



## GLM-based automatic picking of waveforms

Giada Adelfio (1), Marcello Chiodi (1), Antonino D'Alessandro (2), and Dario Luzio (1)

(2) Istituto Nazionale di Geofisica e Vulcanologia, Italy (antonino.dalessandro@ingv.it), (1) Università delgi Studi di Palermo

Large high-quality and data sets are required for accurate earthquake location, to better define seismogenic volumes and enhance the resolution of crustal and upper mantle structures in seismic tomography. Large data sets are generally derived from routine hand-picking and are certainly affected by inconsistencies and blunders. Manual picking of the arrival times of the major body waves phases on full seismograms is a very time consuming activity, practically unusable for these studies, where typical dat sets consists of thousands to tens of thousands of seismograms.

Technique for automatic detection of first-arrival on continuous seismic records have been studied for over thirty years, this being one key issue for improving the efficiency of automatic seismic network.

These are mainly based on STA/LTA algorithms applied to signal envelopes (Allen, 1978, 1982; Baer and Kradolfer). Recent picking techniques are based on the quantification of some attributes of the signal, instantaneous or related to a sliding time window like amplitude, frequency or polarization, and have been applied to single attributes or to smoothed combinations of them. Some authors (Morita and Hamaguchi, 1984; Takanami and Kitagawa, 1988; Kushnir et al., 1990; Takanami and Kitagawa, 2003) modeled seismic time series as a multiple AR process, using the Akaike Information Criterion (AIC). Other seismologist (Der and Shumway, 1999 ) have tested the performance of the CUSUM algorithm (Basseville and Nikiforov 1993) in the first arrivals picking. Approches based on Pattern Recognition (Klumpen and Joswig, 1993) and Neural Networks (Dai and MacBeth 1995, 1997), window threshold detection (Willis and Toksoz., 1983; Coppens, 1985) and variation in fractal dimension (Boschetti et al., 1996) have been also used.

We develop an automatic picking method based on the fit of a generalized linear regression model (GLM) for detecting change-points in the variance of heteroscedastic Gaussian variables, with piecewise constant variance function, associated to the sequence of the major seismic phases in a seismogram. In particular, we propose a breakpoint detection procedure for changes in variation assuming that the variance function can be described by a piecewise constant function with segments delimited by unknown change-points; moreover the method discard the spurious change-points on the basis of a generalized version of the BIC, extending the cumSeg procedure proposed by Muggeo and Adelfio (2011). Although there are many circumstances in which testing for change in variance is crucial, such as waveforms of earthquakes picking, the problem of variance change-point detection has not been widely considered in the literature, but for few papers focusing in autoregressive time-series models (Wichern et al., 1976; Wang and Wang, 2006; Zhao et al., 2010, e.g.), most neglecting the problem of multiple change-points in part because of the difficulty in handling computations.

The proposed simple and efficient method allows to detect and select unknown multiple change-points on the basis of a very basic GLM-based procedure. Simulations have shown good performance of the proposed approach, such as a sample size increases, mean squared errors of the detected number of change-points decrease. Finally, applications to real seismograms provide efficient and fast arrivals picking, detecting multiple change-points that are reasonably identifiable with particular source signal generation.

### References

- Allen, R., 1978. Automatic earthquake recognition and timing from single traces, *Bull. seism. Soc. Am.*, 68, 1521–1532.
- Allen, R., 1982. Automatic phase pickers: their present use and future prospects, *Bull. seismol. Soc. Am.*, 72, S225–S242.
- Basseville, M.&Nikiforov, I.V., 1993. *Detection of Abrupt Changes: Theory and Application*. Prentice Hall Information and System Science Series, Prentice Hall, Englewood Cliffs, New Jersey.
- Boschetti, F.. Dentith, M.D., and List, R. D., 1996, a fractal-based algorithm for detecting first arrivals on seismic traces, *Geophysics*, 61, 1095-1102.

- Baer, M. & Kradolfer, U., 1987. An Automatic phase picker for local and teleseismic events, *Bull. seism. Soc. Am.*, 77, 1437–1445.
- Coppens, F., 1985, First arrival picking on common-offset trace collection for automatic estimation of static corrections, *Geophysics Prosp.*, 33, 1212-1231.
- Dai, H. & MacBeth, C., 1995. Automatic picking of seismic arrivals in local earthquake data using an artificial neural network, *Geophys. J. Int.*, 120, 758–774.
- Dai, H. & MacBeth, C., 1997. The application of back-propagation neural network to automatic picking seismic arrivals from single-component recordings, *J. geophys. Res.*, 102, 15 105–15 115.
- Klumpen, E. & Joswig, M., 1993. Automated reevaluation of local earthquake data by application of generic polarization patterns for P- and S-waves, *Computers & Geosciences*, 19(2), 223–231.
- Morita, Y. & Hamaguchi, H., 1984. Automatic detection of onset time of seismic waves and its confidence interval using the autoregressive model fitting, *Zisin*, 37, 281–293.
- Muggeo, V. M. R. and Adelfio, G. (2011). Efficient change point detection for genomic sequences of continuous measurements. *Bioinformatics*.
- Takanami, T. & Kitagawa, G. (eds), 2003. *Methods and Applications of Signal Processing in Seismic Network Operations*. Springer, Berlin, *Lecture Notes in Earth Sciences*, 98.
- Wang, L. and Wang, J. (2006). Change-of-variance problem for linear processes with long memory. *Statistical Papers*, 47, 279-298.
- Willis, M.E., and Toksoz, M.N., 1983, Automatic P and S velocity determination from full waveform digital acoustic log, *Geophysics*, 48, 1631-1644.
- Wichern, D. W., Miller, R. B., and Hsu, D.-A. (1976). Changes of variance in first-order autoregressive time series models-with an application. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 25(3), 248-256.
- Zhao, W., Tian, Z., and Xia, Z. (2010). Ratio test for variance change point in linear process with long memory. *Statistical Papers*, 51(2), 397-407.