Automatic Seismic Phase Picking and Consistent Error Assessment: Application to Local Earthquake Data in the Alpine Region

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

[1]Institute of Geophysics, ETH Zurich, Switzerland, email: diehl@tomo.ig.erdw.ethz.ch
[2]Swiss Seismological Service, ETH Zurich, Switzerland
[3]Mies, Switzerland

Introduction

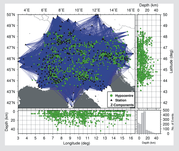
Inconsistencies in Merged Data Sets

Why Automated Repicking of Waveform Data?

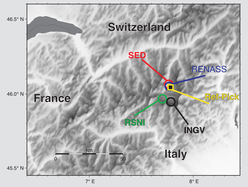
Data sets based on merged routine phase data from several networks usually include a significant degree of inconsistency due to:

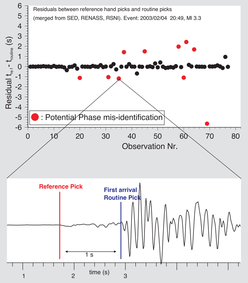
- Timing uncertainties of phase picks is often unknown or inconsistent.
- Phase identification is often not consistent. No consistent error assessment for merged phase data.
- Identification of common events can be biased through significantly different locations

Furthermore, modern data sets are usually too large for consistent manual repicking:



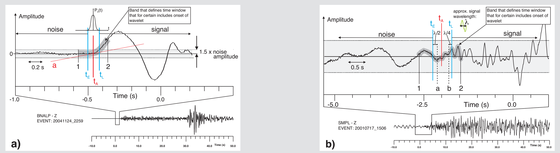
Waveform data set for Alpine region:
Networks: 13
Events: 1521
Stations: 381
Z-Comp: 72252
N-Comp: 35852
E-Comp: 35852





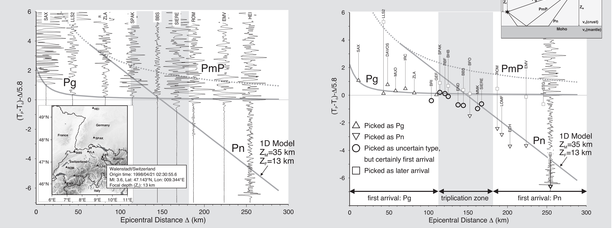
Consistent Quality Assessment

Consistent Timing Quality:



Probabilistic phase picking approach: the "earliest" possible pick corresponds to t_{el}, the "latest" possible pick corresponds to t_l. The most likely arrival time t_e is located within this interval. **a)** Primarily amplitude is used for determination of t_{el} and t_l. **b)** Primarily frequency is used for determination of t_{el} and t_l.

Consistent Phase Identification:



Automatic Picking System Mannekenpix (*MPX*)

Basic Concept

Wiener Filter
Used to determine threshold (picking engine), waveform charcteristics (predictors) for quality assessment, and waveform filtering

Picking Engine:
- Initial pick defines position of search window for repicking (**routine** pick if available or **predicted** from simple model)
- Extended Baer-Kradolfer (1987) algorithm

Quality Assessment Engine:
- True uncertainty class from discriminant waveform characteristics (predictors) via Fischer coefficients
- Weighting scheme has to be calibrated with a (hand picked) reference data set in order to adopt error assessment of seismologist: Multiple Discriminant Analysis (MDA) to find appropriate Fischer coefficients for data set

Consistently picked phases including individual quality assessment

1. Step: Calibration Procedure
=> **Reference Data Set:** consistent manually picked for representative subset of waveform data set
=> Adopt automatic picking and quality assessment to reference data

2. Step: Production Procedure
=> Initial picks necessary for all seismograms
=> Apply calibration scheme to complete waveform data set

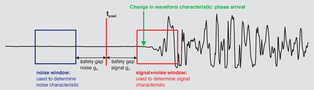
MPX Calibration Procedure

Summary of Picking Parameters to Calibrate:

MPX Parameter	Achieved by	Dependencies	Example from Alpine Ref. Data
Fischer coefficients (#Predictors x #Classes)	MDA, Test with Ref. Data	waveform characteristics of data set	not shown here
Length of analysing window (for Wiener Filter) (s)	Test with Ref. Data	dominant frequency content of signals	2 s
Set of safetey gaps for predicted pick (s)	Test with Ref. Data	scatter of predicted & frequency content	2.5 s, 1.25 s, 0.5 s, 0.08 s
Set of safetey gaps for routine pick (s)	Test with Ref. Data	scatter of routine & frequency content	not used
Lower frequency threshold (Hz) (Frequencies below threshold are not considered for SNR determination)	Test with Ref. Data	dominant signal frequency	0.50 Hz

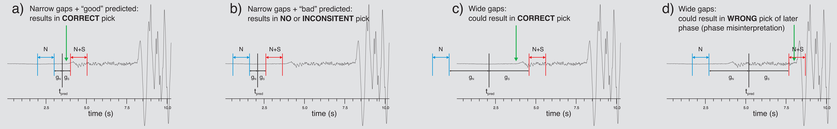
1. Step: Initial run of MPX to determine Fischer coefficients appropriate for specific waveform data set (Aldersons 2004)

2.Step: Adjustment of search window parameters

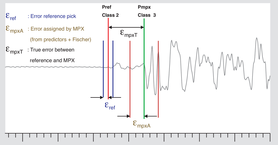


! Length of analysing windows (noise and signal+noise): Depending on signal frequency

! Safety gaps: Depending on scatter of initial pick and on signal frequency
Tradeoff between hit-rate and accurancy as illustrated in a) - d)



3. Step: Check of performance by comparison with reference data set

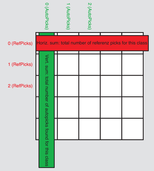


To assess the quality of the automatic pick we have to check:

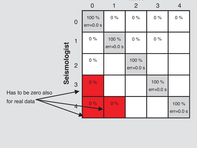
! **True error** ε_{mpxT} (difference between reference and automatic pick).

! **Difference in error classes** ε_{ref} **and** ε_{mpxA} (low quality reference picks should not be upgraded to highest quality class by MPX)

The Performance-Matrix: Smearing of quality classes and standard deviation "err" of ε_{mpxT} for each element



For ideal calibration:



Remark: This is an iterative process of repeating several times step 1-3 until satisfying result is achived

Application to Alpine Region

Reference Data Set

Error Assessment used for Reference Data Set:

Class	Error (s)	Weight
0	0.00 < |ε| ≤ 0.05	100 %
1	0.05 < |ε| ≤ 0.10	50 %
2	0.10 < |ε| ≤ 0.20	25 %
3	0.20 < |ε| ≤ 0.40	12 %
4	|ε| > 0.40	0 %

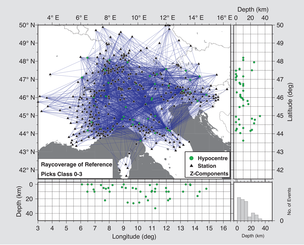
Phase label	Phase is…	Phase used for routine 1D arrival studies
Pg	Direct (crustal)	Yes (if first arrival)
Pn	Moho-refracted	Yes (if first arrival)
PmP	Moho-reflected	No
P1	Unknown type, but certainly first arrival	Yes
P2	Unknown type, second arrival	No
P3	Unknown type, third arrival	No
P	Unknown type, uncertain if first arrival	No

Picking Statistics for Reference Data Set:

Class	Number of Picks
0	773
1	926
2	905
3	502
4	901

Phase	Number of Picks
Pg	301
Pn	1959
P1	591
P	1156

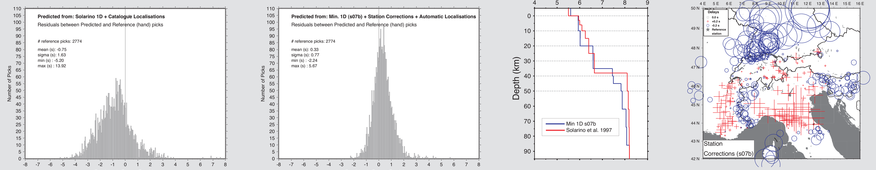
Number of Earthquakes: 43
Number of Stations: 359



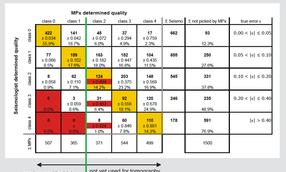
MPX Production Procedure

Initial Picks for MPX-Production Mode

The scatter of predicted arrivals derived from a standard model (Solarino et al. 1997 + catalogue locations) is too large to be used as initial picks (requires too wide safety gaps). The use of a min. 1D model (s07b) + station correction + relocated earthquakes (based on un-weighted automatic picks) leads to a significant reduction of the scatter and allows narrow safety gaps.



Performance of Calibration



Performance of Production

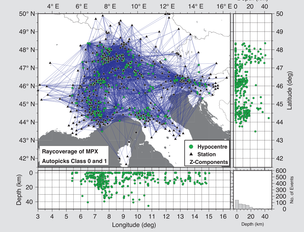
All events used for production: 916

MPX-DeterminedClass	MPX-Picks
0	6242
1	3207
Sum:	9449

Well locatable events with class 0 and 1 picks (Nobs > 5, GAP < 180): 407

MPX-DeterminedClass	MPX-Picks
0	4993
1	2595
Sum:	7588

consistent uncertainty: 0.00 < |ε| ≤ 0.100 s



Discussion & Outlook

The presented calibration and automated production procedure is able to pick the highest quality P-wave arrivals from a waveform data set of over 70,000 seismograms with an average uncertainty < 0.100 s. Such a performance cannot be reached in reasonable time by manual repicking. However, the present calibration downgrades too much of the "class 1" picks towards lower quality classes. In a next step, we will test an iterative calibration procedure (MDA for different subsets of data) in order to improve prediction and discrimination of class 1 and 2.

References

Aldersons, F. (2004) Toward a Three-Dimensional Crustal Structure of the Dead Sea region from Local Earthquake Tomography. *Ph.D. thesis*, Tel Aviv University, pp. 123.

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

Acknowledgements

We wish to thank people from BED, GRSN, GERESS, INGV, LED,OGS, RENASS, RSNI, SED, SISMALP, SNRS, TGRS, ZAMG for providing us with waveform data

Poster presented at SSA 2007, Kona, Hawaii, USA