# Estimating errors in picked traveltimes


Albena Mateeva

*Center for Wave Phenomena*



**ABSTRACT**

Some inversion procedures (like statics estimation and tomography) require picking of traveltimes. Picked traveltimes always contain some observation errors. Having realistic estimates of these errors is of crucial importance for traveltime inversion. The biggest uncertainty is usually due to the picking algorithm. Automated pickers based on neural networks are very attractive because, besides flexibility, they provide an opportunity for picking error estimation.
This paper presents a procedure for estimating the picking error of first arrivals picked by a neural network.


## Introduction

Having realistic estimates of data uncertainties is of great importance for any inversion (Scales et al., 1994). In seismology (tomography, statics estimation), picked traveltimes are inverted for the velocity-depth structure of the subsurface. No reliable solution can be obtained without a good notion of the errors in traveltimes. Error estimates tell us:

• *which data are most reliable* – in a typical inverse problem data are redundant but inconsistent; for a successful inversion, data must be properly weighted according to their quality.

• *when the obtained velocity-depth model fits the data*, e.g., when computed traveltimes match observed traveltimes "well enough"; we would like to extract maximum information from our data but not to fit the noise.

The biggest uncertainty is usually due to the picking algorithm. Since seismic data volumes are enormous, automated pickers are used. Pickers based on neural networks are especially attractive because, besides flexibility, they provide an opportunity for estimating the picking errors.

So far only first arrivals can be automatically picked. I will concentrate on them, though the suggested approach is applicable to any event.

Instead of the actual signal onset, distinct peaks or troughs are usually picked to minimize the influence of signal and noise amplitude variations. In this paper I will refer to some characteristic peak of the signal as "the first arrival". Each peak on a trace, which is a *potential* first arrival, will be called *"feature"*.

## Errors

Traveltime observation errors can be divided into two groups: *picking error* (due to choosing a wrong feature on a trace as first arrival) and *other errors* (errors that will be present even when the first arrival feature is chosen correctly). These *other errors* are due to noise, modeling (we have to correct the peak time to signal onset time but the very correction contains an error because the wavelet is generally unknown) and band-limited data (traveltimes are only precisely defined for a pulse with infinite frequency range).

The total observation error is given by the sum of all these errors. The picking error is usually the biggest one. Its estimation from the output of a neural network picking program (Hart, 1996) is discussed below.

## Neural Network Pickers

In order to be used for first break picking, a neural network must have been trained to recognize first arrivals, e.g., to classify features as "first arrivals" or "not first arrivals". Example traces are provided to the network. Each trace is considered as a series of features (potential first arrivals). Each feature on the example traces has been labeled by a human interpreter by 0 (not first arrival) or 1 (first arrival). The neural network is taught

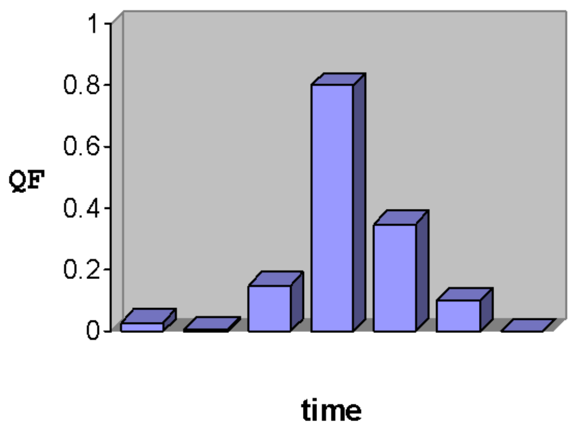


**Figure 1.** Reality – when a trace differs from those in the training data set, the quality factor (QF) distribution along it is not the desired unit spike.

to assign the same labels ("quality factors") to these features. Then it can be used for picking the rest of the traces in a survey. The neural network assigns a quality factor between 0 and 1 to each feature on a given trace. The first arrival is selected as the feature with the maximum quality factor.

## Our Data

The output from the picking program for each trace consists of the first arrival traveltime and its quality factor (the maximum quality factor on that trace).

The closer a quality factor to 1, the greater the similarity between this feature and the first arrivals presented to the neural network during its training. That is how the maximum quality factor carries information about the pick reliability. However this information is not sufficient to determine the picking error. To understand this, let us have a look at the primary output from the neural network (the quality factors of all features on a given trace), instead of at the final output (the feature with the maximum quality factor). In the ideal case the neural network would classify only one feature on a trace as "first arrival" and all other features as "not first arrivals", e.g., on each trace we would like to have one quality factor equal to 1 and the rest – equal to 0. In such a situation the choice of first arrival is doubtless and the picking error is zero (provided the neural network is searching within the right time window).

When a trace differs from those in the training data set, the quality factor distribution along it can deteriorate in two ways:

First, the maximum quality factor can decrease while the rest of the quality factors remain zero. This should warn us that the only interesting event on the trace does not resemble the first arrivals from the training data set. If we are sure the first break is in the examined time window, such a message is not bad news – the selected event must be the first arrival and the picking error goes to zero. Unfortunately, we cannot be always sure that the first break is in the examined window. If there is no first arrival, accepting the pick would lead to a substantial traveltime error.

Second, several features on a trace may have high quality factors. Sometimes this is a sign that the first arrival is not in the examined time window, but even when there is a first arrival, its choice is not reliable (among several strong candidates) and the picking error is high.

Usually these two types of deterioration take place simultaneously, an example of which is shown in Figure 1.

Obviously the maximum quality factor alone is not sufficient to determine the error of a particular trace. Even for maximum quality factor of 1, the error may vary from zero (when there is only one candidate for first arrival) to infinity (when all features on a trace have high quality factors). Error estimates based on the maximum quality factor alone will discriminate poorly between traces of different quality. This is illustrated in Figure 2a. It shows picked versus true traveltimes for an example set of traces. A dot represents a trace. The error bar assigned to each trace is equal to the average error of all traces with the same maximum quality factor. The result is that some of the correctly picked traces have big error-bars while some of the mispicked traces have small error-bars. Such an error estimate will not help us with the relative weighting of the data.

For a better error estimate, *all* quality factors along a trace should be considered somehow. However we do not want to blow up the output volume from the picking program. It would be nice to stick to the old two-parameters-per-trace output – the first arrival traveltime and some quantity from which pick reliability can be estimated. It turns out, this can be achieved if the picking program outputs the **ratio** between the maximum quality factor and the sum of all quality factors along a trace instead of the maximum quality factor itself. I will call this ratio "normalized quality factor". The normalized quality factor is a measure for the sharpness of the quality factor distribution along a trace. Normalized quality factor of 1 corresponds to a spike, no matter what the

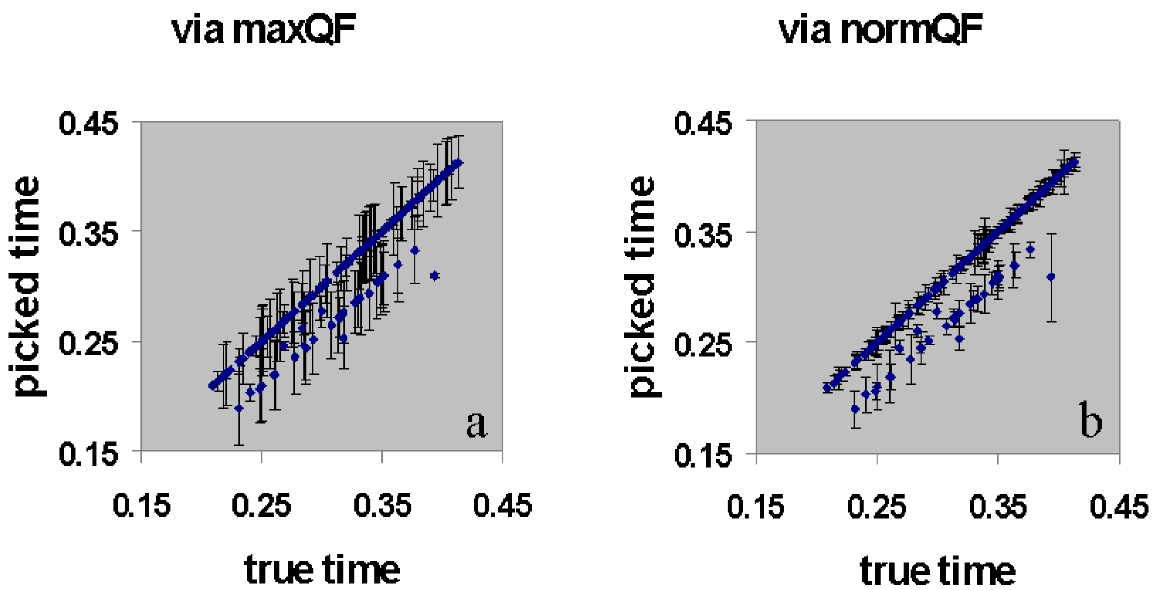


**Figure 2.** First arrivals and their uncertainties estimated via: a) the maximum quality factor and b) the normalized quality factor.

maximum quality factor is. Normalized quality factor of 0 corresponds to infinite number of features with non-zero quality factors.

Let us go back to the example in Figure 2. If we assign to each trace an error-bar equal to the average error of all traces with the same normalized quality factor, we can judge the pick quality quite fair (Figure 2b) – mispicked traces are associated with bigger error-bars than correctly picked traces. Figure 2a and Figure 2b are drawn for the same data set, so the average picking error is the same – just its distribution over the traces is better on Figure 2b.

## Picking Error Estimation via the Normalized Quality Factor

The normalized quality factor can be interpreted as the probability that the picked feature is the first arrival, provided the first arrival is in the considered time window.

This probabilistic interpretation together with a Gaussian assumption about the traveltime uncertainties allows us to derive an estimate of the picking error, effectively averaged over all traces with a given normalized quality factor (Appendix A). Thus, when the first arrival is in the examined time window, its rms picking error is given by:

$$\tau = \sqrt{\frac{-0.125}{\ln(1-p^2)}}T \qquad (1)$$

where $p$ is the normalized quality factor and $T$ is the predominant period of our data (predominant time interval between two successive features). Formula (1) says that for a spike-wise distribution of quality factors along a trace ($p = 1$), the first arrival choice is doubtless ($\tau \to 0$); the broader the distribution, the more uncertain the choice.

Numerical tests show a good correlation between observed and predicted error.

When we know the first arrival is in the examined time window, formula (1) is all we need. However we often do not know whether the first arrival is in that time window. We have to quantify our doubts by big error bars. Calculating them is not easy. Moreover the error cannot be considered as Gaussian anymore (Appendix A) while in tomography it is explicitly assumed to be Gaussian. That is why it would be better to filter out all traces that are at high risk of having no first arrival (traveltime inversion itself would also benefit from such filtering) and, assuming that the rest of traces have first arrivals, use formula (1) to estimate the picking error.

How to recognize the risky traces? One may expect that the chance of having a first arrival depends primarily on the maximum quality factor of the trace since it is a measure for the similarity between the best candidate for first arrival and the first arrivals from the training

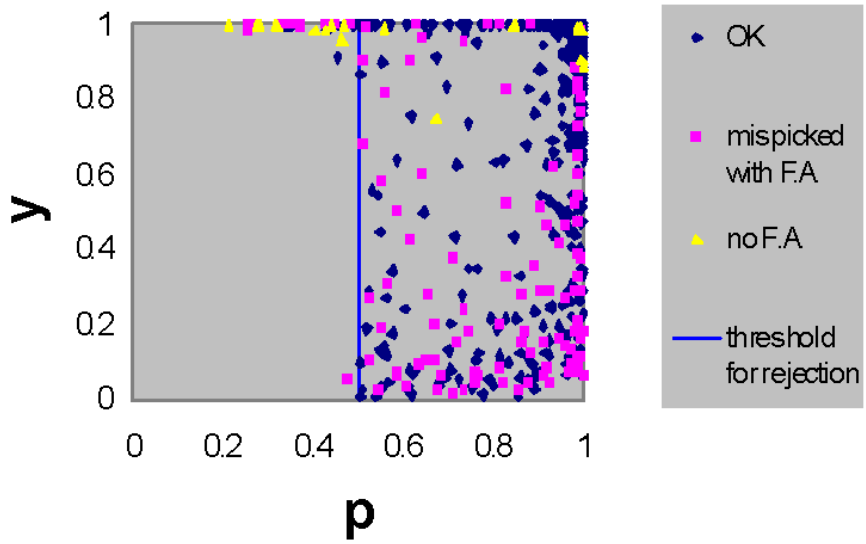


**Figure 3.** Maximum quality factors $y$ and normalized quality factors $p$ of different populations of traces: correctly picked traces (OK), mispicked traces with first arrivals (F.A.) and traces without first arrivals (no F.A.) in the examined time window.

data set. But it turns out the probability for having a first arrival is much more sensitive to *how many* features in that window resemble first arrivals. This information is carried by the normalized quality factor.

Unfortunately, even when both the maximum and the normalized quality factors are available, we cannot filter out all traces without first arrivals since their cluster is not completely separated from the cluster of traces that have first arrivals. Nevertheless, normalized quality factor of about 0.5 provides a reasonable threshold. Most traces with normalized quality factor below 0.5 either have no first arrival or are mispicked (an example is shown in Figure 3). After filtering them out, the picking error of traces that do have first arrivals will be well estimated by (1). The error of traces with normalized quality factor above 0.5 but with no first arrivals will be underestimated. Hopefully, the inversion program will reject them if their traveltimes are too inconsistent with the rest of the data.

## Conclusion

Traveltime inversion requires good estimates of data uncertainties. The most significant errors in traveltimes often arise from the picking procedure.

In this paper, a modified output for neural network pickers has been suggested. It would allow better segregation of picks of different quality and objective error estimation without increasing the output volume from the picking program. An analytical expression for the picking error has been derived. It has already been implemented successfully to statics estimation through diving-wave tomography (Osypov, 1998).

Although only first arrivals were discussed here, the suggested procedure is applicable to any event picked by a neural network.

In future, better picking and error estimation can be achieved if information from adjacent traces is considered.

The picking error is usually predominant but the other types of traveltime errors also need to be well estimated, especially for high quality data with small picking errors. While uncertainties due to modeling and band-limited data strongly depend on the particular experiment, uncertainties due to random noise are always present. That is why the next step should be to develop an objective approach for assessing noise influence on picked traveltimes.

## Acknowledgements

This research was supported by Western Geophysical, Denver. I am grateful for the opportunity to work with them and for the permission to present this paper. Special thanks to Konstantin Osypov and Doug Hart for the helpful discussions.

Also, I would like to thank the reviewers of this manuscript, Vladimir Grechka and Albane Saintenoy (Center for Wave Phenomena, Colorado School of Mines) for their thoughtful and creative suggestions.

## References:

Hart, D., 1996, Improving the reliability of the first-break picking with neural networks: Expanded Abstracts, 66-th Annual Meeting of SEG, Denver, p.1662-1664.

Osypov, K., 1998, A comparative study between 3-D diving wave tomography and head-wave refraction methods: Expanded Abstracts, 68-th Annual Meeting of SEG, New Orleans, p.1222-1225.

Scales, J., Smith, M., 1994, Introductory Geophysical Inverse Theory:Part I: Samizdat Press, pp.193.

## APPENDIX A: Picking Error Estimation via the Normalized Quality Factor

The ratio between the quality factor of a given feature and the sum of all quality factors along a trace can be interpreted as *the probability that this feature is the first arrival, provided there is a first arrival in the considered time window.* I will call this ratio "**p-ratio**". The normalized quality factor $p$ is just the maximum p-ratio on a trace.

If the first arrival is in the considered time window, the picking error $\tau(|is)$ can be estimated from the p-ratio distribution along a trace (the chances of each feature to be the first arrival). However, the final output from the neural network contains only one parameter of this distribution – its maximum (through the normalized quality factor). We have to make some assumptions about the rest. Specifically, we are interested in the standard deviation. To determine it uniquely from the distribution maximum, we have to assume the p-ratio distribution is Gaussian (which is consistent with the Gaussian-error assumption in tomography). Although it is rarely Gaussian along each individual trace, it resembles Gaussian when many traces with the same normalized quality factor are considered. The error $\tau(|is)$, estimated as the standard deviation of this distribution, is effectively averaged over all traces with a given normalized quality factor.

Since features along a trace correspond to discrete time values, the p-ratio distribution is approximated by a histogram. Each column of the histogram is centered on one feature and has a width $T$, where $T$ is the predominant period in our data (the predominant time interval between two successive features). The normalized quality factor $p$ corresponds to the highest column and is approximately equal to the area locked into $[t_0 - T/2; t_0 + T/2]$ bellow the probability mass curve, where $t_0$ is the picked traveltime. It is convenient to choose the time axis so that $t_0 = 0$ (since errors are always considered with respect to the picked feature). Then, using the generic expression for Gaussian distribution, we can write:

$$p = \int_{-0.5T}^{0.5T} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-t^2}{2\sigma^2}\right] dt \tag{A1}$$

from which it follows that:

$$\sigma^2 = \frac{-0.25T^2}{2\ln(1-p^2)} \tag{A2}$$

and consequently:

$$\tau(|is) = \sigma = T\sqrt{\frac{-0.125}{\ln(1-p^2)}} \tag{A3}$$

The above discussion treats the case when the first arrival is in the considered time window. In the general case, the probability for having a first arrival $P(is)$ is not one and the rms picking error is given by:

$$\tau = \sqrt{P(is)\ \tau^2(|is) + [1 - P(is)]\ \tau^2(|isnot)}, \tag{A4}$$

where $\tau(|isnot)$ is the average error we would make accepting the pick if there is no first arrival. When $P(is) \neq 1$, the distribution of potential first arrivals along a trace has a secondary lobe outside the picking window at $\tau(|isnot)$ from the picked feature and traveltime uncertainties are substantially not Gaussian. In most inversion procedures, however, they are explicitly assumed to be Gaussian. In addition, even if all quality factors within the picking window were available, to estimate the probability for having a first arrival $P(is)$ would be difficult (strongly dependent on the particular data set). Since $\tau(|isnot) \gg \tau(|is)$, small variations in $P(is)$ have strong influence in (A4), e.g., the difficulties in determining $P(is)$ make the whole estimate of the picking error $\tau$ unstable. All this suggests that, instead of computing the picking error from (A4), we should try to reject traces that are at high risk of having no first arrival. Assuming that the rest of the traces have first arrivals (e.g., $P(is) = 1$ for them), the picking error $\tau = \tau(|is)$ can be computed from (A3).