Geophys. J. Int. (2006) **165,** 121–134 doi: 10.1111/j.1365-246X.2005.02799.x

Automatic seismic phase picking and consistent observation error assessment: application to the Italian seismicity

R. Di Stefano,[1,2] F. Aldersons,[3] E. Kissling,[1] P. Baccheschi,[2] C. Chiarabba[2] and D. Giardini[1]

[1] *ETH, Geophysik, Zürich, Switzerland*
[2] *INGV, CNT, Rome, Italy. E-mail: raffaele.distefano@ingv.it*
[3] *Tel Aviv University, Tel Aviv, Israel*

Accepted 2005 September 7. Received 2005 August 15; in original form 2004 June 8

SUMMARY

Accuracy of seismic phase observation and consistency of timing error assessment define the quality of seismic waves arrival times. High-quality and large data sets are prerequisites for seismic tomography to enhance the resolution of crustal and upper mantle structures. In this paper we present the application of an automated picking system to some 600 000 seismograms of local earthquakes routinely recorded and archived by the Italian national seismic network. The system defines an observation weighting scheme calibrated with a hand-picked data subset and mimics the picking by an expert seismologist. The strength of this automatic picking is that once it is tuned for observation quality assessment, consistency of arrival times is strongly improved and errors are independent of the amount of data to be picked. The application to the Italian local seismicity documents that it is possible to automatically compile a precise, homogeneous and large data set of local earthquake *Pg* and *Pn* arrivals with related polarities. We demonstrate that such a data set is suitable for high-precision earthquake location, focal mechanism determination and high-resolution seismic tomography.

Key words: automatic picking system, event location, focal mechanism, high quality, seismic tomography.

1 INTRODUCTION

Body wave arrival times of high quality, consistency in data error assessment and uniformly dense sampling of the target volume are the three main prerequisites for high-resolution traveltime tomography. While density and uniformity of sampling strongly depend on source and receiver distributions and are easily assessed, consistency and observation quality of data are not. Routinely hand-picked data are human dependent and errors are difficult to estimate, since manual picking is performed by several seismologists over the years. Nonetheless, consistency in error estimation is of primary importance because the estimated average observation error (the noise level in the data) normally defines the termination of the iterative inversion process of seismic tomography and hypocentres calculation. Large data sets are generally derived from routine hand-picking and are certainly affected by inconsistencies and blunders. The lack of reasonably consistent error estimates severely reduces the resolving power of such data set.

Recent teleseismic tomography studies of the Italian region clearly show that the velocity models depend on a trade-off between quantity and quality of data. Large low-quality data sets allowed to improve the sampling of the target volume (Piromallo & Morelli 2003) at the cost of a low resolution of *Vp* images due to large data errors. On the other hand, small high-quality data sets (Lucente *et al.* 1999; Lippitsch *et al.* 2003) yield high-resolution tomographic images even with fewer rays, but only for limited regions. When targeting the whole shallow lithosphere–asthenosphere system beneath Italy, the uneven distribution of sources and receivers (Fig. 1) requires the use of all available data to achieve a sufficiently dense and homogeneous sampling. During 20 years of seismic monitoring, the Istituto Nazionale di Geofisica e Vulcanologia (INGV) located as much as 50 000 local and regional earthquakes with a total of more than 500 000 *P* and 250 000 *S* arrival times picked by several analysts. To improve the quality of such a data set, manual repicking is not a feasible option. Automated pickers, however, normally lack the possibility to consistently estimate individual observation errors on the arrival times they produce. This serious shortcoming explains, partially at least, why such procedures are still rarely used today in tomography and related studies. In this paper we present the results obtained by the use of a recently developed picking system called MannekenPix (Aldersons 2004[1]), capable of producing highly accurate first arrival *P* picks and related polarity determination. Furthermore, MannekenPix (MPX) includes a weighting mechanism rigorously calibrated on a series of reference picks with related

[1] Freddy Aldersons' PhD thesis is available at http://faldersons.net

© 2006 The Authors
Journal compilation © 2006 RAS

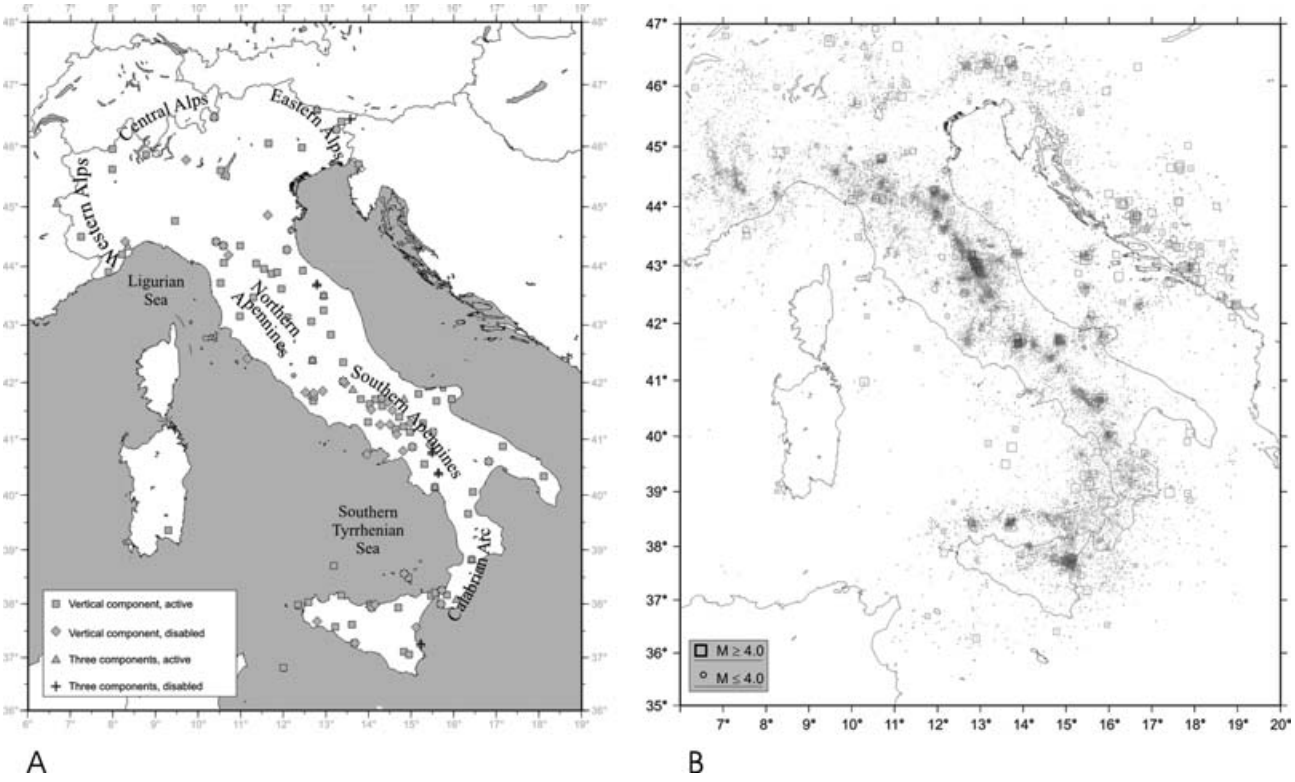

Figure 1. (a) Italian seismic stations network between 1988 and 2001: squares denote vertical component seismometers while triangles are three-components stations; (b) Italian seismicity recorded at INGV network between 1983 and 2001; dots indicate magnitude lower than 4.0 while squares denote events with $M \geq 4.0$. Note the uneven distribution of epicentres.

observation error estimates provided by the user. The quality weighting mechanism of MPX can be tuned toward predicting the same picking uncertainties as those that would be estimated by the user. As such, it acts as a consistent extension of the manual approach and not as a totally independent method possibly in conflict with the legacy. To calibrate MPX on INGV data, we selected a reference data set (RD) of about 700 waveforms that was picked and quality-weighted by an expert seismologist. We tuned the MPX procedure to mimic the behaviour of the seismologist, by comparing the results of the manual and the automatic picking (AP) on the RD. Finally, we applied MPX to a large data set of seismograms (seismicity for the period 1988–2002), testing the performance of the weighted MPX arrival times in standard earthquake location and in local earthquake tomography.

2 INGV STATION NETWORK AND SEISMICITY IN ITALY

The INGV seismic network (Fig. 1a) located over 50 000 local and regional earthquakes between 1983 and 2003. As shown in Fig. 1(b), the seismicity distribution is uneven, clustered along the axis of the Alps and Apennines and in the Calabrian Arc, with some large gaps in the Tyrrhenian and Adriatic regions (Amato *et al.* 1997). In addition to the abundant crustal seismicity, there are two regions of subcrustal earthquakes. One such region is located in the Northern Apennines, where foci reach 90 km depth (Selvaggi & Amato 1992), the other region being in the Calabrian Arc, where a Benioff zone is identified down to 500 km depth (Giardini & Veloná 1991; Selvaggi & Chiarabba 1995). The geometry of the INGV network significantly changed between 1980 and today and digital waveforms are only available since 1988. Only 60 stations were available in 1980 but their number is about 160 today. As a consequence, the sampling power of the data set has also changed considerably over the years. During the 1984–2001 period, most stations were equipped with short-period 1 s seismometers (S13), and signals were sampled at 50 Hz. At INGV, seismograms are picked visually by a team of analysts for routine location and magnitude determination. These parameters are stored in the waveform-event parameters database and represent the INGV seismic bulletin. During periods of intense seismic activity as in 1997, not all seismograms have been hand-picked and thus some *P*-wave arrivals are missing.

3 THE AUTOMATIC *P*-PHASE PICKING SYSTEM MPX

There are basically two main approaches to automated picking. A first way is to pick each seismogram independently from the others and one event at a time. Single-trace picking can be seen as an extension of the detection process, and some methods work in near-real time. It is also the first way analysts usually work with during the routine task of picking seismograms. Traditional methods quantify some attributes of waveforms like amplitude, frequency or polarization, and apply their detection and picking algorithms on these attributes or on a smoothed combination of them. Among the wide variety of traditional methods for picking first arrival *P*-waves of local and regional events, Allen (1978, 1982) designed a STA/LTA (short-term-average long-term-average) algorithm applied on an envelope function sensitive to both amplitude and

frequency of seismograms. Despite its age, Allen's picking system is still part of Earthworm (Johnson *et al.* 1994) and Sac2000 (Goldstein *et al.* 1999). Baer & Kradolfer (1987) derived their picking engine from Allen's work but they use the square of a modified envelope to generate a characteristic function whose value is tested against an adjustable threshold. The Baer–Kradolfer picker is integrated into Pitsa (Scherbaum & Johnson 1992). Autoregressive methods (Morita & Hamaguchi 1984; Takanami & Kitagawa 1988; Kushnir *et al.* 1990; GSE/JAPAN/40 1992; Takanami & Kitagawa 2003) work also well for picking local events, but they can also be used at regional and teleseismic distances (Leonard & Kennett 1999; Sleeman & van Eck 1999). The Cusum algorithm (Basseville & Nikiforov 1993) appears to be an attractive alternative to autoregressive methods at regional distances (Der & Shumway 1999; Der *et al.* 2000). Klumpen & Joswig (1993) apply pattern recognition on polarization images for *P*- and *S*-onset picking of three-component local data. Non-traditional methods like neural networks can sometimes work directly on seismograms (Dai & MacBeth 1995, 1997), avoiding the need to compute attributes or characteristic functions.

A second approach is to work on several seismograms at once, exploiting the similarity of waveforms from nearby events. This multitrace approach derives basically from controlled source seismic methods (Peraldi & Clement 1972) used to compute static corrections for seismic reflection data. In seismology, seismograms are typically organized in common station gathers (Dodge *et al.* 1995; Shearer 1997). A good account of the methodology and its historical progress is provided by Aster & Rowe (2000). The single-trace approach appears today to be more versatile than the multitrace approach in the sense that it is not based on the restrictive criterion of waveforms similarity. A multitrace approach is inherently better adapted to relatively restricted volume studies like those where clusters usually occur. Aster & Rowe (2000) suggest, however, that further developments in adaptive filtering and event clustering (Rowe *et al.* 2002) might significantly extend the range of data sets that can benefit from cross-correlation picking techniques.

Because MPX was originally developed to pick local events in the Dead Sea region (Aldersons 2004), where locally a great variety of waveforms is observed, it follows the single-trace approach. Despite the impressive number of picking methods reported, the need for complementary components has probably not received enough recognition in the literature. Examples of such components are high-fidelity filters of seismograms and methods for estimating time uncertainties associated with picked phases. In order to be meaningful, every physical measurement needs an assessment about its own uncertainty. Estimating picking uncertainties is an equally important function of an automated system as determining arrival times. MPX integrates the robust Baer–Kradolfer (1987) single-trace picking algorithm into a three-step automatic procedure (Aldersons 2004) consisting of pre-picking (spectral analysis and filtering), picking (and polarity) determination and error assessment.

Spectral analysis and waveform filtering

Filters are sometimes used in seismological analysis to reduce the level of noise that can obscure a precise recognition of phase onsets. In fact, commonly used filters often distort also the signal part while attempting to increase the signal-to-noise ratio (SNR). For this reason and in order to preserve as much as possible the true shape of onsets when filtering is used, MPX allows to apply an adaptive Wiener filter (Douglas 1997; Aldersons 2004) on the waveforms in the pre-picking step. Although the distortion by the Wiener filter of MPX is neglectable, we have not filtered the INGV data set because filtering did not significantly improve the picking accuracy or success rate. However, whether the Wiener filter is applied or not, SNRs derived from spectral densities analysis are always computed by the Wiener filter routines in the pre-picking step of MPX as they generate important predictors for the quality weighting engine. MPX is designed, in its present release, to enhance the quality of a catalogue of routinely located earthquakes when event recognition has already been performed by standard procedures. Thus, it works on waveforms associated to known, though approximate, earthquake locations. A theoretical arrival time (TP) can, therefore, be calculated and it is used in MPX as a reference to place two time windows in the noise and in the signal + noise part of the waveform, respectively (Fig. 2). The first window (N) is set before TP and is used to evaluate the noise spectral density while the second window is set after the pick and is used to evaluate the noise and signal spectral density. Spectral densities are calculated using a maximum entropy method (see Aldersons 2004). Since the predicted arrival time is used simply to focus MPX in the 'region' near the *P* onset, the TP can be determined by a simple 1-D ray tracing routine and the hypocentre–receiver distance. TP uncertainties as large as ± 3 s do not prevent MPX to converge to the real onset determination. Noise and signal + noise windows must be safely separated by gaps (Safety Gaps, Fig. 2) which width depends on the expected uncertainty in the TP calculation.

Arrival time picking

After the spectral density analysis, the waveform is used as the input for the picking engine (Baer & Kradolfer 1987). In this processing phase the characteristic function CF_i of the waveform (see Appendix A for details) is calculated and used by the picking engine to find the onset. The onset is accepted as a pick for a value of CF_i larger then a pre-defined *Threshold1* for a certain amount of time (*TupEvent)*. *Threshold1* is automatically determined by MPX based on the noise

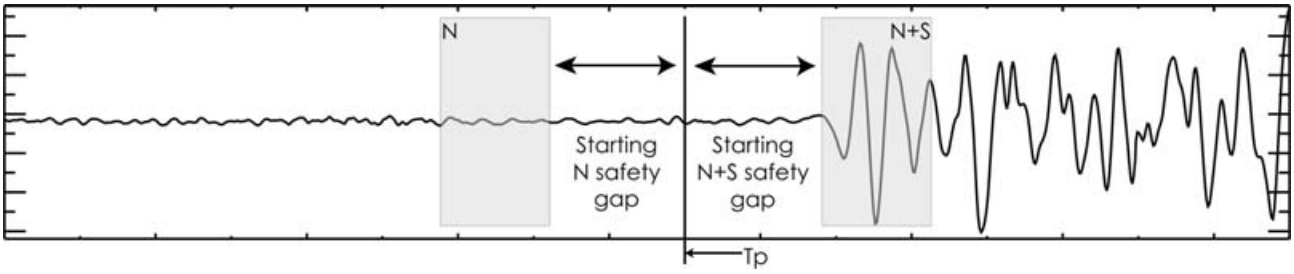

Figure 2. Calibration procedure setup: spectral density windows (N and $N + S$), TP theoretical arrival time predicted with IASP91 1-D earth velocity model, and the Safety Gaps set taking into account the prediction error, related to the TP.

© 2006 The Authors, *GJI*, **165,** 121–134
Journal compilation © 2006 RAS

window analysis, while *TupEvent* has a fixed value of 0.5 sec derived from the dominant frequencies of P onsets observed at local and regional distances (Appendix A). A typical frequency range for *Pg* and very short distance *Pn* from local earthquakes is between 1.0 Hz and 4 Hz. Polarity of the *P*-wave onset is also determined by the picking engine at this stage (Baer & Kradolfer 1987).

Error assessment

Finally, the observation error estimation follows an *a priori* scheme tuned to a reference waveform data set. Uncertainties of arrivals are automatically computed by the quality weighting algorithm taking into account the waveform sampling rate, the spectral densities analysis and the SNR. The goal of the weighting mechanism of MPX is to predict the uncertainty class for an automatic onset pick time from quantitative variables evaluated by the program. A calibration phase is needed to properly set up the weighting engine. The calibration is an iterative process consisting of four steps (summarized in Fig. 3):

(1) Building the RD: selection of a subset of waveforms from earthquakes representative of the whole data set (magnitude, hypocentre location and depth); definition of appropriate observation classes (quality weighting) hand-picking and accurate quality weighting of the RD;
(2) MPX setup: determination of the onsets dominant frequency; Safety Gaps setup;
(3) First run of MPX and
(4) Multiple discriminant analysis (MDA). This part is of specific importance for the topic dealt with in this work, and it is further discussed in the following paragraph.

4 CALIBRATION OF THE WEIGHTING SCHEME ON INGV DATA SET

The weighting mechanism of the Baer & Kradolfer (1987) algorithm implemented in Pitsa (Scherbaum & Johnson 1992) is based on the ratio between the *P*-wave amplitude within a certain time window and the largest amplitude of the seismogram before the pick flag is set for the first time. This kind of SNR is certainly important for the weighting of arrival times, but any variable by itself cannot carry all the information necessary to properly discriminate among time uncertainty groups such as those routinely used by seismological services. For this reason in MPX the quality weighting mechanism is based on the discriminating power provided by the simultaneous use of several predicting variables as in the Allen picking system. Adapting the weighting mechanism of the Allen system to a particular network and data requires however a lot of trial-and-error fitting at the very least. In MPX, a calibration is made according to a statistical technique known as discriminant analysis on a set of reference picks and associated weights supplied by the user. Picking uncertainties are expressed as standard labels 1 to *N*, with *N* not greater than 5 and a precise uncertainty duration defined by the user is associated with each of these labels. This calibration cannot be performed internally by MPX at the moment, but it can be easily performed from the automatic picking output of the RD with a package like SAS, SPSS or equivalent.[2] The RD of the INGV data consists of about

[2]Both SAS (http://www.sas.com/) and SPSS (http://www.spsstools.net/) are popular software packages for statistical analysis. Open source packages are also available.

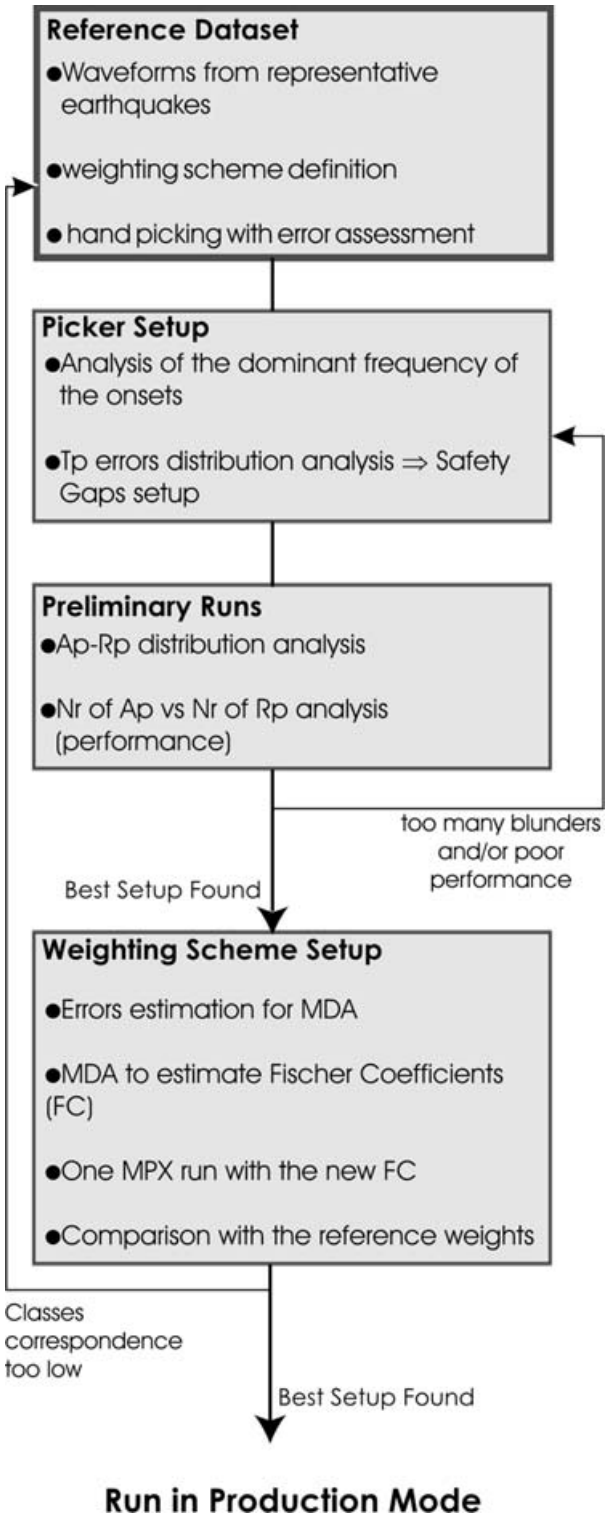

Figure 3. The four main steps of MPX calibration procedure; step 1, the creation of the data set, includes the conversion of the waveforms to the SAC format and the fulfilling of the header with important information like the reference picking (RP), the theoretical arrival time (TP), the origin time (TO), station and event coordinates and station name. The 'Picker setup' might be performed a few times depending on the width of the distribution of errors in the theoretical traveltimes (TP) and on the width of the distribution of the dominant frequencies of the *P*-wave onsets.

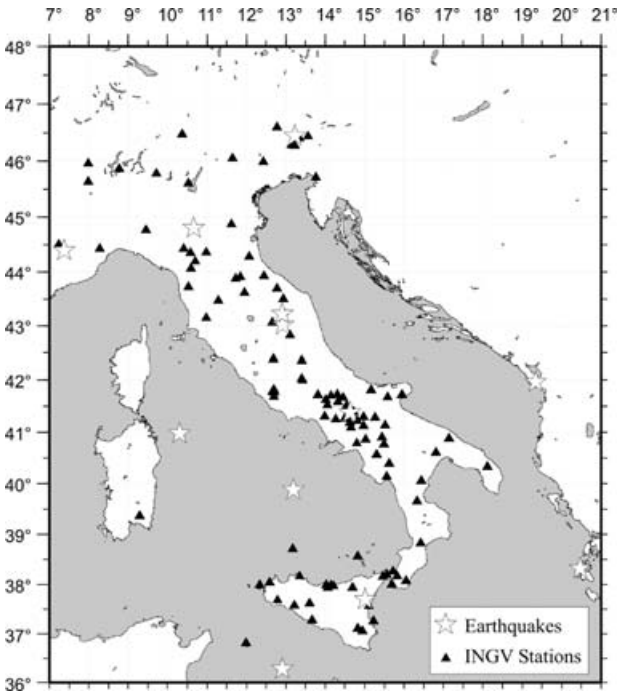

Figure 4. Reference data set; 12 epicentre locations (stars) and stations (triangles); the distribution of the stations, epicentres, hypocentres and the magnitudes giving a broad range of frequencies for the *P* onset in Italy are considered representative of the whole INGV data set.

700 waveforms from 12 events (Fig. 4) with magnitudes ranging from 3 to 5.5, and from shallow crustal depth to 600 km depth in the southern Tyrrhenian sea.

The application of MPX weighting engine on the INGV data demonstrated that enlarging the RD does not improve calibration results. More important is in fact to have a data set representative of the spectrum of frequencies characteristic for the first arrivals due to different sources, stations' locations and regional structure heterogeneities affecting the wave propagation. The waveforms of the RD have been accurately picked by hand (Fig. 5) and the errors have been estimated based on the weighting scheme shown in Table 1. We defined the error boundaries of each class according to the typical noise level in the waveforms and the sampling rate. The characteristics of the seismicity and of the stations network in the Italian region introduce the problem of dealing with different kinds of signals for the first onset, whose frequency content changes mainly with the hypocentral distance. *Pg*, short distance *Pn* and long distance *Pn* are present in the data set together with phases from deep events of the southern Tyrrhenian sea Benioff zone. In order to better calibrate MPX we have split the data set into *Pg*, *Pn* and deep events and we have run MPX separately on the three parts of the RD. The reference hand-picking (RP) and the AP (Fig. 5) are the key values in the MDA, together with variables (predictors), calculated by MPX for each AP (see Appendix B for details about MDA and predictors). The difference between AP and RP is the error ε associated to the AP. A weight class is then attributed to each $|\varepsilon|$ according to Table 1. A statistical relation between predictors and user-defined weight classes is found through MDA. Resulting Fischer's coefficients for the INGV reference dataset are provided in Table 2. Fischer's coefficients (Fischer 1936, 1938) actually represent the memory of the errors association, allowing MPX to make a prediction of the target weights on unseen cases (Aldersons 2004). Fischer's coefficients newly derived from the MDA are used in the MPX input file for the second run. This loop is stopped when the correspondence between the reference weights and MPX weights is maximum and the number of blunders introduced by the AP is minimized.

To assess the quality of the weighting scheme results, we compare the classes attributed by MPX to those determined by the seismologist for the same seismic signals (Table 3). The number of class 1, 2, 3 and rejected RP that fall into MPX class 1, 2, 3 and rejected, are counted to appraise the number of RP correctly weighted (bold), upgraded (grey cells), or downgraded (white cells) by MPX. A correct picking quality estimation (class 1 in class 1, class 2 in class 2, ...) being our goal, the higher are the main diagonal elements in Table 3 with as small as possible off-diagonal numbers, the better is the MPX performance. To establish high-quality data set, however, downgrading of a certain number of best quality picks by MPX (white cells in Table 3) can be endured, while an upgrading of even a few actually lowest quality RPs (light grey cells) cannot. As well, a few upgrades of medium quality picks (dark grey cells below the main diagonal in Table 3) are still acceptable in common geophysical applications. The calibration output shows that MPX correctly rejects 75 per cent of waveforms also rejected by the seismologist. Among the recognized arrival times ∼88 per cent of potentially high-quality picks (reference classes 1 and 2) are picked by MPX while ∼12 per cent is lost due to MPX misidentification. MPX recognizes also 65 per cent of reference classes 3 and 4 as arrival times, rejecting ∼35 per cent. Such evidence (Table 3) demonstrates that although we lose 12 per cent of potentially high-quality picks, MPX does not introduce mistakenly identified phases into high-quality classes from waveforms considered useless by the seismologist. Another important property is that ∼91 per cent of picks attributed by MPX to classes 1 and 2 correctly belong to reference classes 1 and 2. Fig. 6 displays traveltime versus distance values for each MPX weight class. The fit for MPX values of class 1 is nearly perfect both for shallow and for deep events. *Pg* and *Pn* traveltimes of shallow events can be represented by:

- $Y = 0.1653X + 1.2962$ (*Pg*)
- $Y = 0.1199X + 7.1858$ (*Pn*)

with a crossover distance of about 130 km corresponding to a flat Moho located at a mean depth of 25 km. High-quality picks are present for both *Pg* and *Pn*. MPX class 1 results fall into reference class 1 error range of ±0.1 s except 5 that actually should have been attributed to reference class 2 ($|\varepsilon| \approx = 0.15$ s) and 1 belonging to reference class 3 ($|\varepsilon| \sim = 0.25$ s). MPX class 2 results also display an almost linear fit. Errors are larger but mainly contained within the error range of class 2, except for about 11 arrival times out of 85, that should have been attributed by MPX to reference class 3 and 6 that belong to lower classes. Classes 3 and 4 are essentially dominated by downgraded picks. To better comprehend the reason for larger errors in MPX class 2 we must note that an important focus must be put on the meaning of the ε, whose distribution is shown in Fig. 7: ε is the difference between the RP and the AP assuming the hand-pickings and in particular their observation errors as correct. The formula of ε for the AP can be correctly rewritten as:

$$\varepsilon_a = AP - (RP + \varepsilon_r),$$

© 2006 The Authors, *GJI*, **165,** 121–134
Journal compilation © 2006 RAS

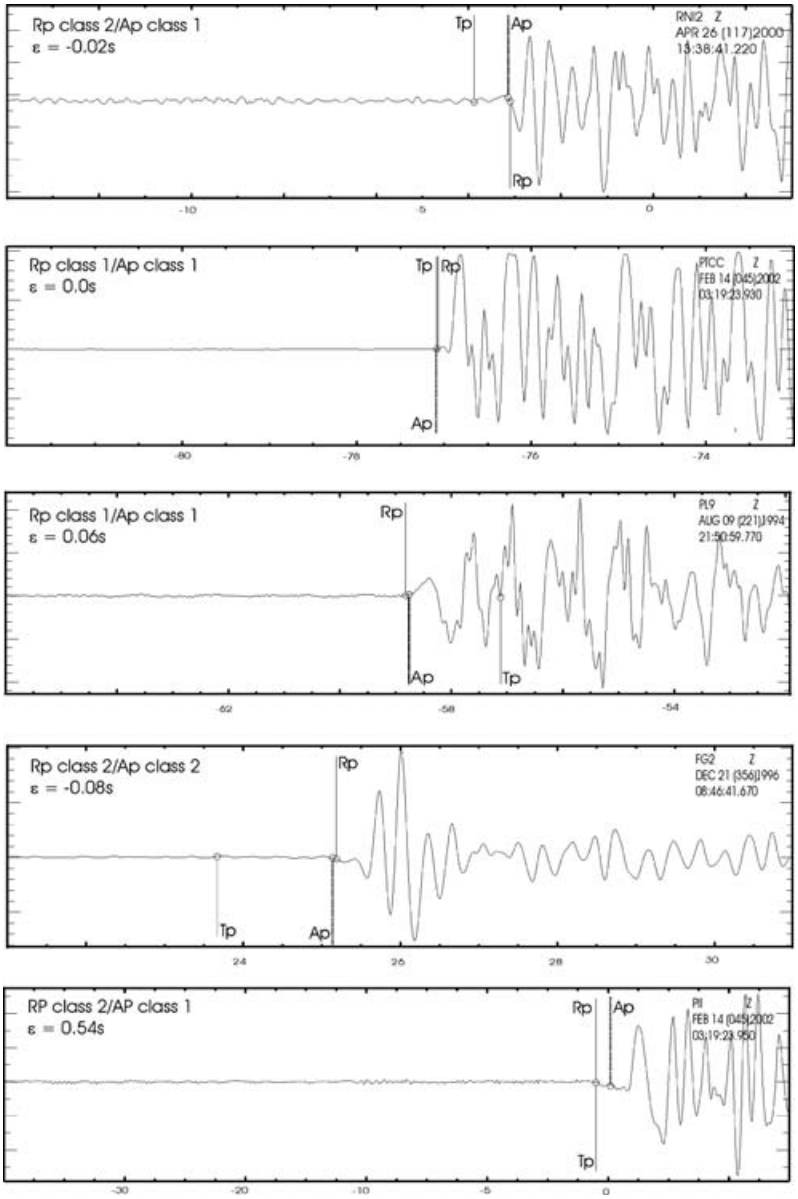

Figure 5. Examples of picked waveforms. Automatic and reference picks with related errors are compared. AP = automatic pick; RP = reference Pick; TP = theoretical arrival time.

where ε_r is the error in the hand-picks, while ε_a is the error in the AP. For class 1 of the hand-picks ε_r is small enough to have $\varepsilon_a \approx AP - RP$. In this case the ε_a distribution in Fig. 7 is significant as a quantitative error estimation. On the other hand, in classes 2, 3 and 4 of the hand-pickings, ε_r becomes progressively larger and the uncertainty in estimating AP errors increases. For this reason, the error estimation in these classes can be either too optimistic or too pessimistic. Fig. 7 shows that for class 1 of the MPX, errors are contained within the window ± 0.1 s, except for the six outliers mentioned above. This window exactly denotes the error boundaries of class 1 of the user defined weighting scheme (Table 1). Seventy-six per cent of picks in class 2 are contained within the range of reference class 2, while 88 per cent are within the range ± 0.4 s, that takes into account also the expected error for the RP. Thus we can be confident of high quality of classes 1 and 2 but we less accurately assess errors in classes 3 and 4. As a consequence, we define the high-quality data set for applicative uses to consist of classes 1 and 2 picks only. In the subsequent production stage we target to pick automatically all classes 1 and 2 signals.

5 PRODUCTION STAGE OF THE AUTOMATIC PICKING ON INGV DATA

We applied MPX with the calibrated weighting scheme to some 650 000 waveforms from about 78 000 local and regional events (1988–2002) of known location published (Chiarabba *et al.* 2005) in the new Italian Seismicity Catalogue (CSI). The events of the CSI catalogue were located by using bulletin data from INGV stations merged with data from other networks. We limited ourselves to repick INGV data only for which waveforms were already associated

Table 1. Weighting scheme applied to the Italian dataset: ε is the uncertainty attributed to a pick. ε between ± 0.10 and ± 0.20 means that the uncertainty bar is smaller than 0.4 s and greater than 0.2 s.

ε (s)	Class
$0.00 \div \pm 0.10$	1
$\pm 0.10 \div \pm 0.20$	2
$\pm 0.20 \div \pm 0.40$	3
$\pm 0.40 \div \pm 0.80$	4
$> \pm 0.80$	Rej

Table 2. Fischers linear discriminant coefficients of the Italian dataset (see Appendix B).

		Classes			
		1	2	3	4
Predictors	WfStoN	0.4283	.4228	.4235	.3954
	GdStoN	0.0217	.0376	.0145	.0126
	GdAmpR	0.0321	.0370	.0565	.0386
	GdSigFR	0.3331	.2948	.2173	.2396
	GdDelF	−0.0086	−.0117	.1066	.1009
	ThrCFRat	−0.2815	−.4683	−.4522	−.4177
	PcAboThr	−1.2612	−1.1877	−1.3602	−1.0344
	PcBelThr	0.0593	.0611	.0673	.0625
	CFNoiDev	6.8090	6.9343	7.1238	7.1325

to events and easily accessible. MPX recognized 217 435 *P*-wave arrival times (71 per cent of the 306 334 onsets recognized by the analysts) from 60 251 out of the 78 000 events. The first result is that while 90 per cent of bulletin readings are unweighted, we are now able to discriminate higher quality from lower quality picks. 37 929 out of the 217 435 *P* phases (17.44 per cent) fall into class 1 and 33 759 phases (15.52 per cent) fall into class 2. These percentages are slightly lower than for the calibration subset due to the presence of a larger number of very small magnitude events (noisy signals) in the full data set, yielding a higher number of weights 3 and 4 phases. Fig. 8 shows the time versus distance plots for the four MPX classes of picks. Due to the higher number and wider geographic distribution of events, a larger variability in depth and a much higher number of *Pg* phases with respect to the calibration data set (Fig. 6) are observed. On the other hand, the slopes and intercepts of the linear regressions of *Pg* and *Pn* for the full data set are in good agreement with the ones obtained for the calibration data set (Fig. 6). This fact and the results for classes 1 and 2, confirm the reliability of the method and the representativeness of the chosen calibration data set.

6 QUALITY ASSESSMENT OF THE AP RESULTS: 1-D LOCATION AND LOCAL EARTHQUAKE TOMOGRAPHY

To assess the quality of the repicking we first compare the 1-D locations obtained with only *P* phases from INGV stations picked with MPX to the data set obtained for the same events with bulletin phases. In order to restrict the analysis to rather well-constrained events, we selected 7878 events having more than four *P* readings, small location errors and an azimuthal gap smaller than 180°, both with bulletin data and with MPX data. MPX recognizes an overall number of 81 256 *P*-wave onsets, representing 78 per cent with respect to the 104 458 INGV bulletin readings related to the same events. 15 735 (19,37 per cent) of MPX picks fall into class 1, 14 679 (18.06 per cent) fall into class 2, 17 748 (21.84 per cent) fall into class 3 and 33 094 (40.73 per cent) fall into class 4, yielding thus

Table 3. Performance table (confusion matrix) showing the correspondance between MPX (Ap) and hand picks (Rp) classes.

		Ap Classes				
		1	*2*	*3*	*4*	*rej*
Rp Classes	*1*	**116**	47	22	13	13
	2	5	**21**	46	38	28
	3	1	11	**26**	29	30
	4	0	6	9	**18**	24
	rej	0	0	8	37	**134**

∼38 per cent (30 414) of high-quality arrival times and related polarities. The availability of such a high number of high-quality weighted polarities makes this subset also suitable for well-constrained focal mechanism determinations with a higher confidence than for the bulletin data (Di Stefano *et al.* 2002). The number of available polarities in the bulletin are in fact only 30 per cent with respect to the MPX data set and no error estimation is given. Since a simple 1-D model was used to locate events, long-distance ray paths travelling in complex structures such as those of the Italian region, may accumulate very large residuals. So the *P*-wave residuals after location depend also on the lateral seismic velocity heterogeneities. Nevertheless, the comparison of residual distributions for the two different data sets leads to some important conclusions. In Fig. 9, we show distributions of arrival time residuals (events with at least 4 phases), the difference in depth and the absolute depth calculations for both MPX and bulletin locations. Plot 1 in Fig. 9(a) represents bulletin data, for which the standard deviation σ is 0.91 s while for MPX data (plot 3) including low-quality classes 3 and 4, σ is 1.27 s. The same calculation performed on high-quality MPX picks (plot 2) gives a much smaller σ of 0.63 s. Despite the fact that location residuals suffer from misfits due to the 1-D model, these σ values show that most of the large residuals from MPX are due to classes 3 and 4. These results mark the importance of an accurate weighting scheme. On the other hand, due to the lack of proper weights, poor bulletin picks (CSI) cannot be separated from good picks leading to a higher level of noise. In Figs 9(b–d) we show the effect of the automatic picking and weighting on the location of events with either MPX or bulletin readings using the same method and parameters setup (rms cut, distance weighting, etc.). It is widely known that hypocentral depth is the most sensitive and less constrained focal parameter. Fig. 9(b) represents the distribution of absolute values of ε_z. By comparing these results with the histograms for MPX depths (Fig. 9c) and bulletin depths (Fig. 9d) we argue that most of the deepening is due to the fact that about 1000 events located with bulletin data are badly located at the surface while the same events are correctly located within the upper crust with MPX. Fig. 9(c) shows clustering of events around 7 km and 14 km, in agreement with what is known about the crustal seismicity in the target region (Selvaggi & Amato 1992; Amato *et al.* 1997). We thus verified the improvement on the depth variable when locating with MPX weighted picks. The mean of the distribution for MPX depths is around 10 km, (within the upper crust) while the epicentres follow the Apennines belt. Subcrustal seismicity is present to the east of the Northern Apennines. Such locations are consistent with those obtained using data from temporary local networks.

We have verified the resolution enhancement in local earthquakes tomography achievable by using the automatic picks, classes 1 to 4, along with a selection of weighted picks from other local and regional networks. We selected 8206 events that occurred from 1988 to 2002 with at least one MPX pick from the INGV network and we added *P*-wave arrival times with high-quality manual weights from

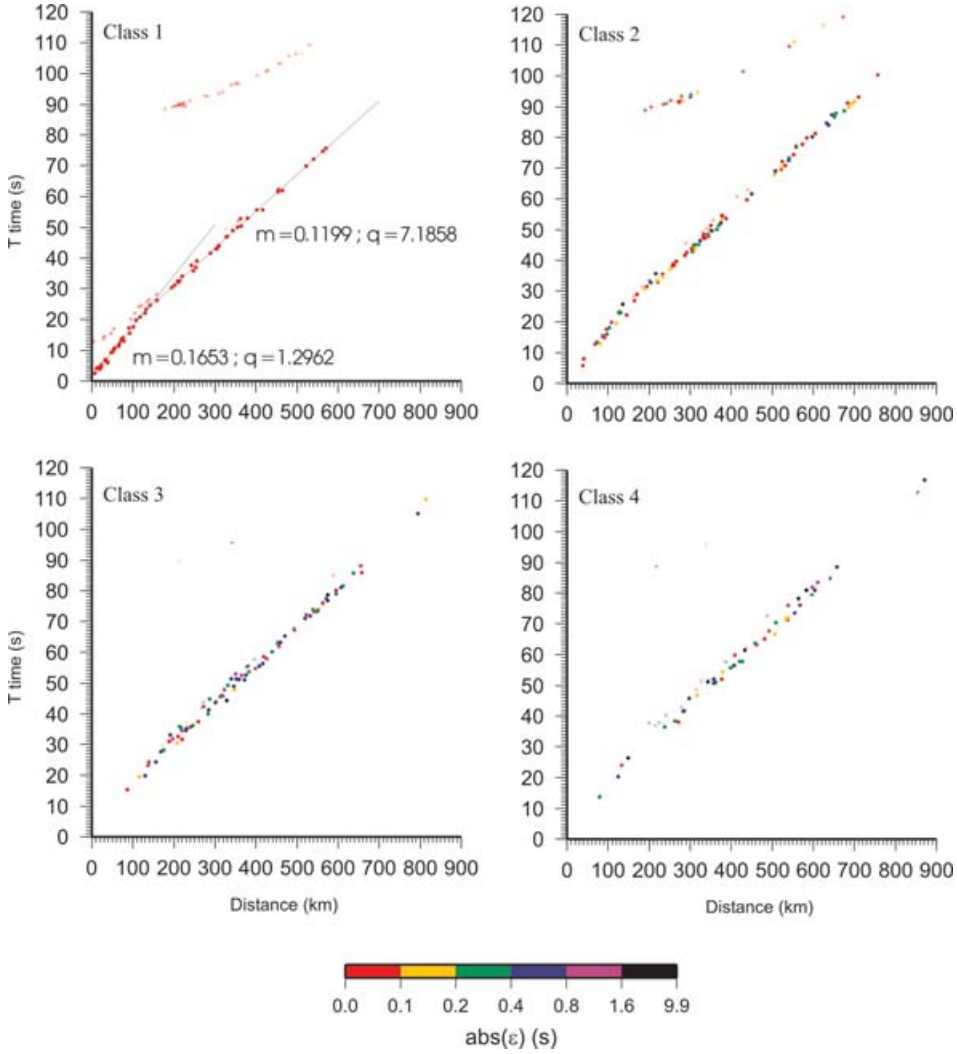

Figure 6. Plots for the MPX classes. Crosses represent events below 35 km depth, while circles are events with focal depth above 35 km. Colours are related to differences, in absolute value, with respect to hand reference picks.

other networks. The majority of phases consists of MPX picks. Selected events have more than 15 phases, azimuthal gap smaller than 180°, 1-D location errors less than 5 km (x, y, z) and 1-D RMS residuals less than 0.6s. The total number of *P* observations is 165,968. Based on the number of picks belonging to each class we estimate an rms $\leq$ 0.2 s for our high-quality data set, significantly smaller than the one of bulletin data. In previous studies, Chiarabba and Amato (1996) estimated that more than 45 per cent of the bulletin data (Italian seismicity 1975–1997) have errors higher than 0.2 s and that most of the largest errors are found at stations as distant as 60 km and above and especially on large distance *Pn* where errors reach $\pm$0.8 s. For this reason Chiarabba & Amato (1996), and Di Stefano *et al.* (1999) excluded long distance (200 km) *Pn* from their tomographic inversion. Mainly four factors contribute to final residuals rms in local source tomography: random data errors, systematic data errors, misfit to the velocity model, and misfit to the hypocentral parameters. Large random and systematic data errors must be properly identified and removed from the data set, as there is no way to separate them from real anomalies and hypocentre locations during the inversion procedure. This is the major source of errors in seismic tomography and it is in this part that benefits of using MPX are greatest, rejecting a large amount of noisy useless data and enhancing the effect of good data. Moreover, the use of MPX allows us to include long- and very long-distance *Pn* travel-times, previously affected by large errors, thus sampling large scale and deep structures, without blurring crustal anomalies revealed by high-quality *Pg* phases. Di Stefano *et al.* (1999) used some 48 000 bulletin *P*-wave arrival times to image the lithosphere beneath Italy with layers at 8, 22 and 38 km depth and a cell size of about 27 km in latitude and 21 km in longitude. With our data set we performed sensitivity checkerboard tests, decreasing the cell size to 15 km in latitude and longitude and adding three more layers at mantle depth (54, 66 and 80 km). We observe very high-to-fair resolution for all the layers except at 80 km depth where resolution is high only beneath the Calabrian Arc. Fig. 10(a) shows results of a synthetic test at 8 and 22 km depth, where anomalies are 30 km wide. The resolution is much higher than the one obtained by Di Stefano *et al.* (1999) especially at 8 km depth (Fig. 10b). The enhanced resolution of the crust and the low noise level in our automatic picks encourages future tomographic studies.

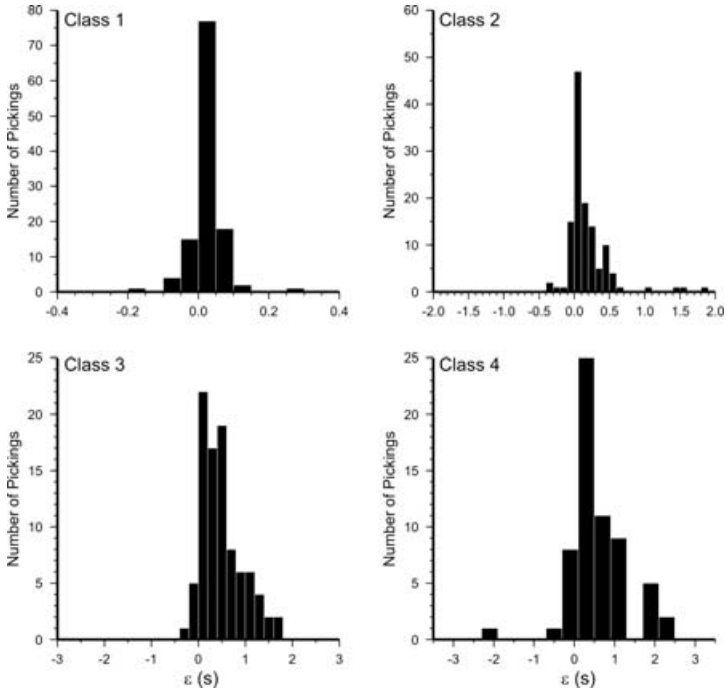

Figure 7. Distribution of time differences (AP − RP), in seconds, between MPX arrival times (AP) and reference picks (RP) within each weighting class.

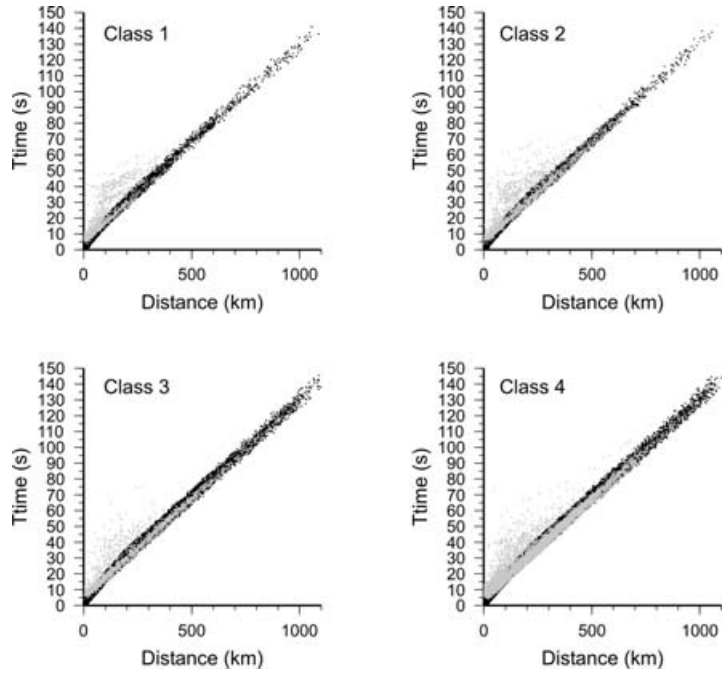

Figure 8. Time versus distance plot of automatic picks, classes 1 to 4. Black dots represent arrival times from shallow events ($Z \leq 35$ km) while grey dots are related to deeper events ($Z > 35$ km). The depth distribution being more continuous respect to the reference data set (Fig. 6), the grey dots overlay the shallow events arrival times distribution.

7 DISCUSSION AND CONCLUSIONS

The application of MPX to the very large and inhomogeneous data set of waveforms recorded in Italy from 1988 to 2002 yields 81 256 picks and polarities with observation error estimation associated, from about 7878 events. Though MPX applied to a large number of noisy data has a hit rate of 'only' 78 per cent with respect to an expert seismologist whose criteria are based on a visual inspection, the consistency of computed arrivals and polarities and their rapid estimation are suitable to substitute bulletin data, solving the typical problem of extending consistency and quality to large data sets and saving a very large amount of time. Results discussed in this paper

© 2006 The Authors, *GJI*, **165,** 121–134
Journal compilation © 2006 RAS

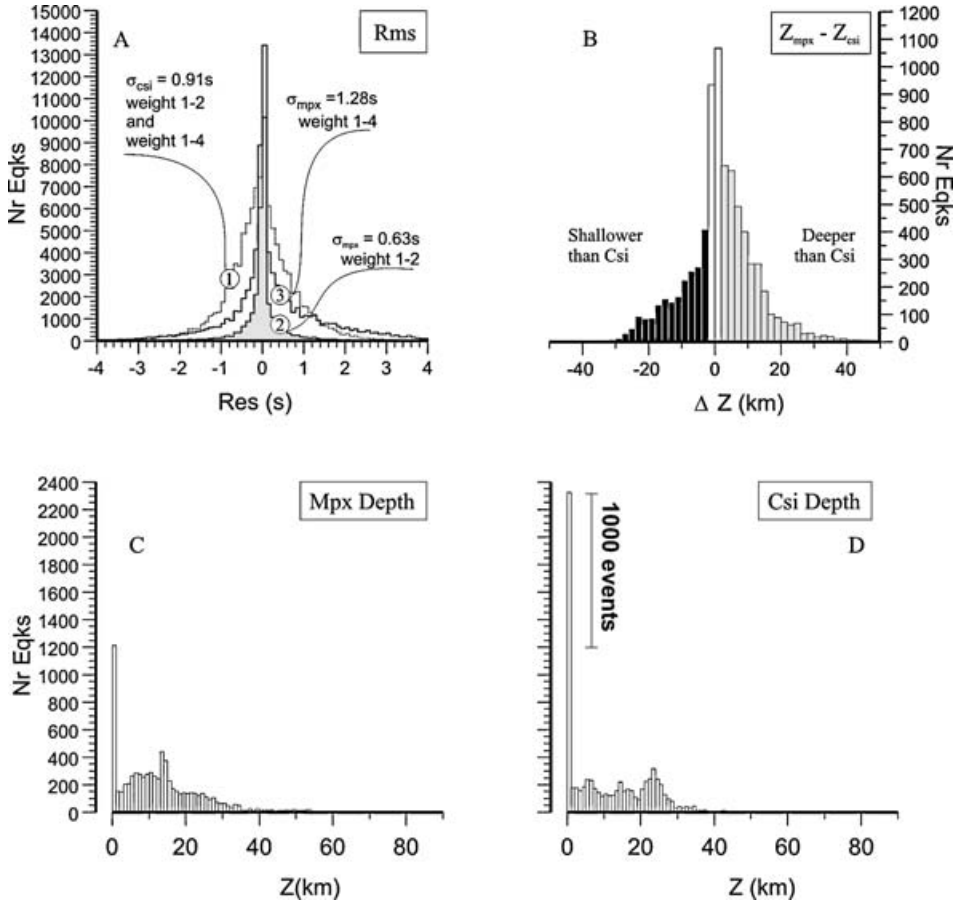

Figure 9. Distribution of residuals and depths after 1-D location. (a) Residuals from phases recorded only by INGV stations. Plot 1: CSI bulletin picks (classes 1-4). Plot 2: MPX picks (classes 1-2). Plot 3: MPX picks (classes 1-4). (b) Difference between depths derived from MPX and bulletin depths. Note that most depths derived from MPX picks are deeper than bulletin depths. (c) Depths derived from MPX. (d) Bulletin depths. Note how the number of events located at the surface is reduced with MPX.

suggest that 1-D locations obtained by using the MPX arrival times are more accurate and that hypocentral depths are better constrained than those obtained with INGV bulletin data only. The analysis of 1-D location residuals and the time versus distance plots show the effectiveness of the picking system. By using MPX we were able to strongly decrease the noise level of data, enhancing the resolution and the depth range of tomographic studies, demonstrating also that lower quality observations (classes 3 and 4) when properly weighted are still useful, complementing high-quality data in tomographic applications.

The successful application of the automatic picking system MPX to the INGV data represents a very significant test for its general applicability to other regional data sets. Our next step will be to significantly increase the number and consistency of high-quality phases by repicking waveforms recorded at other local and regional permanent networks to build the highest quality and largest data set of *P*-wave readings and polarities in the Italian region.

ACKNOWLEDGMENTS

We are grateful to the responsibles of the INGV national network who provided us with 20 years of digital recordings used in this study, and to the analysts and responsibles of the INGV-OV and INGV-CT, DipTeris Genova, CRS-OGS, University of Calabria, ENI-AGIP, Umbria Resil, Marche and Abruzzo regional networks (the last three sponsored by the National Seismic Survey) who provided complementary hand-picked data.

Our work has been strongly improved by the useful suggestions of two anonymous referees. A special thank is also due to P. De Gori, F. Bernardi, L. Chiaraluce, M.G. Ciaccio, B. Castello for helpful comments and E. Boschi for continuous encouragement.

REFERENCES

Aldersons, F., 2004. Toward a three-dimensional crustal structure of the Dead Sea region from local earthquake tomography, *PhD thesis,* Tel Aviv University, Israel, 120 pages. http://faldersons.net

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.

Amato, A., Chiarabba, C. & Selvaggi, G., 1997. Crustal and deep seismicity in Italy (30 years after), *Ann. Geofis.,* **XL**(5), 981–993.

Aster, R. & Rowe, C., 2000. Automatic phase pick refinement and similar event association in large seismic datasets, in *Advances in Seismic Event Location*, **18,** pp. 231–263, eds Thurber, C.H. & Rabinowitz, N., Modern approaches in geophysics, Kluwer Academic Publishers,Dordrecht, the Netherlands.

Baer, M. & Kradolfer, U., 1987. An Automatic phase picker for local and teleseismic events, *Bull. seism. Soc. Am.,* **77,** 1437–1445.

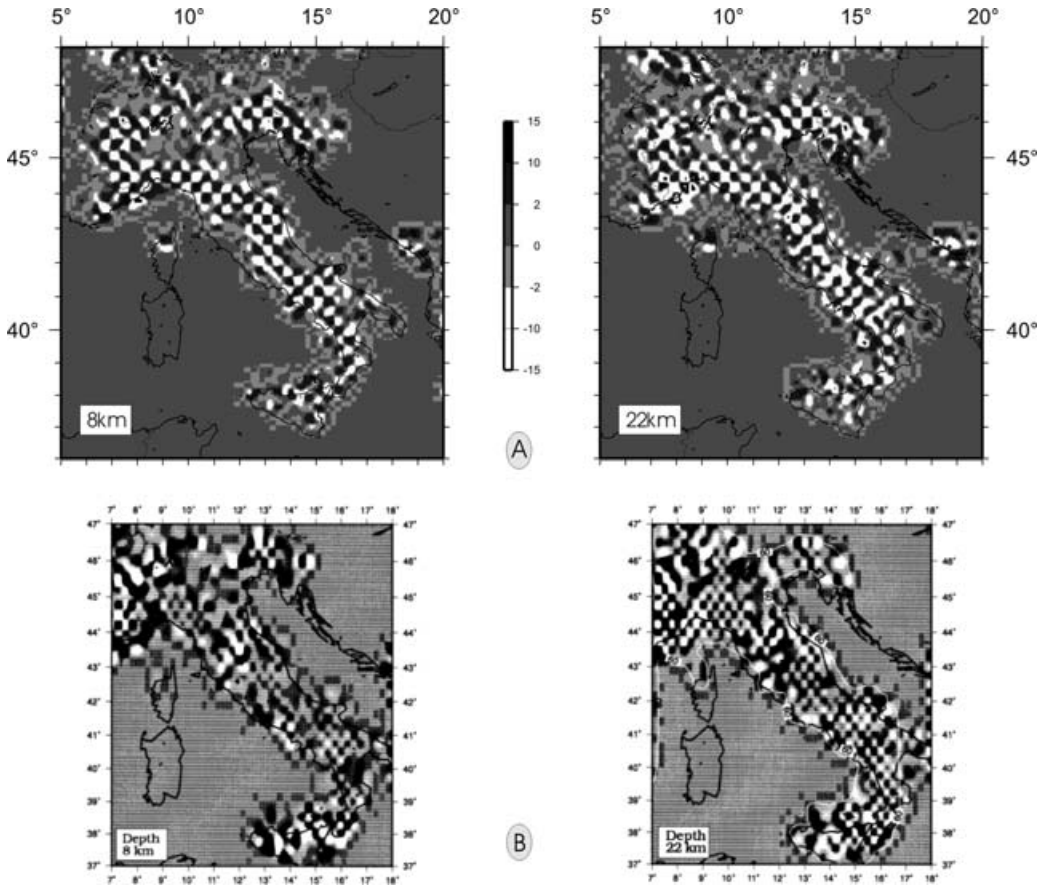

Figure 10. Comparison between sensitivity checkerboard tests. (a) *Pg* and *Pn* automatic picks from the INGV network merged with high-quality manual picks from other networks. Grid spacing is 15 km in latitude and longitude. The grey scale from −15 to 15 represents Vp perturbations in percentage. (b) *Pg* and very short distance *Pn* bulletin data from Di Stefano *et al.* (1999). Grid spacing is 0.25° in latitude and longitude.

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.

Chiarabba, C. & Amato, A., 1996. Crustal velocity structure of the Apennines (Italy) from P-wave travel time tomography, *Ann. Geofis.,* **39,** 1133–1148.

Chiarabba, C., Jovane, L. & Di Stefano, R., 2005. A new view of Italian seismicity using 20 years of instrumental recordings. *Tectonophysics,* **395,** 251–268.

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.

Der, Z.A. & Shumway, R.H., 1999. Phase onset time estimation at regional distances using the CUSUM algorithm, *Phys. Earth planet. Inter.,* **113,** 227–246.

Der, Z.A., McGarvey, M.W. & Shumway, R.H., 2000. Automatic interpretation of regional short period seismic signals using the CUSUM-SA algorithm, Proceedings of the 22nd DoD/DoE Seismic Research Symposium. New Orleans, LA, Louisiana.

Di Stefano, R., Chiarabba, C., Lucente, F. & Amato, A., 1999. Crustal and uppermost mantle structure in Italy from the inversion of P-wave arrival times: geodynamic implications, *Geophys. J. Int.,* **139,** 483–498.

Di Stefano, R., Amato, A., Aldersons, F. & Kissling, E., 2002. Automatic re-picking and re-weighting of first arrival times from the Italian Seismic Network waveforms database, *EOS Trans. Am. geophys. Un.,* **83,** S71A–1059, Fall Meet. Suppl.

Dodge, D.A., Beroza, G.C. & Ellsworth, W.L., 1995. Foreshock sequence of the 1992 Landers, California earthquake and its implications for earthquake nucleation, *J. geophys. Res.,* **100,** 9865–9880.

Douglas, A., 1997. Bandpass filtering to reduce noise on seismograms: is there a better way? *Bull. seism. Soc. Am.,* **87**(4), 770–777.

Fischer, R.A., 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugenics,* **7,** 179–188.

Fischer, R.A., 1938. The statistical utilization of multiple measurements, *Ann. Eugenics,* **8,** 376–386.

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 GSE/RF/62.

Giardini, D. & Veloná, M., 1991. The deep seismicity of the Tyrrhenian Sea, *Terra Nova,* **3,** 57–64.

Goldstein, P., Dodge, D. & Firpo, M., 1999. SAC2000: signal processing and analysis tools for seismologists and engineers, UCRL-JC-135963, Invited contribution to the IASPEI International Handbook of Earthquake and Engineering Seismology.

Johnson, S.J. & Anderson, N., 1978. On power estimation in maximum entropy spectrum analysis, *Geophysics,* **43,** 681–690.

Johnson, C.E., Lindh, A.G. & Hirshorn, B., 1994. Robust regional phase association, U.S.G.S. Open File Report 94–621.

Joswig, M. & Schulte-Theis, H., 1993. Master-event correlations of weak local earthquakes by dynamic waveform matching, *Geophys. J. Int.,* **113,** 562–574.

Klumpen, E. & Joswig, M., 1993. Automated reevaluation of local earthquake data by application of generic polarization patterns for P- and S-onsets, *Computers & Geosciences,* **19**(2), 223–231.

Kushnir, A., Lapshin, V., Pinsky, V. & Fyen, J., 1990. Statistically optimal event detection using small array data, *Bull. seism. Soc. Am.,* **80**(6b), 1934–1950.

Leonard, M. & Kennett, B.L.N., 1999. Multi-component autoregressive techniques for the analysis of seismograms, *Phys. Earth planet. Int.,* **113,** 247–263.

Lucente, F.P., Chiarabba, C., Cimini G.B., Giardini, D., 1999. Tomographic constraints on the geodynamic evolution of the Italian region, *Geophys. Res. Lett.,* **104,** 20 307–20 327.

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.

Lippitsch, R., Kissling, E. & Ansorge, J., 2003. Upper mantle structure beneath the Alpine orogen from high-resolution teleseismic tomography, *J. geophys. Res.,* **108**(B8), 2376, doi:10.1029/2002JB002016.

Peraldi, R. & Clement, A., 1972. Digital processing of refraction data study of first arrivals, *Geophys. Prospecting,* **20,** 529–548.

Piromallo, C. & Morelli, A., 2003. P wave tomography of the mantle under the Alpine-Mediterranean area *J. geophys. Res.,* **108**(B2), 2065, doi:10.1029/2002JB001757.

Rowe, C.A., Aster, R.C., Borchers, B. & Young, C.J., 2002. An automatic, adaptive algorithm for refining phase picks in large seismic data sets, *Bull. seism. Soc. Am.,* **92**(5), 1660–1674.

Selvaggi, G. & Amato, A., 1992. Subcrustal earthquakes in the Northern Apennines (Italy); evidence for a still active subduction? *Geophys. Res. Lett.,* **19**(21), 2127–2130.

Selvaggi, G. & Chiarabba, C., 1995. Seismicity and P-wave velocity image of the Southern Tyrrhenian subduction zone, *Geophys. J. Int.,* **121,** 818–826.

Shearer, P.M., 1997. Improving local earthquake locations using the L1 norm and waveform cross correlation: application to the Whittier Narrows, California, aftershock sequence, *J. geophys. Res.,* **102,** 8269–8283.

Scherbaum, F. & Johnson, J., 1992. Programmable Interactive Toolbox for Seismological Analysis (PITSA). IASPEI Software Library, **5,** Seismological Society of America, El Cerrito.

Sleeman, R. & van Eck, T., 1999. Robust automatic P-phase picking: an online implementation in the analysis of broadband seismogram recordings. *Phys. Earth planet. Int.,* **113,** 265–275.

Takanami, T. & Kitagawa, G., 1988. A new efficient procedure for the estimation of onset times of seismic waves, *J. phys. Earth,* **36,** 267–290.

Takanami, T. & Kitagawa, G. (eds), 2003. *Methods and Applications of Signal Processing in Seismic Network Operations.* Springer, Berlin, *Lecture Notes in Earth Sciences,* **98**.

APPENDIX A: AUTOMATIC PICKING

After the application of a Wiener filter in step 1 of MPX, the automatic determination of the first *P*-wave onset time is done by the picking algorithm of Baer & Kradolfer (1987). This algorithm defines first an approximate squared envelope function E_i^2 of the seismogram x_i, given by

$$E_i^2 = x_i^2 + \dot{x}_i^2 \frac{\sum_{j=1}^{i} x_j^2}{\sum_{j=1}^{i} \dot{x}_j^2}, \tag{A1}$$

where $\dot{x}$ denotes the time derivative of x. The algorithm uses then a characteristic function CF_i, defined as

$$CF_i = \frac{E_i^4 - \overline{E_i^4}}{\sigma^2\left(E_i^4\right)}, \tag{A2}$$

in which $\overline{E_i^4}$ is the mean of E_i^4 from $j = 1$ to i and $\sigma^2(E_i^4)$ is the variance of E_i^4. This characteristic function differs only from the statistical Z score of E_i^4 by the use of the variance σ^2 as denominator in (A2) instead of the standard deviation σ used in the Z score. A pick flag is set if CF_i increases above a given value *Threshold1*. The onset is accepted as a valid pick only if the value of the characteristic function stays above *Threshold1* for a certain amount of time *TUpEvent*. If the characteristic function drops below *Threshold1* after a shorter duration than *TUpEvent*, the pick is rejected and the pick flag is removed. The pick flag is however not removed if the characteristic function drops below *Threshold1* for an amount of time shorter than *TdownMax*. This happens frequently for events with low-to-moderate SNRs. The variance $\sigma^2(E_i^4)$ is continuously updated to account for variations of the noise level, except when CF_i exceeds the given value *Threshold2* usually chosen greater than *Threshold1*. The variance of E_i^4 is thus frozen when the validity of the onset is examined, shortly after the pick flag has been set. In MPX, the *Threshold1* value of the characteristic function that triggers the pick flag is determined in a fully adaptive and automatic way. For each seismogram, *Threshold1* is first set equal to the highest value of CF_i within the time window corresponding to the noise segment of the Wiener filter. If the SNR after the application of the Wiener filter is greater than 14 dB, the value of *Threshold1* is increased according to an internal table derived empirically. Regarding the amount of time *TUpEvent* required to validate an onset after a pick flag has been set, Baer & Kradolfer (1987) recommend to use at least one full cycle of the longest signal period expected. In the case of local events that occurred in the vicinity of the Dead Sea basin, *TUpEvent* has an optimal value at 0.5 s. A value of 1.0 s could be a sensible choice when regional and local events are mixed together in one single picking set. The amount of time *TdownMax* during which CF_i can drop below *Threshold1* without clearing the pick flag has an optimal value in MPX at half the signal period of highest power density as measured by the Wiener filter routine. This value is quite consistent with Baer & Kradolfer (1987) who recommend half the mean of the two corner periods of their bandpass filter. *Threshold2,* the threshold value for freezing the variance update of E_i^4, is optimal in MPX at a value of $2 \times$ *Threshold1*. This value is again in perfect agreement with the recommendations of Baer and Kradolfer.

Delay corrections

One known shortcoming (e.g. Sleeman & van Eck 1999) of the Baer–Kradolfer algorithm is that the raw onset time provided is always somewhat late compared to what an analyst would determine as the onset time of a valid phase. For local earthquake data recorded by short-period instruments, this delay can be as small as one sample for the best seismograms. For most seismograms, the delay is greater due to the interference of noise. In order to reduce the delay of the Baer–Kradolfer onset times, all versions of MPX include a primary delay correction. This correction derives from how the onset is determined by the Baer–Kradolfer algorithm. When a pick flag is set and validated, the characteristic function value is generally far above the background level observed in the segment of noise immediately preceding the onset. The idea for the correction is to move the raw automatic onset back in time as long as the characteristic function decreases significantly toward earlier samples. The delay correction stops when $(CF_i - CF_{i-1})$ is smaller than 0.01, or when this condition cannot be met after moving back the onset by three samples. This process is not perfect since the characteristic function may not decrease monotonically toward its background level, or because a clear background level may not even exist in the vicinity of the onset due to the disturbance by strong noise. Nevertheless, this simple

correction usually provides good to very acceptable results for local earthquakes recorded by short-period instruments. When regional and teleseismic data are picked with the original Baer–Kradolfer algorithm, raw picking delays can be much longer than a few samples. This behaviour of the algorithm has also been observed by Sleeman & van Eck (1999). It results apparently from the fact that little or no frequency contrast ($\dot{x}^2$ terms in A1) between the *P* wave train and the noise contributes then to the build-up of the characteristic function for these events. At lower frequencies typical of greater epicentral distances, the characteristic function is slow to reach the picking threshold level. The primary correction becomes then insufficient. In order to better correct the delay for regional and teleseismic arrivals, a secondary delay correction exists in MPX. This secondary correction is applied immediately after the primary correction. It does not appear to interfere adversely with the primary correction, so important for local earthquake data. A simple moving average $SMA_{i,P}$ of period *P* is given at the seismogram sample index *i* by

$$SMA_{i,P} = \frac{\sum_{j=0}^{P-1} x_{i-j}}{P}, \quad i \geq P. \tag{A3}$$

The corresponding simple moving standard deviation $SMSTD_{i,P}$ is then

$$SMSTD_{i,P} = \left[\frac{\sum_{j=0}^{P-1} (x_{i-j} - SMA_{i,P})^2}{P} \right]^{1/2}, \quad i \geq P. \tag{A4}$$

MPX 1.7 uses a number of samples *P* in (A3) and (A4) corresponding to the period of highest noise spectral density as determined by the Wiener filter routine. The standard deviation band of MPX 1.7 is defined as

$$STDB_{i,P} = SMA_{i,P} \pm 2 \times SMSTD_{i,P}, \quad i \geq P. \tag{A5}$$

For a positive onset (increasing amplitude of first motion), the precise condition is that the delay correction stops at the first sample x_i below the higher value of $STDB_{i,P}$ in (A5). This condition is usually met at an intermediate value between the two $STDB_{i,P}$ values. For a negative onset, the delay correction stops at the first sample x_i above the lower value of $STDB_{i,P}$. The secondary correction is however skipped if the seismogram value x_i at the onset time corrected by the primary correction is not out the deviation band $STDB_{i,P}$ by more than $0.1 \times SMSTD_{i,P}$.

APPENDIX B: DISCRIMINANT ANALYSIS

Discriminant analysis is a statistical technique whose general purpose is to identify quantitative relationships between two or more criterion groups and a set of discriminating variables also called predictors. It requires the criterion groups to be mutually exclusive, and it assumes the absence of collinearity among discriminating variables, the equality of population covariance matrices and multivariate normality for each group. The mathematical objective is to weight and linearly combine the predictors in a way that maximizes the differences between groups while minimizing differences within groups (Fischer 1936, 1938). When more than two groups are involved, the technique is commonly referred to as multiple discriminant analysis (MDA). As a *descriptive* technique, discriminant analysis serves to explain how various groups differ, what the differences between and among groups are on a specific set of discriminating variables and which of these variables best account for the differences. As a *predictive* technique, it is used to predict the unknown group membership of cases based on the actual value of the discriminating variables. The problem at hand in the weighting mechanism of MPX is to predict the uncertainty group of AP times from quantitative variables evaluated by the program. In order to be able to perform these predictions, it is however necessary to derive classification rules from an initial descriptive discriminant analysis on data for which the group membership is known. For a given data set, this discriminant analysis provides valuable guidelines to the user regarding both the definition of appropriate uncertainty groups and the selection of the most relevant predictors. The group membership predictions on unseen data are not done in a statistical package. MPX performs this prediction internally by computing Fischer's linear discriminant functions from the coefficients determined during the initial descriptive discriminant analysis.

The discriminating variables of MPX

The lack of discriminating power resulting from the use of only one predictor can be easily understood. For instance, any SNR derived from a segment of signal and a segment of noise implies already that a limiting choice has been made. Basically, the length of the segments can be chosen to be long or short compared to the dominant period of the signal. If long segments are selected, chances are that a general characterization of the overall quality of the seismograms will be gained. This overall quality might then discriminate rather well between picks located far away from true onsets and those located closer from true onsets. What is missing, however, in the prediction process is a localized SNR value derived from very short segments. Used in isolation, this localized SNR might reflect the accuracy of the picked features, but little will be known about the possibility of gross errors. When the information provided by both predictors is combined, some discriminating power is gained because a decision matrix can then be derived with information about both the possibility of gross errors and a very localized quality factor. Similarly, the discriminating power can be further increased by adding appropriate variables in the prediction process. Naturally, defining the most appropriate predictors is not a straightforward task. The benefits of using discriminant analysis for this purpose are that trial and error is reduced to a minimum, no decision matrix needs to be coded explicitly and the best predictors for specific data sets can be easily selected from the pool of standard discriminating variables already gathered. Discriminant analysis is thus useful both to define new predictors, and to select the available ones most appropriate to tackle specific data sets. The pool of standard predictors available so far includes nine variables evaluated internally by MPX. On specific data sets, the optimal number of predictors usually ranges from six to eight among the nine available.

Predictor 1 (*WfStoN*) is a SNR derived from the signal and noise power spectra determined by the Wiener filter routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_1 = 10 \log_{10} \frac{\sum_{f=0}^{F_{Ny}} P_{S,S}(f)}{\sum_{f=0}^{F_{Ny}} P_{N,N}(f)}, \tag{B1}$$

where $P_{S,S}(f)$ and $P_{N,N}(f)$ are the power spectral densities of the signal and the noise, respectively, and F_{Ny} is the Nyquist Frequency. Although the Wiener filter uses short data segments (about 2 s for local data), the length of these segments is long compared to one apparent cycle of the *P* wave train. Predictor 1 is an overall quality estimate around the final automatic onset time.

Predictor 2 (*GdStoN*) is a SNR determined by the delay-correction routines. Its value is evaluated at the final automatic onset

© 2006 The Authors, *GJI*, **165,** 121–134
Journal compilation © 2006 RAS

time and it is given in decibels by

$$P_2 = 20 \log_{10} \frac{\sum_{f=0}^{F_{Ny}} A_S(f)}{\sum_{f=0}^{F_{Ny}} A_N(f)}, \tag{B2}$$

where $A_S(f)$ and $A_N(f)$ are the magnitude spectra of the signal and the noise, respectively, and F_{Ny} is the Nyquist frequency. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the *P* wave train. Predictor 2 is a localized measure of quality at the final automatic onset time.

Predictor 3 (*GdAmpR*) is a SNR determined in the time domain by the delay-correction routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_3 = 20 \log_{10} \frac{Amp_S}{Amp_N}, \tag{B3}$$

where Amp_S and Amp_N are the peak-to-peak maximum amplitude of the signal and noise, respectively. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the *P* wave train. Predictor 3 is also a localized measure of quality at the final automatic onset time.

Predictor 4 (*GdSigFR*) is the SNR evaluated at the dominant frequency of the signal, as determined by the delay-correction routines. Its value is evaluated at the final automatic onset time and it is given in decibels by

$$P_4 = 20 \log_{10} \frac{A_S(F_S)}{A_N(F_S)}, \tag{B4}$$

where $A_S(F_S)$ is the value of the magnitude spectrum of the signal at the dominant frequency of the signal $f = F_S$, and $A_N(F_S)$ is the value of the magnitude spectrum of the noise at frequency $f = F_S$. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the *P* wave train. Predictor 4 is also a localized measure of quality at the final automatic onset time.

Predictor 5 (*GdDelF*) is the difference between the dominant frequency of the signal and the dominant frequency of the noise, as determined by the delay-correction routines. Its value is evaluated at the final automatic onset time and it is given in hertz's by

$$P_5 = F_S - F_N, \tag{B5}$$

where $f = F_S$ is the dominant frequency of the signal and $f = F_N$ is the dominant frequency of the noise. The length of the noise segment is half the length of the Wiener filter noise segment. The signal segment extends from the picked onset until the second zero-crossing after the onset. Its length is about one apparent cycle of the *P* wave train. Predictor 5 is a localized measure of the frequency contrast between the signal and the noise.

Predictor 6 (*ThrCFRat*) is the natural logarithm of the ratio between the first maximum of the characteristic function after picking to the *Threshold1* value, as determined by the AP routines. It is given by

$$P_6 = \ln \frac{CF_{\mathrm{Max1}}}{\mathrm{Threshold1}}. \tag{B6}$$

Predictor 6 is a localized measure of quality derived from the characteristic function.

Predictor 7 (*PcAboThr*) is the percentage of characteristic function samples above the *Threshold1* value before the picked onset, as determined by the AP routines. The data segment has the length of the Wiener filter noise segment and stops one sample before the picked onset. A Predictor 7 value different from zero is often associated with a possible earlier pick time, or with a high level of pre-arrival noise.

Predictor 8 (*PcBelThr*) is the percentage of characteristic function samples below the *Threshold1* value after the picked onset, as determined by the AP routines. The signal data segment length is 1/10 the length of the Wiener filter mixed signal-and-noise segment and starts at the picked onset. A Predictor 8 value different from zero is often associated with low-to-moderate SNRs.

Predictor 9 (*CFNoiDev*) is a measure of deviation of the characteristic function before the picked onset, as determined by the AP routines. It is given by

$$P_9 = \ln \frac{\frac{\sum_{i=1}^{N} |CF_i - CF_{\mathrm{median}}|}{N}}{|\mathrm{Threshold1} - \mathrm{CF}_{\mathrm{median}}|}, \tag{B7}$$

where the CF_i are the N samples of the characteristic function in a noise segment of length equal to the length of the Wiener filter noise segment, and CF_{median} is the median value of CF_i over the N samples.

© 2006 The Authors, *GJI*, **165,** 121–134
Journal compilation © 2006 RAS