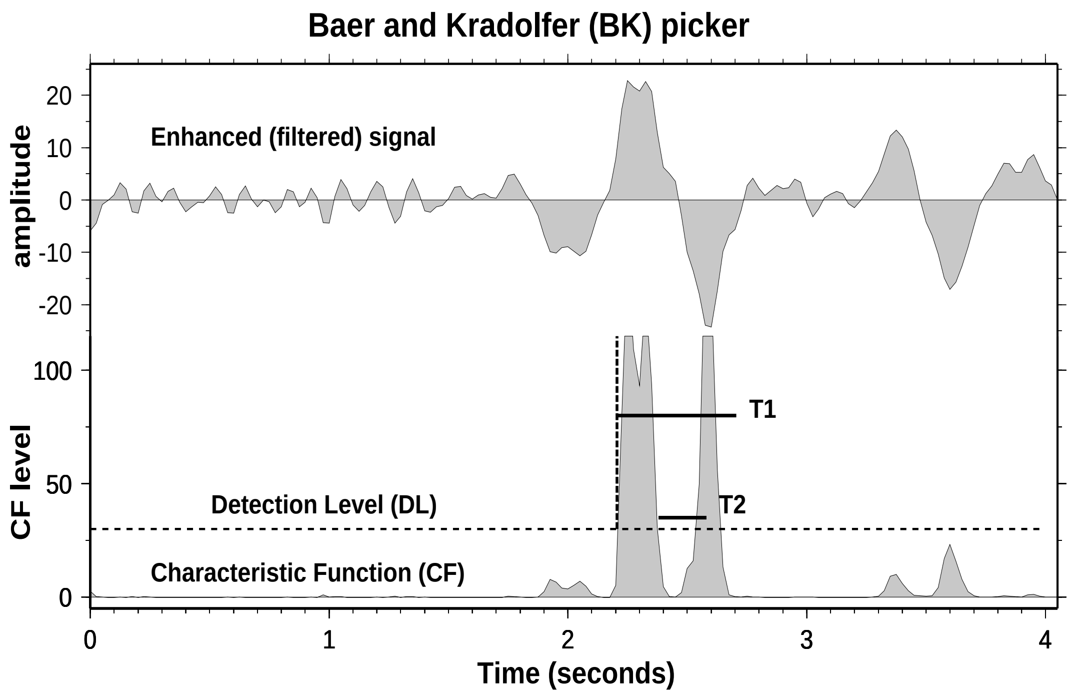


Figure 6: *Some basic parameters for the BK picker. Upper trace shows the enhanced (filtered) trace, and the lower trace the corresponding characteristic function (CF). Indicated are the detection level (DL), the time window (T1) within which CF should exceed DL and a second time window (T2) in which DL may drop temporarily below DL. Consequently, the signal above is declared a detection with the indicated onset time. Note the slight delay introduced.*