Combining arrival classification and velocity model building using expectation-maximization

Cericia Martinez
CSIRO
Perth, WA
Cericia.Martinez@csiro.au

James Gunning
CSIRO
Clayton, VIC
James.Gunning@csiro.au

Juerg Hauser
CSIRO
Perth, WA
Juerg.Hauser@csiro.au

SUMMARY

Probabilistic inversions of wide angle reflection and refraction data for crustal velocity models are regularly employed to understand the robustness of velocity models that can be inferred from these data. It is well understood that the uncertainties associated with the picks of individual arrivals contribute to overall model uncertainty. Typically only a modicum of effort is devoted to quantifying uncertainty in the traveltime picks; a constant noise estimate is commonly assigned to a given class of arrivals. Further, determining the class of arrivals is often left to the behest of the interpreter, contributing additional uncertainty to the data that is both difficult to quantify and may be altogether incorrect. Given the crucial role data uncertainty plays in characterising model robustness, there is a need to thoroughly and appropriately quantify uncertainty in the traveltime data which itself is inferred from the waveform. Here we propose a method that treats arrival or phase classification as part of the velocity model building (inversion) framework using the well-established expectation-maximization (EM) algorithm.

Key words: seismic traveltime; inversion; data uncertainty; data classification; machine learning

INTRODUCTION

Over a number of years, Geoscience Australia (GA) has conducted an extensive seismic acquisition campaign that has resulted in multiple seismic transects across the continent (Kennett et al., 2013). While there have been some targeted velocity model building efforts (Fomin and Goleby, 2006; Finlayson et al., 2002; Goncharov et al., 2000), the broad application of seismic velocity model building techniques has not kept pace with the release of these high-quality datasets. This may in part be due to the labour intensive process of generating the traveltime data necessary for seismic tomography. Often the generation of traveltime datasets is done by manually picking arrivals and assigning phase classifications, or via an experienced interpreter using a semi-automated approach (Reading et al., 2001).

Regardless of how the traveltime picks are generated, reliable measures of data uncertainty are needed to obtain reasonable estimates of the robustness of the resulting velocity models. If the data uncertainty is not well characterised, inversion may result in incorrect representations of the true subsurface velocity including an over or under estimation of uncertainty. Given the crucial role data uncertainty plays in characterising model robustness, there is a need to appropriately quantify uncertainty for traveltime picks. We put forward that a holistic assessment of traveltime pick uncertainty should consider three factors:

- The nature of the onset of the arrival (peak, amplitude, etc.)
- The identification of picks that are ray-theoretical arrivals (timing)
- The class (e.g. reflection or head wave) of arrivals a pick belongs to

Much of the effort in semi-automated picking revolve around the first two items, nature of the onset and quantifying uncertainty in the time of the arrival. Allen (1982) summarises standard methods for onset recognition and arrival picking on individual traces. In the context of ray-tracing based inversion, verifying that picks are ray-theoretical arrivals can be readily done using an appropriate velocity model and forward modelling arrivals. The problem here is that an appropriate velocity model is needed, and this is exactly what we wish to resolve with tomography. Applications around appropriately identifying the class or phase alongside the onset have been largely in the earthquake seismology fields where P-wave arrivals are identified and picked, with less success for identifying and picking S-wave arrivals or converted phases (Reading et al., 2001).

As noted by Finlayson et al. (2002), "The confidence with which seismic phases can be identified places constraints on the robustness of the interpretation". Interestingly, many inversion algorithms assume the arrival classifications (phases) are fully known and correct and any potential mis-identifications are considered outliers (Rawlinson et al., 2014). The trouble with this treatment of the third aspect of data uncertainty is that all data could be misidentified, e.g. an inexperienced interpreter manually handpicking. It is challenging to quantify data uncertainty for this third aspect as it is more of a categorical classification unique to the inverse problem formulation that few have considered (Myers et al., 2009). As an overlooked and neglected contribution to traveltime data uncertainty, we devise a method to incorporate uncertainty of the class a pick belongs to into traveltime inversion. We use the expectation-maximization (EM) algorithm as the overarching mechanism to incorporate arrival classification of traveltime picks into inversion.

TRAVELTIME DATA

Traveltime datasets are created by identifying and selecting the appropriate arrivals of interest. While the arrivals of interest being picked depend on the application and inversion scheme being used, all applications rely on some mechanism to extract the traveltime picks from acquired seismic data.

Seismic traveltime tomography techniques abound for building velocity models by picking and inverting refraction arrivals (Rawlinson et al., 2010). Others have developed techniques to invert other or additional arrival types such as turning waves, Rayleigh waves, or reflections (Rawlinson et al., 2014; Xia et al., 1999).

Inversion techniques tend to utilise multiple arrivals in order to construct a velocity model, potentially requiring multiple traveltime picks from the same seismic trace. An example of this could be the desire to have the first arrival on a shot gather in addition to refracted and reflected phases. If there are multiple reflecting layers and multiple shot gathers, it could be challenging to appropriately pick the same reflection arrivals across all shot gathers. The uncertainty in appropriately identifying the desired pick contributes to data uncertainty and needs be accounted for in an inversion. Further, this uncertainty exists for picks that are manually or semi-automatically generated. Often when it comes time to invert traveltime data, the uncertainty in the arrival types are ignored or neglected and the available data are assumed to be absolutely correct with little error.

With inversion algorithms readily available to invert one or more types of traveltime arrivals (Zelt and Ellis, 1988), it is worthwhile to consider how this form of data uncertainty contributes to uncertainty in the velocity model and account for it. For this reason, we propose to incorporate traveltime pick classifications into the inversion itself. We do this by employing the expectation-maximization algorithm.

METHODS

The expectation-maximization (EM) algorithm is a general technique used when there is incomplete or unobserved data (Dempster et al., 1977). In the context of our problem, the incomplete data is assumed to be the class of the traveltime as well as the layer it is associated with. In this way, the expectation step is used to classify the membership (arrival type and layer) of each traveltime and the maximization step consists of determining the general least-squares model solution.

For our purposes, multiple iterations of the expectation and maximization steps are needed to determine both the arrival classifications and the velocity model. In this sense, the two steps combined are analogous to a standard iterative inversion. In the description that follows, this means that a cycle through one expectation and one maximization step consists of one EM inversion iteration.

Parameterization

To demonstrate our proposed method, we assume a simple 1D layered velocity model. The layer velocity, v_i, is represented as slowness such that s = 1/v and $\mathbf{s} = [s_0, s_1, s_2, ..., s_m]^T$ is the constant velocity earth with *m* layers, including a halfspace, composing the 1D velocity model where layer thickness is assumed to be known.

We use standard 1D ray tracing to generate traveltimes for three types of arrivals from a given interface associated with a velocity discontinuity. The three arrival types are: direct, reflection, and refraction. We assume these basic arrival types for now in order to demonstrate the ability to differentiate and classify the arrival types while simultaneously building a velocity model.

Complete Data

We next provide a definition for the notion of *complete data* in the context of our EM inversion algorithm. A complete observation has the following information:

- traveltime, t
- offset (lateral location of observed traveltime)
- arrival classification (e.g. reflection, refraction)
- layer correspondence (e.g. reflection from interface at bottom of first layer)

A complete dataset would have each of these four pieces of information for every traveltime arrival. In reality, we are able to more readily determine the first two data items (traveltime and offset) while the arrival classification and layer correspondence are generally estimated and based on interpreter experience. These latter two pieces of information are the incomplete data which we use the EM algorithm to quantitatively determine while building a velocity model.

More formally, we assume the traveltime data, $\mathbf{t} = [t_1, t_2, \ldots, t_n]^T$, each belong to a finite unobservable set of states that are the three arrival types (direct, reflection, refraction). The incomplete data **Z** is a matrix of indicator variables for the unobservable state that **t** belongs to, which is the arrival classification and layer correspondence. For each datum t, there is a membership vector, $\mathbf{Z}_{i:}$, of the unobservable states. The complete data are then **y** and **Z** consisting of each possible t and corresponding arrival classification.

The traveltime data can be calculated via $y_{il} = f_{il}(s)$ for offset *i* and wave/layer-type *l* where *l* is an index over all possible wave and layer correspondences and **s** is the slowness model. The Jacobian $X_{il,j} = \partial f_{il} / \partial s_j$ for all possible wave types is required for the M-step below.

Expectation step

In the expectation step, the memberships (arrival classification and wave-type correspondence) are computed. The **Z** coefficients provide a measure of how likely it is that each traveltime i is generated by wave-type/layer *l*. They are defined by

$$Z_{il} = \frac{exp\left(-\frac{(f_{il}(s) - y_i)^2}{2\sigma_d^2}\right)}{\sum_p \ exp\left(-\frac{\left(f_{ip}(s) - y_i\right)^2}{2\sigma_d^2}\right)}$$

where p is an index over all possible wave-type and layer correspondences. During each expectation step the membership matrix **Z** is updated based on the current velocity model. For each traveltime (t_i) the vector $\mathbf{Z}_{i:}$ contains values corresponding to the likely arrival/layer membership of the traveltime. Near convergence, usually one of the memberships z_{il} is close to one.

Maximization step

The maximization step consists of a generalized least squares minimization to obtain an update to the velocity model, where we use a Gauss-Newton method based on local linearization of travel time computations. The **Z** membership coefficients are used as weights in the minimization via a reweighted Gauss-Newton algorithm. The forward traveltime model at the *k*th iteration taken as $f(\boldsymbol{s}) = X^{(k)} (\boldsymbol{s} - \boldsymbol{s}_k) + f(\boldsymbol{s}_k)$, and the membership reweighting amounts to replacing each datum and Jacobian row by a block of entries for each wave mode, thus

$$y_i \to \tilde{y}_{\{il\}} = \sqrt{z_{il}}y_i$$
$$f_i \to \tilde{f}_{\{il\}} = \sqrt{z_{il}}f_{il}$$
$$X_{ij} \to \tilde{X}_{\{il,j\}} = \sqrt{z_{il}}X_{ij}$$

The Gauss-Newton update is then of form

$$\boldsymbol{s}_{\{k+1\}} = \left(\tilde{X}'C_d^{-1}\tilde{X}\right)^{-1}\tilde{X}'C_d^{-1}(\tilde{y} - (\tilde{f}(\boldsymbol{s}_k) - \tilde{X}\boldsymbol{s}_k))$$

with $C_d = \sigma_d^2 I$ representing the augmented data covariance with the identity matrix given as I. This expresses the uncertainty of the traveltime, **y**. Only a few Gauss-Newton iterations are usually needed within each M-step. After final convergence of the EM algorithm, the slowness model is available in **s**, and the dominant membership for each datum identifies the relevant wave-modes/layers.

EXAMPLES

As an example, we generate a synthetic shot gather and traveltimes for a four layer model. The true slowness model is **s** = 1/[1.5, 2.0, 3.5, 6.0] km/s with corresponding layer thicknesses of [0.5, 1.0, 1.752, 2.248] km. The true traveltimes and arrival classifications are shown in Figure 1 for 20 offset locations spaced evenly from 0.1km to 10km. The colour of each arrival indicates the corresponding layer. Errors are simulated from a normal distribution with standard deviation of 0.02s and added to the true traveltimes in order to generate observed traveltimes for this example. The noisy traveltime data are shown in Figure 2.

The recovered traveltime classifications from the EM inversion are shown in Figure 3. Figure 4 emphasizes the difference between the EM arrival classifications and the true classifications. The major differences in arrival classifications occurs where the traveltimes overlap. The higher noise levels also contribute to the overlap in traveltimes. The membership probabilities for the arrivals at the far offset (10km) are shown in Figure 5. The true arrival class is denoted by the colours while the probability of belonging to each classification is shown by the membership probability. In all cases, the highest membership probability corresponds to the true arrival classification.

The recovered velocity model from the EM inversion is shown in Figure 6. The true velocity model is plotted in red while each iteration is plotted in shades of grey, with the final model underlying the true red line. After a single EM iteration, the recovered velocity model is close to the true model with only three EM iterations needed to converge to roughly the true velocity model.

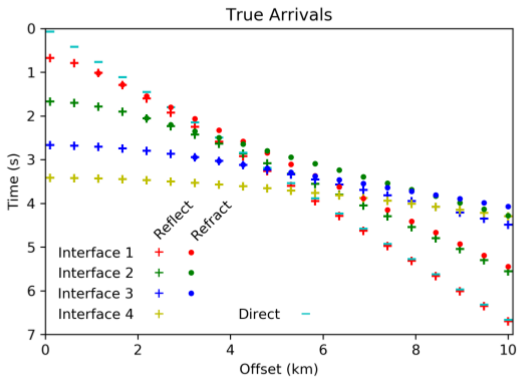

Figure 1. True arrival classifications for four layer model. The colour of each arrival indicates the corresponding layer.

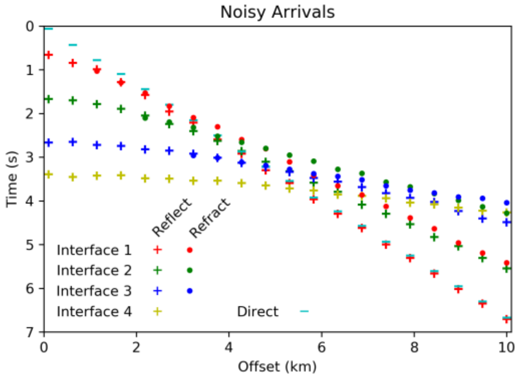

Figure 2. Noise affected traveltime data.

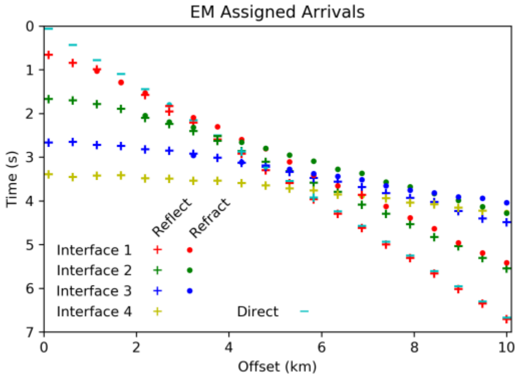

Figure 3. Recovered arrival classifications for four layer model.

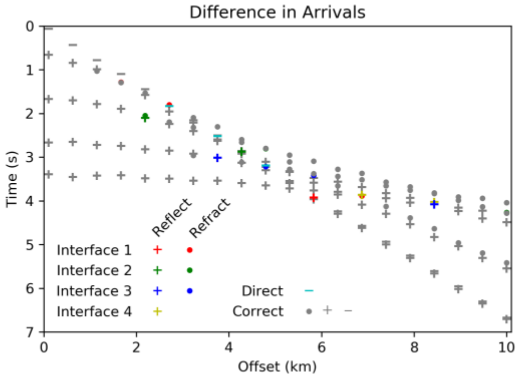

Figure 4. Difference between the recovered arrival classifications and true classification for our four layer model.

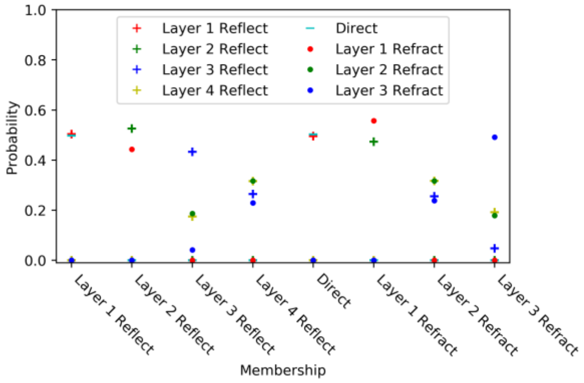

Figure 5. Membership probabilities (Z) for the arrivals at 10km offset. True arrival classifications are noted by the colour.

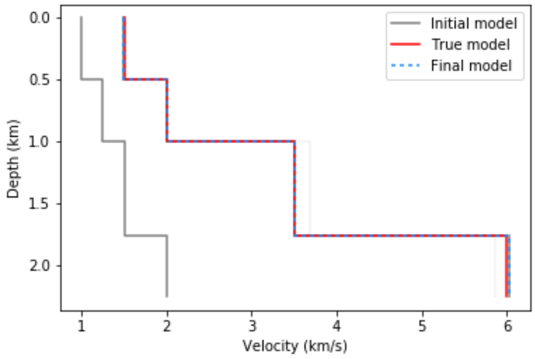

Figure 6. Velocity model recovered by applying our EM inversion for our four layer model.

DISCUSSION

The example demonstrates the potential of combining arrival classification and velocity model building. Other sources of data uncertainty will play a role in the ability to use our method for arrival classification. In particular, the traveltime pick (t) and uncertainty associated with the picked time (C_d) are important. The robustness of our method is predicated on the ability to pick traveltimes with a reasonable level of uncertainty. In the case of our 1D results, tests have shown that if the traveltimes are highly uncertain, the arrival classifications and velocity model are more likely to be in error. This is in part due to the fact that some arrival traveltimes may be coincident in time when timing errors are high. The initial velocity model may also be a factor in the inability to recover fully correct classifications. It may be that the initial model is close to a local minimum and the errors are significant enough that the EM inversion settles on the local minima as the solution. This could be mitigated by using a layer-stripping technique where the velocities in each layer are solved for in sequence starting with the shallowest layer rather than solving for all layer velocities simultaneously as we do here.

CONCLUSIONS

We have developed a method to incorporate additional uncertainty in traveltime tomography data into an inversion framework. As a step in the inversion process that is often determined manually by an interpreter, we propose to have the traveltime arrivals classified as part of the inversion. A variation on the Expectation-Maximization algorithm has provided the framework to do this. The expectation step is used to classify the traveltime arrivals while the maximization step generates the velocity model via a generalized least squares model update. A simple example illustrates the feasibility of such a method to classify traveltime arrivals while generating a velocity model.

REFERENCES

Allen, R., 1982, Automatic phase pickers: their present use and future prospects: Bulletin of the Seismological Society of America, 72, S225–S242.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977, Maximum likelihood from incomplete data via the EM algorithm, 39.

Finlayson, D. M., R. J. Korsch, R. A. Glen, J. H. Leven, and D. W. Johnstone, 2002, Seismic imaging and crustal architecture across the Lachan Transverse Zone, a possible early cross-cutting feature of Eastern Australia: Australian Journal of Earth Sciences, 49, 311–321.

Fomin, T., and B. R. Goleby, 2006, Lessons from a joint interpretation of vibroseis wide-angle and near-vertical reflection data in the northeastern Yilgarn,Western Australia: Tectonophysics, 420, 301–316.

Goncharov, A., G. O'Brien, and B. Drummond, 2000, Seismic velocities in the north west shelf region, australia, from near-vertical and wide-angle reflection and refraction studies: Exploration Geophysics, 31, 347–352.

Kennett, B. L. N., E. Saygin, T. Fomin, and R. Blewett, 2013, Deep crustal seismic reflection profiling, Australia: 1976- 2011.

Myers, S. C., Johannesson, G., and Hanley, W., 2009, Incorporation of probabilistic seismic phase labels into a Bayesian multiple-event seismic locator. Geophysical Journal International, 177(1), 193–204.

Rawlinson, N., A. Fichtner, M. Sambridge, and M. K. Young, 2014, Seismic Tomography and the Assessment of Uncertainty: Advances in Geophysics, 55, 1–76.

Rawlinson, N., S. Pozgay, and S. Fishwick, 2010, Seismic tomography: A window into deep Earth: Physics of the Earth and Planetary Interiors, 178, 101–135.

Reading, A. M., W. Mao, and D. Gubbins, 2001, Polarization filtering for automatic picking of seismic data and improved converted phase detection: Geophysical Journal International, 147, 227–234.

Xia, J., R. D. Miller, and C. B. Park, 1999, Estimation of near surface shearwave velocity by inversion of Rayleigh waves: Geophysics, 64, 691–700.

Zelt, C. a., and R. M. Ellis, 1988, Practical and efficient ray tracing in two-dimensional media for rapid traveltime and amplitude forward modeling: Canadian Journal of Exploration Geophysics, 24, 16–31.