Topic	Seismogram Picking Error from Analyst Review (SPEAR)
Author	**Cleat Zeiler**, University of Texas at El Paso, El Paso, Texas, USA; now: Air Force Technical Application Center, Patrick AFB, Florida, USA, E-mail: cpzeiler@miners.utep.edu **Aaron Velasco**, University of Texas at El Paso, El Paso, Texas, USA, E-mail: aavelasco@utep.edu
Version	July 2011; **DOI: 10.2312/GFZ.NMSOP-2_IS_11.5**

1 Introduction

The Seismogram Picking Error from Analyst Review (SPEAR) project was initiated in 2006 to determine the reliability and accuracy of picks made by analysts. The process of picking is defined as the timing, measuring and naming of seismic phases used to locate and classify seismic events. Prior to this study only one analyst-based study had been conducted (Freedman, 1966a). While the general practice and theory of picking has not changed significantly from 1966, the technology and tools used to analyze waveforms has improved, along with the resolution of digitized waveforms (Lomnitz, 2006). Additionally, the 1966 study was limited to a small subset of analysts that used paper records. To continue and modernize the study of analyst based picks, we implemented a two-part study by first looking at available catalogue data (Zeiler and Velasco, 2009) and then by creating a common data set for analysts to pick.

The initial looks at archived data provided a limited data set to further understand pick error. Zeiler and Velasco (2009) established a pick model for a single analyst and found that picks from different institutions could be biased. However, they could not establish a complete error model. To produce a more reliable error model, a more robust technique to test and study the effects of different analysts reviewing the same seismograms must be implemented. The current SPEAR effort is designed to provide a common data set for analysts from varying experience levels to pick. An updated model is generated with each new participant. The current approach is designed to focus on signal quality and analyst experience. While many studies have tried to estimate pick error (Freedman, 1966a, 1966b, 1967a, 1967b and 1968; Buland, 1976; Pavlis, 1986 and 1992; Billings et al.; 1994, Sipkin et al., 2000), our approach can uniquely remove pick errors from the travel-time modeling errors.

2 History of the study

Zeiler and Velasco (2009) studied catalogue data to test the reliability of a single analyst's picks and compare the repeated picks from different institutions reported in the catalogue of the International Seismology Centre (ISC). The single analyst picks were obtained from a local network operator, who reviewed or made each pick on local to near regional events. With over twenty years' experience picking the waveforms, the operator showed consistency within 0.10 seconds when working on waveforms with high signal-to-noise ratio (SNR > 10). Since, the signal quality is not documented in the ISC catalogue, Zeiler and Velasco (2009) were only able to show that there can be a bias based on institution or who is making the picks. The initial study confirmed that the study of pick error could not fully be conducted on pre-existing data base information.

The current iteration of the SPEAR study relies on a common data set of varying types of waveforms for analysts to pick. The waveforms include the same event recorded at various distances, different types of sources, and synthetic pulses with varied SNR. A broader suite of seismograms would be ideal, but limiting the time constraint to pick the seismograms has been a primary goal of the experiment. Currently the data set can be picked in under an hour by most analysts. Since the project is an on going effort, **we encourage interested institutions or analysts to contact the authors** for any related question **and download the common data set and user guidelines** via the link *Download programs and files* from the NMSOP-2 front page.

The SPEAR study has collected picks during two main efforts (Table 1). In each effort the analysts have only picked phases they could identify. Then, after a minimum of six months, the analysts have re-picked the data to test their consistency of picking over time. The commitment to pick the data set twice is not required for the study, but is encouraged to help understand picking precision and its improvement with time and experience.

Table 1 The growth of the SPEAR project from the inception to the last review of all picks that have been submitted by 2011.

Year	No. of Analysts	No. of institutions	No. of waveforms	No. of picks
2006	12	4	25	647
2011	27	7	39	1633

3 Participation in the study and picking procedure

3.1 Assumptions

Within the NMSOP (Chapter 11) one finds general guidelines of how to analyze and pick seismograms. We therefore assume that an analyst is familiar with these techniques and the IASPEI approved standard phase nomenclature (IS 2.1). Moreover, IS 11.4 offers a "Tutorial for consistent phase picking at local to regional distances". However, we do not want to discourage other possible techniques and **welcome participation from all analysts and automated systems.** In part, the study is to collect the methodology employed in analyzing seismograms and we encourage participants to provide their methodology once they have picked the common data set. This would help in generalizing and spreading the awareness of **best practice** and to identify and avoid in future frequently made mistakes. Since, most routine analysis is performed on events, we understand that not all picks can be associated to a particular phase; the naming of phases will need to be addressed with future work. We are interested in where an analyst records the time of an arrival. This picking process is initially performed on the raw seismograms and only then, if an analyst would like to introduce filtering and waveform enhancement to refine their picks, we will allow these procedures to identify additional biases. Windowing or zooming of the seismogram is encouraged to improve the recognition and timing of onset arrivals.

3.2 General procedure

After downloading or receiving the data set, follow the naming convention outlined in Table 2. If you are using a picking program that does not follow this convention, you will need to include a conversion table of the naming convention used when returning your picks. If additional phases are observed, place additional markers as needed and record the observed phase to be reported with the submission of the picks.

Table 2 Naming convention for picks using SAC header variables.

SAC header	Arrival type
Amarker	P or Pg
T0	S or Sg
T1	Pn
T2	Sn
T4	Lg
T5	Rg

In addition to the recorded phase picks from the data set, we record basic information on each analyst. We use the analyst information to derive correlation of picks based on years of experience, average number of seismograms picked in a year, number of seismograms picked in the past six months, institution/organization, and types of seismograms picked (local, regional, teleseismic, explosion). This profile, along with participant names and results, will not be released for comparison of analysts or institutions.

4 Results from SPEAR

The most striking result from the experiment is the trend observed in pick difference versus years of experience (Figure 1). The cut-off of having more than five years experience seems to group the analysts with the lowest pick difference and measurement repeatability. This phenomenon of 5 years experience has been observed in other research as the minimum time to achieve expertise in a given skill (Gladwell, 2008). However, we also note that the years experience is not always indicative of how the overall population will pick and further work is needed to clearly identify analysts that have achieved a competent level of picking.

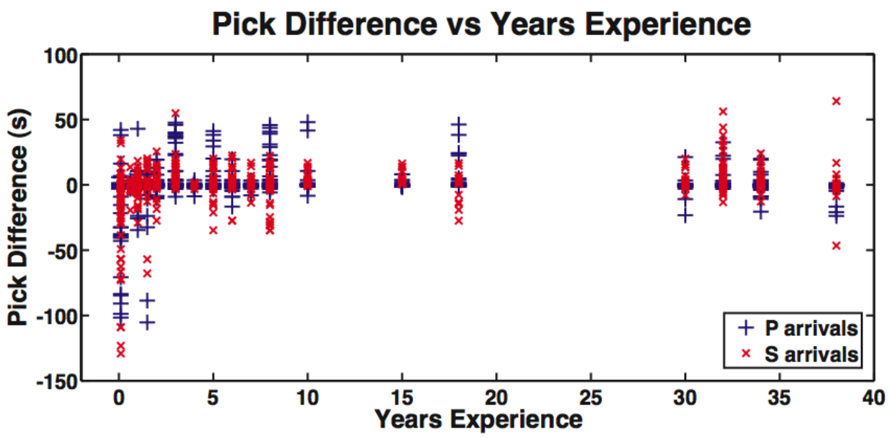

Figure 1 The years experience of each analyst versus the pick difference calculated from the mean arrival time for S and P picks.

4.1 Pick difference

The goal is to determine the arrival time of all phases recorded and we assume each analyst is measuring the same observation. This assumption has resulted in pick differences over a minute (Figure 2). While for standard locating procedures this magnitude of error would be unacceptable, for our purposes we can identify features in the noise field that cause false detections. In many cases, the observations that are over a minute from the mean arrival time have detections that are within seconds of them; this suggests that multiple analysts observed the same feature in the waveform. When participating in the study we encourage analysts to only record observations they can clearly identify. In every case there is at least one event recorded, but if an analyst is unable to detect a significant arrival, we encourage them to not make any measurements. However, if multiple events are observed or if additional arrivals are observed, we encourage the analyst to time and record these detections. In the case an arrival is misidentified we will correct this type of error and add the observation to the correct population.

The pick difference is calculated from the average arrival time for each common observation. When new data is added the average arrival time is recalculated and a new pick difference is established (Figure 2). Since the true arrival time is known for the synthetic seismograms we also compute the pick error for stations 27-40; pick error for these stations can be calculated as an absolute term (Figure 3). Since, the other stations do not have a known arrival, only a relative pick error can be determined. To establish the reliability of the relative pick error, the synthetic measurements will be compared to the mean pick time (average arrival time measured by the analysts) and the known arrival time.

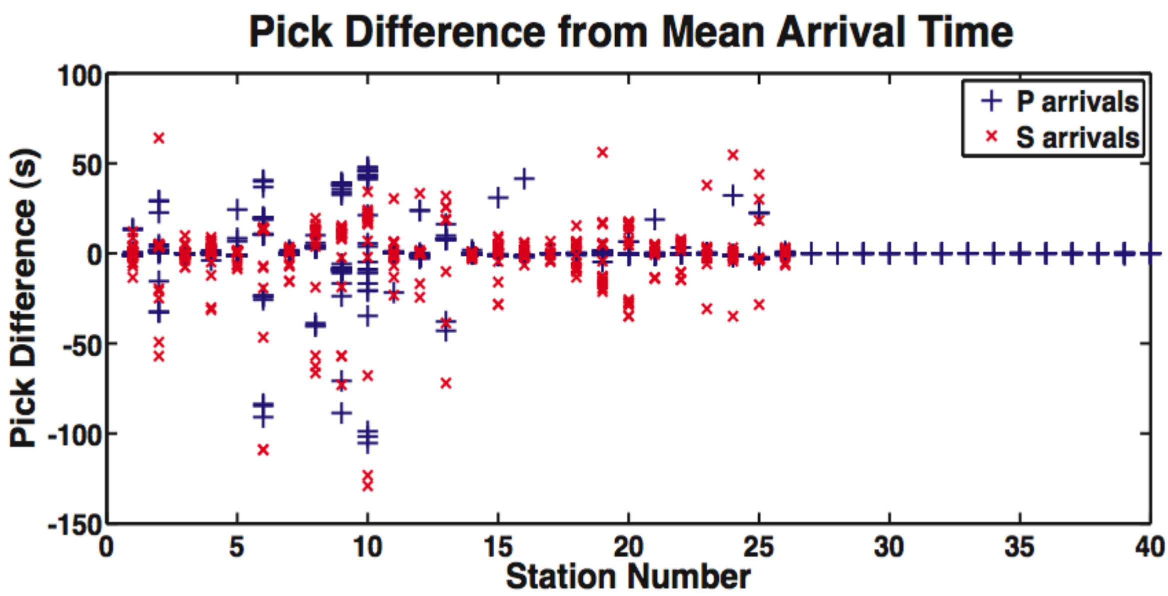

Figure 2 The pick differences of P and S phases computed from the average arrival time. The synthetic traces did not have an S phase to pick.

When you compare the analyst made picks on the synthetic seismograms, you see that overall the analysts made similar picks. However, when you compare the analyst picks to the known arrival time, stations 28, 29, 32, 33, 35 and 38 show a bias of up to 0.5 seconds (Figure 3). All of the late picks are a result of emergent arrivals and analysts choosing to measure the arrival in the coda energy of the phase (the other stations that are biased late are due to the SNR). This presents a difficult variable to the model, since the traditional SNR cannot capture the true amplitude difference between the measured signal and noise.

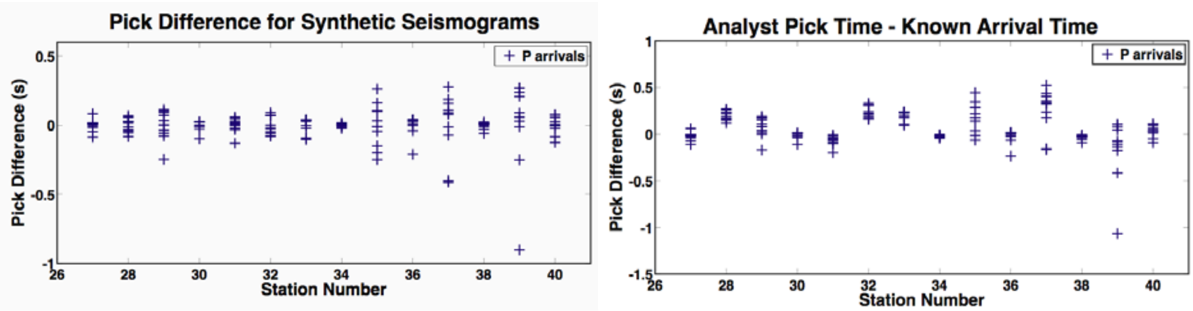

Figure 3 The pick difference for the synthetic pulses, (left) based of the pick difference from the mean arrival time and (right) based off the known arrival time.

Additionally we have placed a repeated seismogram in the data set to determine an analyst's repeatability (Figure 4). Overall the analysts are able to repeat their measurements with a high level of precision, with the type of phase measured and the years experience being the main factors influencing repeatability.

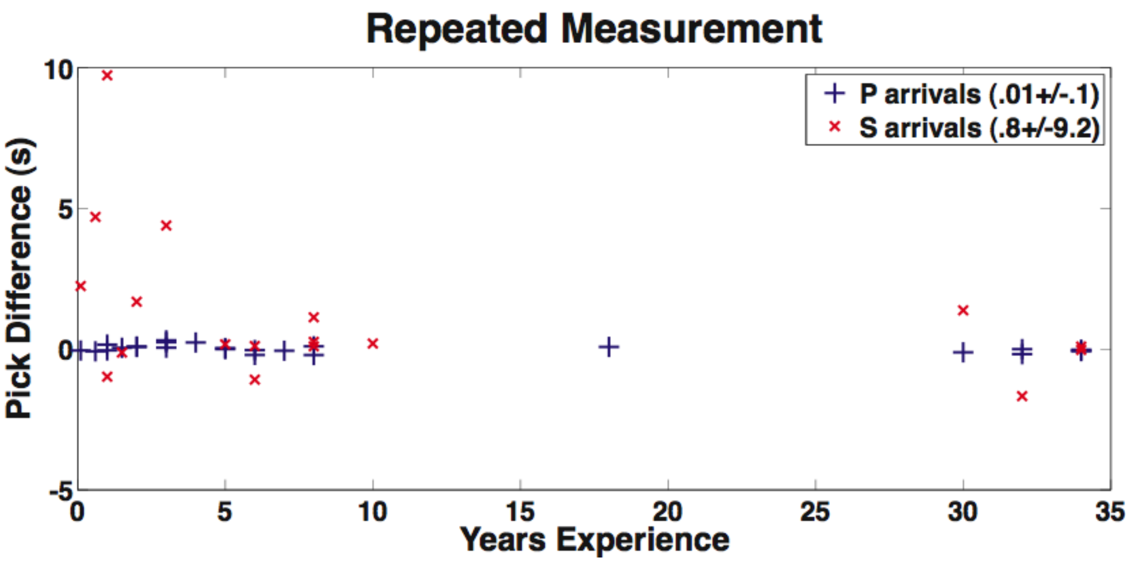

Figure 4 The pick difference between repeated measurements on the same seismogram during the initial round of picking.

The overall goal of the experiment is to derive a pick model based on signal quality. However, no signal quality measurement has been found that adequately models the pick error associated with the process of timing seismic phases. We applied the traditional Short Term Average versus Long Term Average (STA/LTA) using an RMS method of computing the SNR (Figure 5). The STA/LTA model was broke into a two part model of picks with an SNR of 2 or less could have up to a 20 second pick error and all other picks have a pick error of 0.1 seconds. This model is a bit disappointing because all picks under SNR 2 would be thrown out of location solutions. We tested other models to establish signal quality, however, to keep the focus and brevity of this information sheet, this topic needs to be studied further and reported in additional papers.

The best model we found so far to model the pick error is using the peak average ratio of the Power Spectral Density (PSD) of the signal and noise. This approach is similar to the Wide-Spectral band Ratio (WSR) described in Zeiler and Velasco (2009). The PSD ratio provided a broader window to establish a linear relationship between the lower quality signals and then a general pick error for high quality signals (Figure 5). The PSD ratio shows that a signal 10 times larger in amplitude than the back ground noise can be picked with the highest precision.

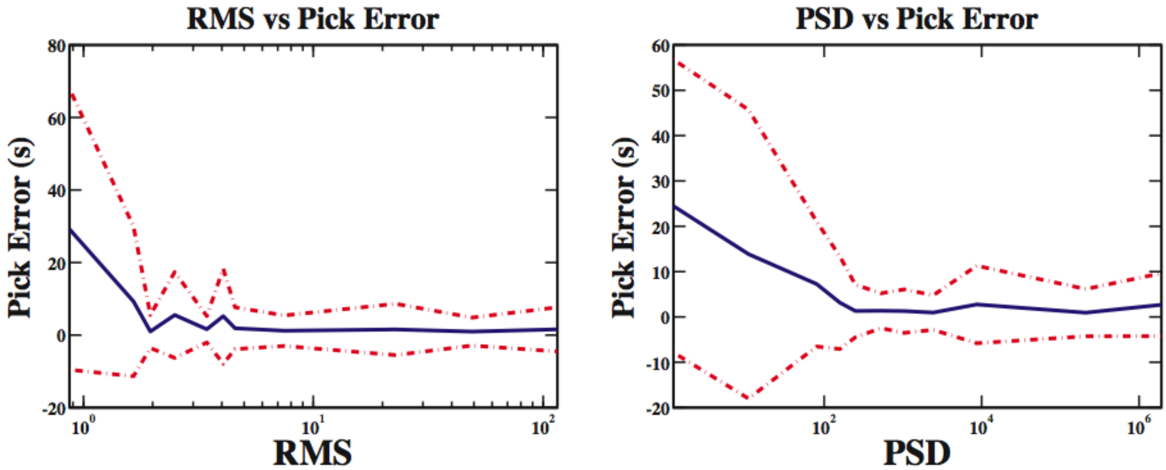

Figure 5 The traditional signal quality of RMS (left) and peak average PSD ratio between signal and noise (right) plotted against the pick error to establish a pick model. The red dashed lines mark the standard deviation of the pick error for the signal quality measurement.

4 Summary

To be successful the SPEAR project needs the continued support of the seismological community by picking the common data set (obtained by contacting the authors or downloading from the NMSOP site). The initial results of SPEAR are promising, but we still need more participants to establish the statistical significance of the measurements. Therefore, we invite all interested seismological institutions, data analysis centers and observatories to participate in the SPEAR project. For **user guidelines** see the **Annex**.

Additional work is also needed to establish the best criteria for quantifying signal quality. The initial measurements are crucial to the event location processes along with the process of event refinement and classification. By fully understanding the errors associated with picking, seismological observatories can better tune their automated systems and improve velocity models. Along with these crucial elements a better understanding of achievable location precision can be assessed from the quality of picks. However, the biggest contribution will be to facilitate the sharing of picks for global studies and work load reduction.

Annex

SPEAR – User Guidelines

If you are a single analyst, you can download the data sets via the link *Download programs and files* from the NMSOP-2 front page. For institutions with more than 2 analysts that will be participating in the study please contact the author for additional data sets. The data needs to be randomly ordered to prevent any learning bias. Once the data sets have been downloaded follow the guidelines in section 3 of IS 11.5. Also, please do not use any of the results provided in IS 11.5 to influence the method of picking.

Depending on the software that is used to analyze the data, common picking procedures need to be adopted. If you are using SAC the provided shell script will allow you to scroll through and pick the events while recording your picks in an output file. However, if you are using a different program the standard procedure is to review the three component data for each station. Only review one stations data at a time and proceed to the next stations data once all picks have been recorded. Do not return to any of the picks to correct or modify them once you have proceeded to the next station. The only enhancement to the waveforms is zooming or windowing the data. Please mark any phases you can identify (following Table 2 in IS 11.5), whether you can name them or not. Generally there is only one event per station, however, if additional phases are observed please measure the arrival time and report them. Stations 27-40 have only a single pulse to identify. Save the measured arrival times and e-mail them back to the first author: cpzeiler@miners.utep.edu.

Once you have recorded the initial picks, feel free to re-pick the waveforms with any filtering or waveform analysis techniques that are common to your observatory. Send these picks in a separate file with a description of the technique used to record the observations. With the pick files also please submit the following information:

- years of experience
- average number of seismograms picked in a year
- number of seismograms picked in the past six months; institution/organization
- types of seismograms picked (local, regional, teleseismic, explosion)
- identify if you would be willing to re-pick the waveforms in 6 months to a year.

Feel free to contact the author for any additional help or clarification of the procedures.

Acknowledgment

The authors would like to thank the analysts that have voluntarily contributed time and energy to pick the common data set. The analysts are highly tasked and their donated time is greatly appreciated. The authors also thank the editor P. Bormann for valuable comments that helped to focus and improve the manuscript as well as his offer to make the SPEAR data set and user guidelines accessible via the NMSOP website.

References

Billings, S. D., M. S. Sambridge, and B. L. N. Kennett (1994). Errors in hypocenter location: picking, model, and magnitude dependence, *Bull. Seism. Soc. Am.* **84**, 1978-1990.

Buland, R. (1976). The mechanics of locating earthquakes, *Bull. Seism. Soc. Am.* **66**, 173-187.

Freedman, H. W. (1966a). The "little variable factor" a statistical discussion of the reading of seismograms, *Bull. Seism. Soc. Am.* **56**, 593-604.

Freedman, H. W. (1966b). A statistical discussion of *Pn* residuals from explosions, *Bull. Seism. Soc. Am.* **56**, 677-695.

Freedman, H. W. (1967a). Estimating the accuracy of source "Parameters", *Bull. Seism. Soc. Am.* **57**, 373-379.

Freedman, H. W. (1967b). A statistical discussion of *P* residuals from explosions: Part II, *Bull. Seism. Soc. Am.* **57**, 545-561.

Freedman, H. W. (1968). Seismological measurements and measurement error, *Bull. Seism. Soc. Am.* **58**, 1261-1271.

Gladwell, M. (2008). *Outliers: the story of success*, Little, Brown and Company, New York, 309.

Leonard, M. (2000). Comparison of Manual and Automatic Onset Time Picking, *Bull. Seism. Soc. Am.* **90**, 1384-1390.

Lomnitz, C. (2006). Three theorems of earthquake location, *Bull. Seism. Soc. Am.* **96**, 306-312.

Pavlis, G. L. (1986). Appraising earthquake hypocenter location errors: a complete, practical approach for single-event locations, *Bull. Seism. Soc. Am.* **76**, 1699-1717.

Pavlis, G. L. (1992). Appraising relative earthquake location errors, *Bull. Seism. Soc. Am.* **82**, 836-859.

Sipkin, S. A., W. J. Person, and B. W. Presgrave (2000). Earthquake bulletins and catalogs at the USGS National Earthquake Information Center, *Incorporate Research Institutions for Seismology Newsletter*, **2000**, No. 1, 2-4.

Steck, L. K., A. A. Velasco, A. H. Cogbill, and H. J. Patton (2001). Improving regional seismic event location in China, *Pure Appl. Geophys.* **158**, 211-240.

Tull, J. E. (1987). SAC reference, tutorial guide for new users, Lawrence Livermore National Laboratory.

Zeiler, C. P. And Velasco, A. A. (2009). Seismogram Picking Error from Analyst Review (SPEAR): Single-analyst and institution analysis, *Bull. Seism. Soc. Am.* **99**, 2759-2770.