Cent. Eur. J. Geosci. • 1(2) • 2009 • 227-231
DOI: 10.2478/v10085-009-0016-5

VERSITA

Central European Journal of **Geosciences**

Bayesian analysis of data and model uncertainty in 3D seismic travel-time tomography

Research Article

Michael Arvanitis[1]* Panagiotis Vlamos[2]

1 *Geomorph Instruments, PO Box 3125, 10210, Athens, Greece*

2 *Ionio University, Department of Informatics, Pl. Tsirigoti 7, 49100, Kerkyra, Greece*

Received 3 December 2008; accepted 15 April 2009

Abstract: We present a probabilistic analysis of seismic travel-time equations using the Bayesian Method. The assessment of models and data is crucial in 3D seismic travel-time tomography, and a method quantitatively assess the quality of both the data and the model is necessary in order to attain the most realistic results. The Bayesian method that we propose here is more effective than the frequentist approach, both in analysis time and uncertainty minimization, when processing large sets of tomographic data.

Keywords: uncertainty analysis • Bayesian methods • seismic tomography

1. Introduction

Tomography, in a geophysical context, is a method to interpret the earth's interior, and is based on representative mathematical models of physical systems [1]. Models are however approximate representations of an incomplete set of data. It is therefore necessary to test the reliability of the data and model being used in order to obtain realistic final models. Data and model uncertainty analysis quantitatively defines the uncertainty in the measured data, the accuracy in the modelling of the physical system, and the realism of the final model. In practice, for examination of data uncertainties is used the chi-square test which effectiveness is problematic [2]. It is remarkable that in the great majority of cases, this examination is either ignored, by simply calculating the difference between observed and calculated travel times, or is partially solved by applying a priori data errors.

2. Inverse Problem

In the following account, we use the symbol d for data, where $d \in \mathfrak{R}^n$ ($\mathfrak{R}$ is the set of the real numbers and n is the dimension of the data); m for an infinite dimensional model vector, where $m \in S$ and S is the infinite dimensional space of models; and, δ and μ for the data and model estimators, respectively. Data estimation involves mapping elements from the model space to the data space. Supposing that we have collected a quantity of n data, we determine a function, termed here the forward function ϕ, which maps models into data space. As this function is always an approximation it introduces a systematic error k, where k represents a vector in a n-dimensional space. The equation which connects data and model is as follows:

*E-mail: mike@geomorph.gr

$$d = \phi(m) + \varepsilon + k, \quad (1)$$

where ε is a vector in a n-dimensional space resulting from random measurement errors. In the case of a linear forward function, Equation (1) can be expressed as an infinite series of coefficients

$$\begin{aligned} m &= \{m_i\} : \\ d &= Lm + \varepsilon, \end{aligned} \quad (2)$$

where L is the linear forward function with $L = \mathfrak{R}^{nm}$. In Equation (2) we ignore the systematic error k because we consider L as accurate. According to Trampert and Snieder [3], we can rewrite Equation (2) considering a finite vector m', which contains the first i coefficients, and the sequence N_∞, which contains all subsequent coefficients:

$$d = L'm' + L_\infty N_\infty + \varepsilon. \quad (3)$$

We can consider the factor $L_\infty N_\infty$ as the systematic error k. As L maps an infinite space into a finite one, the kernel of the forward operator ϕ, Ker(ϕ), will be non-trivial, resulting in multiple models that fit the data. This means that the solution is non-unique, a crucial matter for seismic tomography. A solution is to introduce prior information in order to limit the range of suitable models.
Two probabilistic approaches exist for data and model uncertainty analysis; the Bayesian (e.g. [4]) and the frequentist (e.g. [5]). The Bayesian method, the most popular approach for geophysical inverse problems, is based on an a priori probability model, where we know about the model before collecting/applying any data. The frequentist approach is rarely applied and is based on the interpretation of the probability from the frequency of outcomes. While the Bayesian method relies on pre-data information, or how proximate to the reality a hypothesis can be, the frequentist method uses post-data information. In seismic surveys, especially in very shallow experiments, the selection of the prior model may be a very difficult task, because of velocity variations. As prior knowledge is the conversion of deterministic information into probability is challenging. In practice, the prior model, through a series of calculations, tends to regularize the a posteriori solution [2]

3. The Bayesian methodology

In order to address a Bayesian inversion, firstly, we have to choose how the prior information is going to be represented and, secondly, how the a posteriori information will be estimated (e.g. the constraints, parameters and velocity model to be used). The prior information is a subjective choice usually based on individual experience of a certain problem with a specific level of uncertainty. This decision is particularly difficult to make in high-dimensional problems. A more practical approach to this problem is the use of observations to estimate the prior model. This technique, known as empirical Bayesian, is an approximation to a full Bayesian analysis (e.g. [6]). A full hierarchical Bayesian analysis introduces additional dependencies on parameters and this process ends when all the parameters are considered as known. Conversely, the empirical Bayesian analysis ends when the last parameters cannot ever be considered as known.

4. Estimator's performance and risk

Each estimator depends on prior data, and we presently cannot evaluate how much prior information causes error in the a posteriori information [7]. The risk depends on the chosen model. We define the loss function $L(\mu,m)$ as a measure of how good the estimator μ is for the model m. For any other model n, we have $L(m,n) \geqslant 0$ and $L(m,m) = 0$. There are many loss functions [8], but the most common is the square-error loss function:

$$L(\mu, m) = (m - \mu)^2 \quad (4)$$

and the ℓ_p norm loss function:

$$L(\mu, m) = |(m - \mu)|^p. \quad (5)$$

The risk, $R(\mu,m)$, of the estimator μ is given by:

$$R(\mu, m) = E_p\left(L(\mu, m)\right), \quad (6)$$

where p is the probability distribution and E_p is the expectation operator according to the probability distribution. Equation (6) is also known [7] as the weighted average of the risk.

5. Comparing different estimators

To compare different estimators we can use either posterior risk, by taking the expected loss with respect to the posterior distribution $p(m|d)$, or the weighted average of the risk using the prior distribution as weight function. The posterior risk (r) is given by:

$$r_{m|d} = E_{m|d}\left(L(m, \mu(\delta))\right) \quad (7)$$

The weighted average of the risk, known as Bayes risk [2], is given by:

$$r_\rho = E_\rho R(m, \mu), \tag{8}$$

where ρ is the prior model distribution. The estimator with the smallest risk is called Bayes estimator. If f is the joint distribution of models and data, we can obtain the distribution of the data by integrating f with respect to all the models:

$$h(d) = \int_M f(m, d) dm, \tag{9}$$

where M is the space of models. The Bayesian posterior distribution, that is, the conditional distribution of model m given d data, according to Bayes' theorem [9] is given by:

$$p(m|d) = f\frac{p(d|m)\rho(m)}{h(d)}. \tag{10}$$

Equation (10) updates the prior information according to the data.

6. Application to a seismic profile

In a typical tomographic problem, the aim is to interpret the observed travel times in order to reconstruct a model of the examined volume's interior. For each observation, the travel time (t) is given by:

$$t = \int_L \frac{1}{u} dL, \tag{11}$$

where L is the ray path and u is the velocity. The coefficient $1/u$ is also known as slowness [10]. Given the time t, we want to calculate the velocity u or a functional thereof. The velocity u is not constant because the rays are not straight and, so, the problem is non-linear. From N observations we take the following system:

$$Ms = t, \tag{12}$$

where s is the m-dimensional slowness vector denoting the slowness in the j-th cell, t is the n dimensional time vector denoting the total travel time of the i-th ray and M is the $(m * n)$ dimensional matrix of l_{ij}, where l_{ij} is the length of the i-th ray path through the j-th cell. In most cases, matrix M is sparse, with only a few elements being non-zero, and it is highly overdetermined. Under these conditions, $\gamma(t) \gg \psi(s)$, where:

$$\gamma(t) = \dim\left(\left\langle t^{(1)}, \ldots, t^{(m)} \right\rangle\right)$$

and

$$\psi(t) = \dim\left(\langle s_1, \ldots, s_n \rangle\right). \tag{13}$$

The forward modelling operator ϕ(m), which maps models onto data, is the product of the matrix M and the slowness vector s. Let o t_i^o be the travel time for the i-th observation and t_i^c be the i-th calculated travel time. The term t_i^c is a function of slowness s. If e_i is the noise for the i-th observation, then:

$$t_i^o = t_i^c + e_i. \tag{14}$$

The noise term e_i, though, theoretically has a zero mean, and a variance σ_i^2. As we have already mentioned, the system defined by Equation (12) is overdetermined. However, it is also underdetermined because rank *(M)*< ψ*(s)*. Thus, no s exists that exactly satisfies Equation (12). In the case where rank*(M)=*ψ*(s)* Equation (12) has a unique solution. Our approach is to determine the slowness values, s, that make the misfit function smaller than a user-defined tolerance limit. In other words, we seek a vector s, which minimizes the

$$PMs - t, \tag{15}$$

where PI is the Euclidean norm. Since rank *(M)*< ψ*(s)*, there are infinite solutions of s that satisfy the criterion of Equation (15), which is known as the least-squares solution. For a detailed review of the least-squares solution see A. Van der Sluis and H.A. Van der Vorst [11]. A commonly used misfit function is the chi-square function [12]:

$$x^2(s) = \frac{1}{N} + \sum_{i=1}^{N} \left(\frac{t_i^c + t_i^o}{\sigma_i} \right)^2. \tag{16}$$

Another misfit criterion is the normalized chi-square function [12], given by:

$$x^2(s) = \frac{1}{N} + \sum_{i=1}^{N} \left(\frac{\sum_i L_{ij} - d_i}{\sigma_i} \right)^2, \tag{17}$$

where, m_j is the j-th model, and d_i the data for the i-th datum.

We discard all the values of slowness, s, that do not fit the data for the condition $x^2 = 1$.

7. Case study

In order to test the efficiency of our approach we used both synthetic and real data sets. The real data set was the first-break picks from 200 records gathered during geoarcheological examination [13]. The shot interval was 6 m and the receiver interval was 3 m. Each shot recorded 120 channels. After picking the first breaks, an initial velocity model is estimated. Using the ray-tracing algorithm [14, 15], an initial set of travel times is calculated as well as the ray paths. The travel-time misfit is estimated and the velocity model is changed.This process is iterated until the misfit lies within acceptable limits, i.e. when $x^2 = 1$. We used the Bayesian and the frequentist approach on the same data sets in order to test their efficiency. We applied this approach to an existing geoarchaeological problem [16].

Previous excavations at some parts of the area show good agreement between our results and the existing subsurface structure. From the excavations we see the presence of three layers, which comprise a low velocity sedimentary layer underlain by weathered crystalline rock of several metres in thickness, which is in turn, underlain by compact crystalline rock. The tomographic approach using Bayesian estimators gave very close results to the archeological excavation (Figure 1) and managed to process the data in a very efficient manner. The processing procedure lasted some minutes on a Unix based system. The typical calculation time, approximately ten minutes, was half of that needed for the frequentist approach.

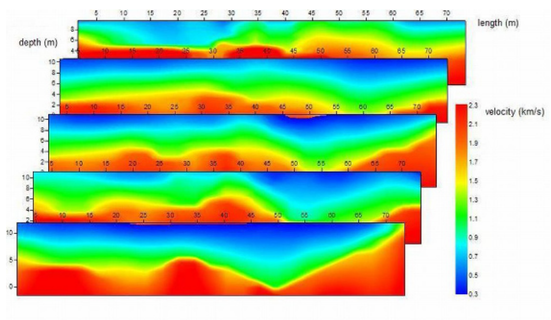

Figure 1. The tomographic results using the Bayesian approach.

Although the frequentist approach also defined three main layers, the tomographic results did not correspond well with the excavation data (Figure 2). Moreover, during the calculation, several delays and breakdowns occurred as the frequentist algorithm cannot efficiently process large data sets. We therefore had to break the data into smaller sets. The Root Mean Square (RMS) error using the Bbayesian approach is approximately 1, while the same data with the frequentist approach gave an RMS error of 1.56. Similarly, the Bayesian method gave a chi-squared value of 1.1, and the frequentist method a chi-squared value of around 2. Our new Bayesian method also corresponded with other excavations in the area according to Arvanitis and Karastathis [1].

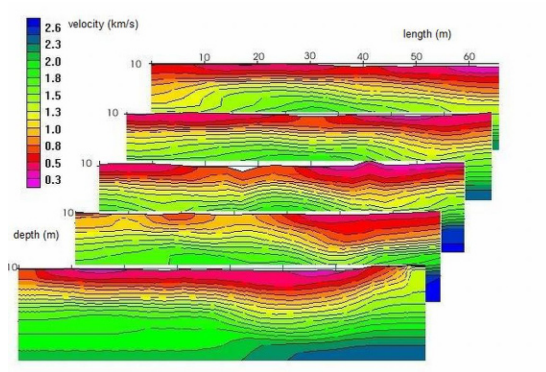

Figure 2. The tomographic results using the frequentist approach.

8. Conclusions

We carried out a probabilistic analysis of seismic travel-time equations using the Bayesian approach. This approach was shown to more efficiently process large tomographic data sets when compared to the frequentist method. We applied both methods to a geoarcheological excavation site in Greece, and found that the Bayesian method resulted in noticeably smaller errors in model estimation. We therefore propose that the Bayesian approach provides an improved method to quantify both model and data uncertainties in 3D seismic travel-time tomography, particularly when applied to large data sets.

References

[1] Arvanitis M., Karastathis V., Uncertainty Analysis of Current 3D refraction tomography on shallow depth applications, Geophysical Research Abstracts, 2003, 5, 05284

[2] Scales J.A., Tenorio L., Prior information and uncertainty in inverse problems, Geophysics, 66, 2001, 389-397

[3] Trampert J., Snieder R., Model Estimations Biased by Truncated Expansions: Possible Artifacts in Seismic Tomography, Science, 5253, 1996, 1257-1260

[4] Tarantola A., Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation, Elsevier, New York, 1987

[5] Parker R.L., Geophysical Inverse Theory, Princeton University Press, 1994

[6] Tang W., Zhang C., Empirical Bayes methods for controlling the false discovery rate with dependent data. In: Regina Liu, William Strawderman and Cun-Hui Zhang (Eds.) Complex Datasets and Inverse Prob-

lems: Tomography, Networks and Beyond, Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2007

[7] Kass R.E., Wasserman L.A., The selection of prior distributions by formal rules, J. Am. Stat. Assoc., 91, 1996, 1343-1370

[8] Barnard G., Studies in the History of Probability and Statistics: IX. Thomas Bayes' Essay Towards Solving a Problem in the Doctrine of Chances, Biometrika, 45, 1958, 293–295

[9] Rao R., Kumar C., Prakasham R., Hobbs P., The Taguchi methodology as a statistical tool for biotechnological applications: A critical appraisal, Biotechnology Journal, 3, 2008, 510–523

[10] Press W.H., Teukolsky S.A., Flannery B.P., Numerical recipes: the art of scientific computing, Cambridge University Press, 1994

[11] Van der Sluis A., Van der Vorst H.A., Numerical Solutions of large sparse algebraic systems arising from tomographic problems: In: G. Nolet (Ed.) Seismic Tomography, D. Reidel Publications Co., 1987

[12] Simon M., Probability Distributions Involving Gaussian Random Variables, Springer, New York, 2002

[13] Stavrinos P., Arvanitis M., Finslerian Generalisation of Ray-tracing method in seismic traveltime tomography, Journal of the Calcutta Mathematical Society, 4, 2005, 133–138

[14] Hobro J.W.D., Three-dimensional tomographic inversion of combined reflection and refraction seismic travel-time data, PhD Thesis, Churchill College, University of Cambridge, 1999

[15] Hobro J.W.D., Singh S.C., Minshull T.A., Three-dimensional tomographic inversion of combined reflection and refraction seismic traveltime data, Geophys. J. Int., 152, 2003, 79-93

[16] Arvanitis M., 3D seismic traveltime refraction tomography and application on a geotechnical problem with high veloicity contrasts, PhD thesis, University of Patras, Patras, Greece, 2008