Bull. Nov. Comp. Center, Math. Model. in Geoph., 10 (2005), 35–43

Numerical algorithms and results of experiments to determine the parameters of the borehole bottom and medium

M.S. Khairetdinov, O.K. Omelchenko, G.F. Sedukhina,
G.M. Voskoboynikova

Abstract. The basic steps of the method of automatic location of borehole sources have been developed. This has been done by measuring the initial wave characteristics used to solve the inverse problem of determining the source coordinates and velocities of near-borehole media.

Introduction

The accuracy of determining seismic parameters of the borehole environment — the in-seam seismic velocities and the geometry of boundaries — is mainly determined by the data for the borehole trajectory in the three-dimensional space. Both of these problems are interrelated: the accuracy of solution to the latter depends on that of the former. The determination of the borehole trajectory, in particular, the inclinometry of inclined boreholes, is rather difficult. It is well known that a solution becomes more complex when the problem is solved in the real time mode. To solve it, the new algorithms and programs for automatic measurement of the arrival times of direct and reflected waves have been developed. This measurement uses the data on the recording of signals from a source by the areal observation system and by solving the inverse problem of reconstructing parameters of the borehole source: the time in the source and its coordinates, as well as the seismic velocity in the medium. The data are processed with the help of a program system called "Astra" realized in the "Matlab" medium.

The efficiency of the algorithms and programs created was estimated by processing the data obtained on the basis of a scheme of direct and inverse VSP from ground-based and pulsed borehole sources. On the whole, the results of numerical and model experiments have shown the accuracy in determining the coordinates within the first meters, which indicates that the methods created to solve the problem are rather effective.

1. The problem of estimating parameters of the borehole bottom and velocity characteristics of the medium

Let the axes x and y in the Cartesian system of coordinates x, y, z be directed along the Earth's surface, and let the axis z be directed down to the

Earth's center. Let ν denote the average propagation velocity of the seismic wave in the vicinity of the borehole. The sensors that record (or radiate) seismic signals are located at the Earth's surface or in small boreholes, at points with the coordinates (x_i, y_i, z_i). Let t_i denote the propagation time of a seismic signal from the source at the borehole bottom (for instance, from the drilling bit) to the i-th point (or vice versa). It is necessary to determine the coordinates (x^*, y^*, z^*) of the borehole bottom and the velocity ν. One can also formulate a problem in which it is difficult to fix the radiation time of a seismic signal, and it has to be included in the unknowns to be determined. Then it will be necessary to determine the coordinates (x^*, y^*, z^*) of the borehole bottom, the time in the source t_s^*, and the velocity ν. Naturally, a minimum number of sensors will increase to five. When estimating the unknown parameters of the borehole bottom, we use a nonlinear system of the so-called conditional equations [1–3]:

$$\vec{t} = \vec{\eta}(X, \vec{\theta}) + \vec{\varepsilon}, \tag{1}$$

where $\vec{t} = (t_1, t_2, \ldots, t_N)^T$ is the vector of travel times of seismic signals, $\vec{\eta}(X, \vec{\theta})$ is N-dimensional vector of measured travel times (a theoretical travel time curve) or a regression function, $\vec{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N)^T$ is the residual vector, $\vec{\theta} = (x, y, z, \nu, t_s)^T$ is m-dimensional vector of the parameters estimated, $X = (\vec{x_1}, \vec{x_2}, \ldots, \vec{x_N})$ is the matrix of the coordinates of sensors (or radiation points), and N is the number of sensors (or radiation points).

Information about the distribution of errors $\varepsilon_i = t_i(\vec{x}_i, \vec{\theta}) - \eta(\vec{x}_i, \vec{\theta})$ is used to estimate the parameters. From here on, we assume that ε_i denotes mutually independent random variables distributed with the zero average and given variances: $E\varepsilon_i = 0$, $E\varepsilon_i\varepsilon_j = \sigma_i^2\delta_{ij}$, $\sigma_i = \sigma(\vec{x_i})$, δ_{ij} is the Kronecker delta, $i = 1, 2, \ldots N$. In case of difficulties with the specification of variances, they are assumed to be equal, and an unbiased estimate of the observation variance with a unit weight in the problem solution is obtained. The latter approach is used in this paper.

2. Methods to solve the problem

The problem of estimating the parameters $\vec{\theta}$ is part of the so-called regression analysis, and estimates of the least-squares method are its solution:

$$\vec{\theta} = \arg\min_{\theta\in\Omega} Q(\vec{\theta}), \qquad Q(\vec{\theta}) = \sum_{i=1}^{N} \sigma_i^{-2}(t_i - \eta(\vec{x_i}, \vec{\theta}))^2. \tag{2}$$

To find a minimum of the functional $Q(\vec{\theta})$, the Gauss–Newton iterative method or its modifications based on a linear approximation of the regression function in the vicinity of the point $\vec{\theta}^k$ is used:

$$J(X,\vec{\theta}^k)\Delta\vec{\theta}^k + \vec{\eta}(X,\vec{\theta}^k) - \vec{t} + \vec{\varepsilon} = 0 \tag{3}$$

where

$$J(X,\vec{\theta}) = \left(\frac{\partial\eta(\vec{x_i},\vec{\theta})}{\partial\theta_1}, \frac{\partial\eta(\vec{x_i},\vec{\theta})}{\partial\theta_2}, \dots, \frac{\partial\eta(\vec{x_i},\vec{\theta})}{\partial\theta_m}\right)_{i=1,2,\dots,n}. \tag{4}$$

The estimates $\vec{\theta}$ are found as a result of the iterative process ($\vec{\theta} = \lim\limits_{k\to\infty}\vec{\theta}^k$):

$$\begin{aligned} &\vec{\theta}^{k+1} = \vec{\theta}^k + \Delta\vec{\theta}^k, \\ &[J^T(X,\vec{\theta}^k)J(X,\vec{\theta}^k) + \alpha^2 I]\Delta\vec{\theta}^k = J^T(X,\vec{\theta}^k)y(X,\vec{\theta}^k), \quad k = 0,1,\dots \end{aligned} \tag{5}$$

Here $y(X,\vec{\theta}) = \vec{t} - \eta(X,\vec{\theta})$, α is the regularization parameter, and I is the unit matrix.

Another approach to solving problem (1)–(4), also used by the authors, is to solve system (3) directly at each step of the iterative process. At the present time, the method of pseudoinversion (or a generalized inversion) based on a singular decomposition (SVD-decomposition) is most widely used to solve this system. Modern versions of the MATLAB system have a built-in function `svd(A)` that implements this decomposition for an arbitrary matrix A of order $n \times m$.

Algorithm and program for automatic determination of wave arrival times. To determine the vector of wave arrival times t in the automatic measurement mode, one uses an algorithm of determining the arrival times of a quasi-periodic sequence of pulses at the background of Gaussian noise and estimating their shape [5]. The following expression is taken as an object function:

$$S_1(t_1,\dots,t_M) = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=0}^{q-1} y_{t_i+k}y_{t_j+k} \to \max_{\Omega} \tag{6}$$

where t_i, t_j $(i,j = 1,\dots,M)$ are arrival times of the first waves; y_{t_i+k}, y_{t_j+k} $(i,j = 1,\dots,M, k = 0,\dots,q-1)$ are the waves of a given duration q;

$$\begin{aligned} \Omega = \{(t_1,\dots,t_M) \mid 0 \le t_1 \le T_{\max} - q - 1,\ N - T_{\max} - q \le t_M \le N - q - 1, \\ q \le T_{\min} \le t_i - t_{i-1} \le T_{\max},\ i = 2,\dots,M\}; \end{aligned}$$

$T_{\min}$, $T_{\max}$ specify a minimum and a maximum values of the quasi-period, and M is the number of seismograms.

Criterion (6) is based on the maximum likelihood method. As a result of some transformations presented in [6], relation (6) is equivalent to the following expression:

$$\tilde{S}(t_1, \ldots, t_M) = \sum_{i=1}^{M} G(t_i) = \sum_{i=1}^{M} \sum_{k=0}^{q-1} \tilde{u}_k(\tilde{u}_k - 2y_{t_i+k}) \to \min_{\Omega}, \tag{7}$$

where

$$\tilde{u}_k = y_{t_1^*+k}, \quad k = 0, \ldots, q-1,$$
$$t_1^* = \arg \max_{0 \le t_1 \le T_{\max}-q-1} S_1(t_1) = \arg \max_{0 \le t_1 \le T_{\max}-q-1} \sum_{k=0}^{q-1} y_{t_1+k}^2.$$

An algorithm based on the method of dynamic programming described in (6) is proposed to solve the minimization problem (7).

The following recurrence formulas of dynamic programming are valid for the minimization problem (7) on the set Ω:

$$\begin{aligned}
S(n) &= 0, \quad \text{if } n \in [-T_{\max}, T_{\max} - T_{\min} - q - 1];\\
S(n) &= \min_{n-T_{\max} \le m \le n-T_{\min}} (S(m) + G(m)), \quad \text{if } n \in [0, N - q + T_{\min} - 1];\\
S(N) &= \min_{N-q \le n \le N-q-1+T_{\min}} (S(n) + G(n));\\
\mathrm{Ind}(n) &= 0, \quad \text{if } n \in [-T_{\max}, T_{\max} - T_{\min} - q - 1];\\
\mathrm{Ind}(n) &= \arg \min_{n-T_{\max} \le m \le n-T_{\min}} (S(m) + G(m)), \quad \text{if } n \in [0, N-q+T_{\min}-1],
\end{aligned}$$

where $S(n)$ and $\mathrm{Ind}(n)$ denote a minimum value of the functional and a minimum indicator at the n-th step.

The number of waves and their location in the sequence is determined by the recurrent calculation in the reverse order by using a minimum indicator:

$$\begin{cases} m_0 = \arg \min_{N-q \le n \le N-q+T_{\min}-1} (S(n) + G(n)), \\ m_i = \mathrm{Ind}(m_{i-1}), \quad i = 1, 2, \ldots \end{cases} \tag{8}$$

and the process terminates at such a step $i = r$, that $\mathrm{Ind}(m_r) = 0$.

As a result of calculation by using formula (8), we obtain a sequence $m_r, m_{r-1}, \ldots, m_1$ such that $(\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{\tilde{M}-1}, \tilde{t}_{\tilde{M}}) = (m_r, m_{r-1}, \ldots, m_1)$. The quantity r yields the estimate $\tilde{M}$ of the number of pulses caught in the frame. As a result of solving the minimization problem, we find an optimal set of the wave arrival times and their number:

$$(\tilde{t}_1, \ldots, \tilde{t}_{\tilde{M}}, \tilde{M}) = \arg \min_{\Omega} \tilde{S}_1(t_1, \ldots, t_M)$$

Taking into account the maximum likelihood estimates and the found parameters $\tilde{t}_i$ $(i = 1, \ldots \tilde{M})$ and $\tilde{M}$, one can easily find the sought for components of the U-wave:

$$\hat{u}_k = \frac{1}{\tilde{M}} \sum_{i=1}^{\tilde{M}} y_{\tilde{t}_i+k}, \quad k = 0, \ldots, q-1.$$

Planning the observation system. No matter how good the methods of solving systems (3) and (5) may be, in practice they are not very effective in the case of bad conditionality of matrix (4). Often, this is caused by poor organization of observations, namely, by inappropriate arrangement of seismic sensors with respect to the borehole bottom. One can conclude that observations should be planned, that is, one should select such an arrangement of sensors within the given territory that could maximally increase the conditionality of matrix (4) and, hence, improve the estimate of the borehole parameters. Specific problem statements of designing seismic observation systems are considered in [1, 2]. A special software has been developed by the authors to solve these problems.

3. Results of experimental material processing

The model data obtained from pulsed sources on the basis of the scheme of direct and inverse VSP were processed by using the method described, realized in the form of a software for PC. The first data were obtained from deep boreholes ("Uraineftegas") and granted by the authors [4]. The test method used the measured arrival times of direct waves in the direction "surface source-borehole receiver". The processing results obtained from 11 operational oil wells (with depths from 1000 to 2300 m) are presented in Table 1. The table shows the borehole number, its depth, a distance from the borehole bottom to its head, errors in determination of the borehole bottom by the coordinates x, y, z, and estimates of the velocities of the first seismic wave. The data in the table demonstrate the effectiveness of the program performance and a high quality of the experimental material. For instance, the root-mean-square errors in the determination of the coordinates of the borehole bottom for all boreholes are within 3–4 m, and for half of them—within 1–2 m. It should be noted that for 7 boreholes with depths within 1685–1948 m, a distance from the borehole bottom to its head is 160–331 m. For borehole No. 880 at a depth of 1846 m, the distance is only 61 m, for borehole No. 735 at a depth of 1725 m it is 471 m, for borehole No. 5252 at a depth of 1042 m it is 515 m (about half the depth) and for No. 436 at a depth of 1836 m it is 700 m.

Results of measurements using the inverse VSP scheme. The model experiments were made with the help of the method of inverse vertical seismic profiling (IVSP) using a water-filled borehole 135 m deep. A scheme of the experiments is presented in Figure 1. Powder explosions of 12.5 g and

Table 1

Borehole No.	Distance from borehole bottom, m	Error in the determination of borehole bottom, m			Velocity of seismic waves, km/s
		x	y	z	
436	700.96	1.12	1.39	0.54	2.53
666	160.87	1.35	1.76	0.60	2.40
753	470.71	3.09	3.07	1.20	2.40
739	257.00	2.83	3.58	1.14	2.43
742	315.00	2.16	2.84	1.12	2.53
751	323.07	0.96	1.26	0.28	2.49
880	60.65	0.97	0.94	0.34	2.51
981	282.96	2.82	3.09	1.51	2.60
1337	331.03	3.35	2.64	1.47	2.51
1338	210.29	1.29	0.97	0.34	2.51
5252	515.61	1.86	3.84	2.84	2.12

30 g, respectively, were used as a source of seismic oscillations. Blasting control was remote, with the electric current passing from a 220 V supply line through a wire in a glass with the explosive. The process of wire burnout initiated the powder blasting. The reference signal was recorded from sensor S1 located at the borehole head. The signal was initiated by a hydroacoustic wave, which propagated from the source along the liquid column filling the borehole cavity. The reference signal was transmitted via the lines to the recording seismic station. A 12-channel digital seismic station “Lakkolit-M” is used to record seismic signals. For each 12-channel arrangement of seismic sensors (Figure 2), explosions at depths of up to 120 m were recorded. As an example, Figure 3 presents the spectral-time function of the first wave recorded from an explosion at a depth of 100 m. It follows from the figure

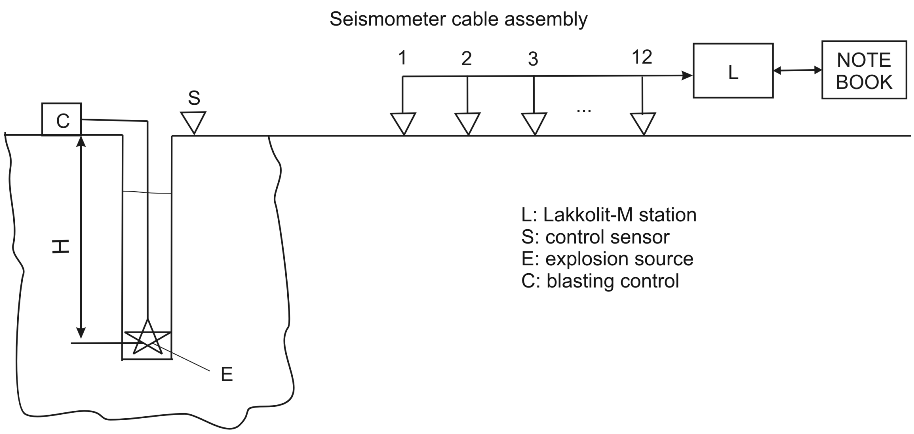

Figure 1. Scheme of model experiments

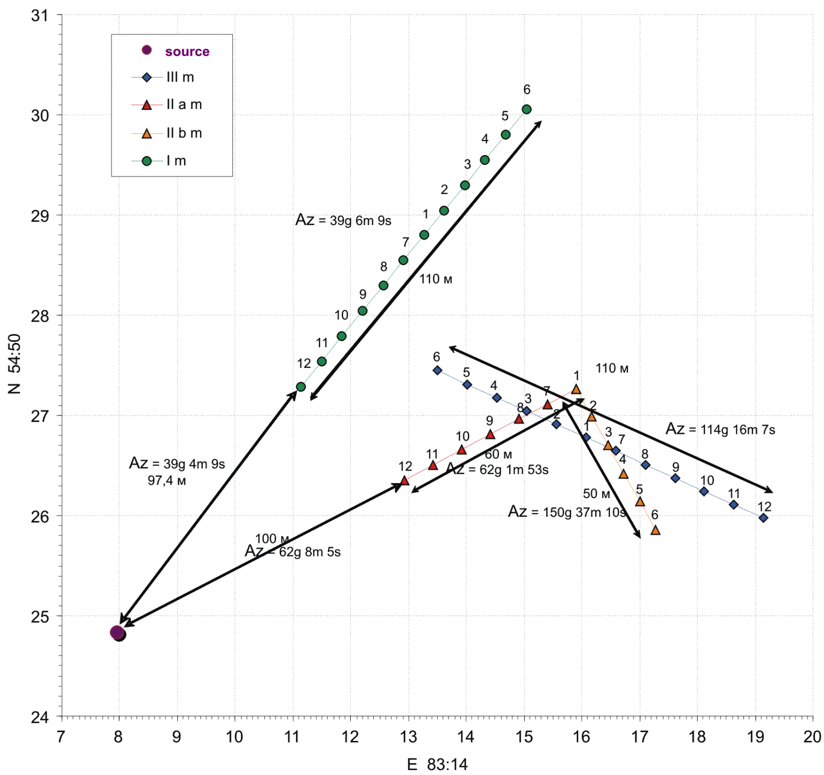

Figure 2. Arrangements of sensors at profiles

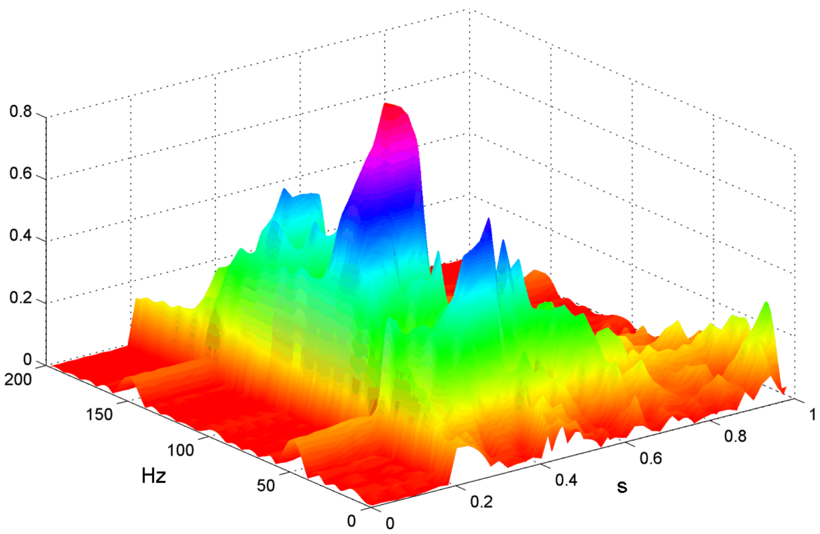

Figure 3. The spectral-time function of an explosion record at a depth of 100 m

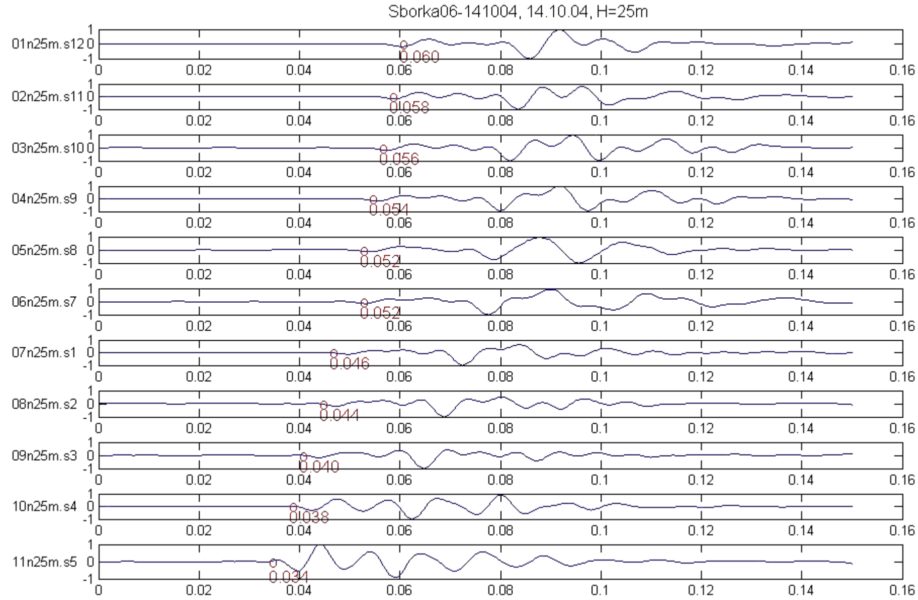

Figure 4

that the bulk of energy of the explosion is concentrated in the frequency range from 80 to 160 Hz. The arrival times of direct waves were automatically determined with the help of algorithm (8). The results of determining the first waves arrival times are illustrated graphically in Figure 4. The arrival times are denoted by points in each of 11 seismograms.

The measured values were used to solve the inverse problem (1) in order to determine errors in the calculation of the coordinates of the borehole bottom and wave velocities for various source depths. The results of the calculations are presented in Table 2. The table presents source depths, errors in determination of the borehole bottom by the coordinates x, y, z, velocity values of direct waves, and errors in their determination, respectively. The

Table 2

No. of source location	Error in the determination of borehole bottom, m			Velocity of seismic waves, km/s	Error in the determination of velocity, km/s
	x	y	z		
br1	0.550	0.757	0.080	1.599	0.0058
br5	0.678	0.927	3.116	1.586	0.0061
br10	0.789	1.070	3.783	1.745	0.0077
br25	0.877	1.208	1.985	2.044	0.0099
br100	1.070	1.478	0.887	3.219	0.0186
br120	1.577	2.163	1.014	3.343	0.0276

table data illustrate rather a high accuracy in determination of the source coordinates (the error along the coordinate z at maximum depths does not exceed 1 %, the horizontal deviation not exceeding 2 m).

Conclusion

A method to determine the coordinates of a borehole source of the pulsed type, the time in the source, and the velocity reconstruction in a medium has been developed, and its efficiency has been evaluated. The results of processing of the data of the model experiments obtained with the use of boreholes show a high potential effectiveness of the method for solving real problems. Final conclusions will be obtained by using the continuous boring data.

References

[1] Omelchenko O.K., Gusiakov V.K. Designing a network of seismic stations for tsunami warning service // Volcanology and Seismology. — 1996. — No. 2. — P. 68–85.

[2] Yermakov S.M. et al. Mathematical Theory of Experimental Design. — Moscow: Nauka, 1983.

[3] Omelchenko O.K. Numerical realization of a wave method to determine the borehole bottom coordinates // Bulletin of the Computing Center, Ser. Mathematical Modeling in Geophysics. — Novosibirsk: NCC Publ., 1996. — Iss. 4. — P. 207–214.

[4] Krivoputsky V.S., Novakovsky Yu.L. A method of borehole bottom location in the process of boring // Bulletin of the Computing Center, Ser. Mathematical Modeling in Geophysics. — Novosibirsk: NCC Publ., 1994. — Iss. 3. — P. 38–43.

[5] Woskoboynikova G.M. Determination of the arrival times of the seismic by the dynamic programming method // 9-th Korean-Russian Intern. Symposium on Science & Technology, KORUS 2005, June 26 – July 2, 2005. — Novosibirsk: Novosibirsk State Technical University, 2005.

[6] Khairetdinov. M.C., Omelchenko O.K., Rodionov U.I. Automatic technology of seismic source location // Intern. conf. "Mathematical Methods in Geophysics". — Novosibirsk, 2003. — Vol. II. — P. 529–535.