MAGYAR ÁLLAMI
EÖTVÖS LORÁND
GEOFIZIKAI INTÉZET

GEOFIZIKAI
KÖZLEMÉNYEK

ВЕНГЕРСКИЙ
ГЕОФИЗИЧЕСКИЙ
ИНСТИТУТ
ИМ Л. ЭТВЕША

ГЕОФИЗИЧЕСКИЙ
БЮЛЛЕТЕНЬ

GEOPHYSICAL

TRANSACTIONS

EÖTVÖS LORÁND GEOPHYSICAL INSTITUTE OF HUNGARY

CONTENTS

VOL. 40. NO. 3–4. APRIL 1997. (ISSN 0016–7177)

BUDAPEST

TARTALOMJEGYZÉK

Lapunk az Ipar Műszaki Fejlesztéséért Alapítvány támogatásával készült.

GEOPHYSICAL TRANSACTIONS 1997
Vol. 40. No. 3-4. pp. 119-143

Hungary's new gravity base network (MGH-2000)

Géza CSAPÓ*

The modernization of the Hungarian Gravity Base Network is being carried out in the framework of the establishment of the unified gravity network of the Central-European countries. The objective of former networks and the necessity for development are reviewed. The scale of the new network is guaranteed in the SI system by the numerous absolute gravity measurements carried out in recent years. The observation methods, data processing and adjustment procedures are presented. The results of the comparison of Czech, Slovakian and Austrian, Unified European gravity networks with the Hungarian network are discussed.

Keywords: gravity, network, absolute measurements, relative measurements, data processing, adjustment of gravity network

1. Introduction

One of the major consequences of the political changes in Hungary was the abolition of the secrecy of gravimetric information during the early 1990s. This made it possible for researchers to take part in the various international projects.

The International Union of Geodesy and Geophysics (IUGG) has long been planning to set up a unified scale and datum gravimetric network which could be applicable throughout the whole continent of Europe. Its conditions have now been established because several countries possess portable absolute

* Eötvös Loránd Geophysical Institute of Hungary (ELGI), H-1145, Budapest, Kolumbusz u. 17-23, Hungary

Manuscript received: 30 April, 1996

gravimeters (AXIS, JILAG, etc.), providing a unified scale in accordance with the current accuracy specifications. At the same time the need for increasing the accuracy of global geodetic reference systems, and the solving of several geodynamic and geotectonic problems have brought about the realization of this objective as a daily routine.

As regards the number of absolute stations and point density (the number of 1st and 2nd order bases as well as their regional distribution), and also their accuracy, the gravimetric base networks of the individual European countries are somewhat heterogeneous. It seems both necessary and expedient to establish a unified network on the basis of the principles recommended at the joint conference of the Geodesy and Geophysics Working Group (GGWG) of NATO and the Mapping Services of East European Armies held in Budapest in collaboration with civilian experts in 1994. Its essence defines the creation of a network consisting of absolute points (zero order points) at a general distance of 100–150 km within which the 1st and 2nd order bases will need to be measured with modern relative gravimeters. It is advisable that the gravity value of the absolute points be checked in three year periods. At the same time it is deemed desirable that zero order points constitute an integrated network with the Hungarian GPS Base Network as well as increasing, where necessary, the number of gravimetric points within a suitable radius of integrated points.

The US Defense Mapping Agency (DMA) began to increase the accuracy of the WGS-84 reference ellipsoid in 1991 and substantially helped Hungary to establish both a Hungarian national Military GPS Network (KGPSH) and an absolute gravimetric base network [Ádám et al. 1994].

2. Antecedents

Hungary's first gravity network (MGH–50) covering the entire territory of the country was established by the Eötvös Loránd Geophysical Institute (ELGI) during the early 1950s. The measurements were carried out by a Heiland gravimeter [Renner, Szilárd 1959]. This network provided a sound basis for implementing a so called 'national reconnaissance gravity survey' launched at the time, but as a result of industrial and infrastructural developments during the 1960s, most of the base points established mainly along national roads were beginning to deteriorate or simply became unsuitable for their original purposes.

During the 1960s ELGI acquired geodetic type Sharpe gravimeters considered to be quite modern at the time, making it possible to re-measure and extend the 1st order network. In 1971 observations were carried out with three CG–2 Sharpe instruments transported by air to nineteen network points established at different airports. Network points were established in a formation so that their connecting sides could constitute triangled- and quadrangled-polygons and they were continually remeasured [CSAPÓ, SÁRHIDAI 1990a].

When the 'Hungarian First Order Levelling Network' was established during the 1970s, gravimetric observations had to be carried out for normal correction. At the same time this made it possible to establish a 2nd order gravimetric network by placing the basis points along the levelling lines in the gardens of long standing buildings such as churches or manor-houses. The base points are monumented by concrete blocks of 60 by 60 by 100 cm set up at ground level and a brass bolt is fixed in the middle of their top surface. The height of the points was determined in collaboration with the Institute of Geodesy, Cartography and Remote Sensing (FÖMI). The gravimetric measuring of the network was completed with two Sharpe CG–2 instruments and a geodetic type LaCoste–Romberg (LCR) between 1980 and 1989. The average distance between the points ranges between 10 and 20 km. The gravimeters and observers were transported by car. The network consists of consecutive triangles. The gravity difference (Δg) of each individual side was determined in the observation sequence A-B-A-B-A. The common adjustment of the measurement was carried out in 1971 and between 1980 and 1989, as well as compiling a catalogue of points was implemented in 1991. The standard deviation of the adjusted network of MGH-80 is ± 16 μGal** [CSAPÓ, SÁRHIDAI 1990b].

The objective of creating a unified gravimetric network which would cover a large territory was laid down in the cooperation projects of the geodetic surveys of the East Central European countries as early as the mid-1960s. The concept of this network was rather similar to the existing ones. Establishment of the first five absolute stations measured by Soviet made GABL absolute gravimeters [CSAPÓ 1981] brought about the partial realization of this plan in Hungary just as a Unified Gravity Network of the

** $1\mu Gal = 10^{-8} ms^{-2}$

Czech Republic, Hungary, and Slovakia (UGN) has been established by now. The observations needed to establish this common network had continuously been carried out by ELGI's experts in bilateral and multilateral forms of cooperation since 1972 [Csapó et al. 1994].

In 1992 and 1993 gravimetric measurements were carried out in the form of interconnecting measurements between Austria and Hungary within the framework of bilateral cooperation. This activity consisted of relative and absolute measurements. Relative measurements were carried out with 4–5 LCR gravimeter, whereas absolute measurements were carried out with a JILAG–6 absolute gravimeter [Csapó et al. 1993] by D. Ruess of 'Bundesamt für Eich- und Vermessungswesen' (BEV).

Implementation of the Unified European Gravity Network (UEGN) was commenced in 1993 in the East Central European countries under the auspices of the joint plan of the International Gravity Commission (IGC) and the Geodesy and Geophysics Working Group of NATO (GGWG) in the form of international cooperation. The reason for organizing an international project was that none of these countries possessed an absolute gravimeter. Absolute measurements were carried out in Hungary between 1993 and 1995 with the AXIS FG5 No.107 gravimeter of DMA and with JILAG–6 equipment of BEV [Friederich 1993, Spita 1994, Krauterbluth 1995]. The job was partly financed by the US-Hungarian Science and Technology Joint Fund (MAKA) under contract JF. No. 369 and it was implemented by the DMA expedition; part of the work was carried out by the Geological Survey of Austria on the basis of a scientific agreement between ELGI and the BEV.

In 1994 the linking of absolute stations with relative gravimeters was commenced with four LCR–G instruments transported by car. Hungary's absolute stations are situated at an average distance of 100 to 120 km, so 1st order and 2nd order points are used as tie points. The project is being financed by the National Commitee for Technological Development (OMFB) by means of contract No. MEC-94-0508. This job is backed by DMA too by their long-term loan to ELGI of two geodetic type LCR gravimeters.

3. Concept of the new base network (MGH-2000)

Previously the most important aspect of gravity base networks was to provide a unified framework for gravity surveys carried out by various institutes and companies for mineral exploration. It was necessary to have an

appropriate number of base points for the economical completion of the work. Strangely enough the requirements for increasing the accuracy of measurements did not increase parallely with the development of the instruments. On the contrary, the requirements diminished as perfomance grew. For this reason, the improvement of the base network became of secondary importance as far as geophysics specializing in mineral exploration was concerned. However, during the course of the preparatory phase of 'International Gravimetric Calibration Polygon' (IGCP) which was launched to compile a unified geophysics map within the framework of cooperation between the Geodetic Surveys of the former Socialist Countries, including certain scientific projects launched by their respective academies of sciences, it turned out that the scale relations of the gravimetric base networks in this countries sometimes had a difference as big as 100–150 μGal. Such experiences led to the planning of the above mentioned unified network. However, its implementation was abandoned mainly for financial reasons in the second half of the 1980s.

As a matter of fact the number of geophysical measurements carried out for mineral exploration has decreased world-wide during the 1990s. At the same time the extension of cooperations requires unified measurements, geophysical maps, and databases in order to be able to resolve scientific as well as economic tasks covering entire continents. The 'advances' of military and civil spheres make it unnecessary to carry out parallel research in the same field. One of the prime examples is the useful cooperation between military and civil institutes in the GPS base network activities in Hungary.

When establishing Hungary's new gravimetric base network, we considered the following aspects important:

— The national gravimetric network should be regarded as a geodetic one, just as triangulation, levelling, or any GPS base network. The protection, maintenance and continuous development of networks should be undertaken by the government.

— Throughout the planning of the zero order base net (absolute stations) we attributed great importance to placing them evenly all over the country as well as setting them quite close to GPS geodynamic points — which had been established some time before — so that an economical system of integrated networks could be set up. We call these points integrated network points which means

that the same point is a member of the base network of GPS, gravity, levelling, etc.

— When new points are established, or those that have deteriorated are replaced, one has to bear in mind the changes of ownership which are closely related to the protection of points.

— MGH–2000 is part of the above mentioned UGN, so the methods of modern network planning could partly be applied [CSAPÓ, SÁRHIDAI 1985, SÁRHIDAI 1986], because the joint form is basically determined by the previous and applicable parts of the networks of the three countries (Czech Republic, Hungary and Slovakia). In any case, we have experienced that the planning of networks for optimal network measurements can give rise to the necessity of establishing connections between far distant points. However, this might not be carried out (transporting instruments by planes) on account of Hungary's present financial situation.

4. Structure of MGH-2000

The network consists of zero order points as well as of points first order and second order (see the top of *Fig. 1*). In addition, the network contains 16 original, so called 'fundamental'. points established in 1954, so that the transformation equation could be given in relation to the old network and the new one while establishing the new network (one of them completely broke down during the past fifty years).

The zero order network

The use of such network is meant to establish the scale of the national (entire) base network as well as enabling the stability of gravity to be checked by repeated observations. The zero order network consist of 12 absolute stations (7750 km^2/point) whose locations are given in *Fig. 2* — including 5 foreign absolute stations near the borders. These points were placed at the ground level of important buildings whose survival and accessibility seem to be ensured well into the future (castles, stately homes, etc.). Monumentation was implemented by concrete blocks with the dimensions 120 by 120 by 100 cm. A brass bolt was fixed to the middle of the top of the block indicating

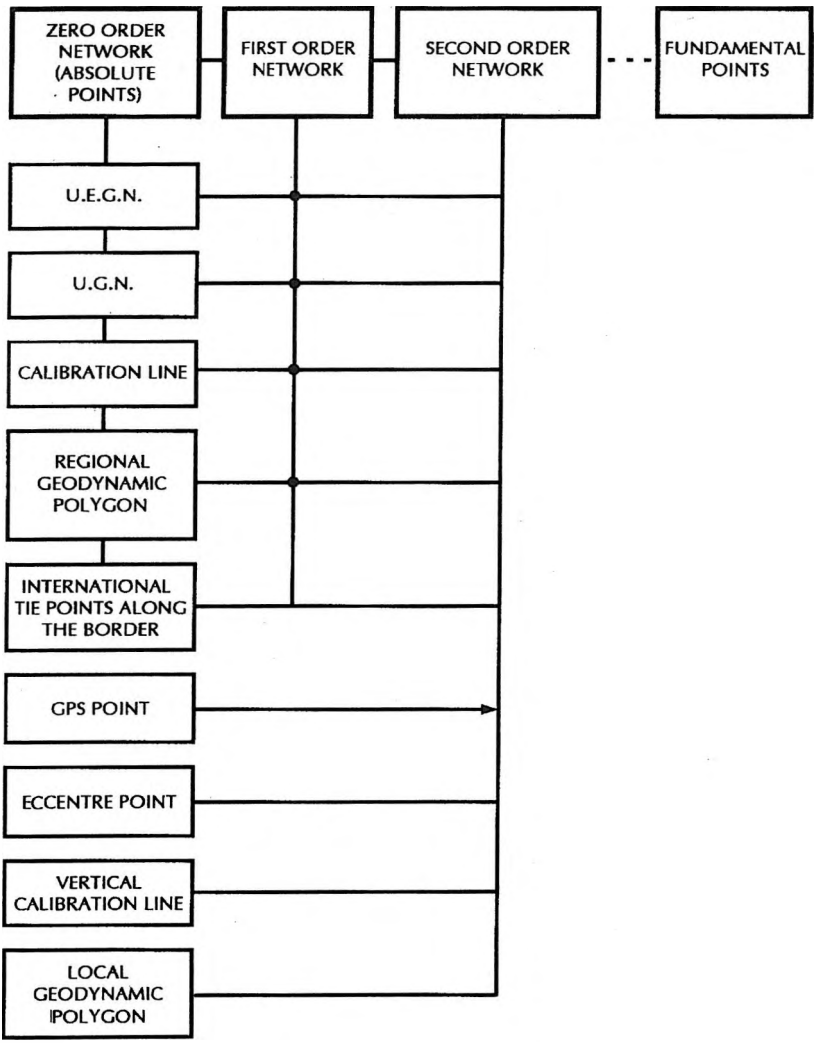


Fig. 1. The structure of MGH-2000
1. ábra. Az MGH-2000 szerkezeti felépítésének vázlata

the height above sea-level based on the Baltic system. The points were tied to two or three points of the national levelling network, allowing ±5 mm confidence limit. The geographical coordinates of the stations were established on the basis of 1:10000 topographical maps with ±1 second confidence limit. Gravity acceleration relating to the reference heights of absolute gravimeters was determined with LCR gravimeters allowing 1.5–3 μGal

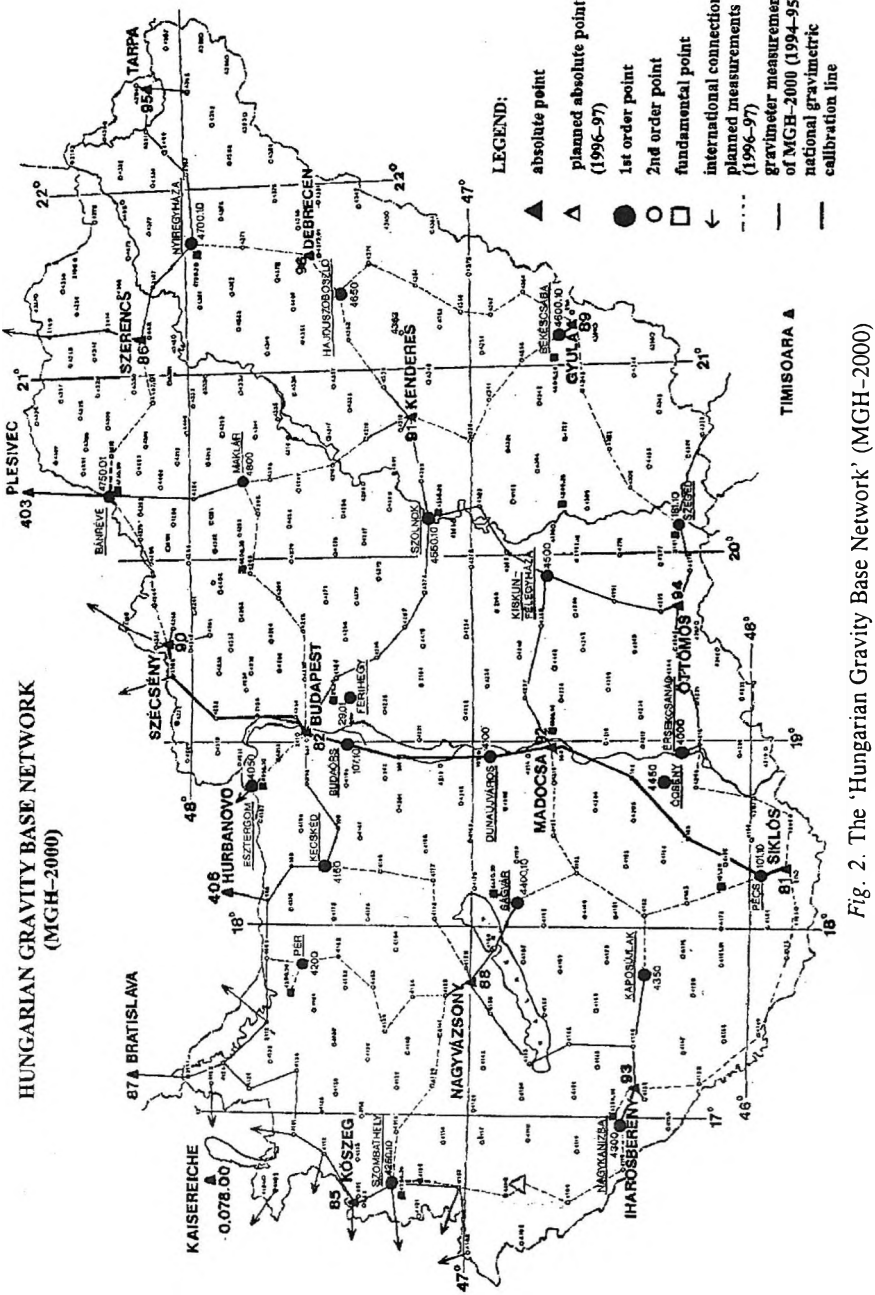


Fig. 2. The 'Hungarian Gravity Base Network' (MGH-2000)

2. ábra. A „Magyarországi Gravimetriai Hálózat" (MGH-2000) pontvázlata

confidence limit. The station established in Budapest is of extraordinary importance because measurements need to be repeatedly carried out on it with different types of absolute gravimeters every two or three years — as has in fact taken place since 1980. At most stations repetitive measurements have been carried out in the past three or four years as well. Each station has at least one so called 'eccenter point' which is fixed with a concrete block of 80 by 80 by 100 cm outside the building. The relative confidence limit of its 'g' value is better than 5 μGal. The most important parameters of the zero order network can be seen in *Table I.*

NUMBER	STATION	LATITUDE [° ' "]	LONGITUDE [° ' "]	HEIGHT [m]	DATE OF THE FIRST/ LAST OBSERVATION	GRAVIMETER
81	SIKLÓS	45-51-10	18-07-55	128.489	1978 / 1995.	GABL / JILAG-6
82	BUDAPEST	47-32-00	19-01-00	201.563	1980 / 1995.	GABL / AXIS FG5
85	KŐSZEG	47-23-24	16-32-33	284.461	1980 / 1993.	GABL / JILAG-6
86	SZERENCS	48-09-56	21-12-21	111.243	1980 / 1993.	GABL / JILAG-6
88	NAGYVÁZSONY	46-59-23	17-42-00	241.085	1993.	AXIS FG5
89	GYULA	46-38-42	21-17-14	89.053	1987 / 1995.	GABL / JILAG-6
90	SZÉCSÉNY	48-05-07	19-31-08	166.888	1993.	AXIS FG5
91	KENDERES	47-14-54	20-40-37	83.450	1993.	AXIS FG5
92	MADOCSA	46-41-19	18-57-40	93.758	1994.	AXIS FG5
93	IHAROSBERÉNY	46-21-48	17-06-17	203.898	1994.	AXIS FG5
94	ÖTTÖMÖS	46-17-04	19-40-47	124.042	1994.	AXIS FG5
95	TARPA	48-06-14	22-31-40	110.778	1995.	AXIS FG5

Table I. Most important parameters of the zero order network

I. táblázat. A nulladrendú hálózat legfontosabb paraméterei

First order network

The 21 points included in Fig. 2 are, by and large, the same as the bases of MGH–80 placed at airports [CSAPÓ, SÁRHIDAI 1990b]. The distance varies between 50 and 70 km and the density of points is 4400 km^2/point. Bigger point monuments (100 by 100 by 100 cm) were placed exactly beside the MGH–80 points before UGN measurements started in Hungary in 1982, so that several gravimeters could be set up at the same time by the staff of air expeditions. Determination of the geographical coordinates of the points was similar to the methods described in the section on zero order points. Levelling (altitude determination) between the individual points and the national levelling base points was performed by the Budapest Geodetical and Cartographical Enterprise and by ELGI, with 1–10 mm confidence limit.

Second order network

As has already been mentioned, these points were established by ELGI in the 1970s. The distance between the individual points is 10 to 15 km in hilly areas whereas it ranges from 15 to 25 km in the plains. The average density of points is 220 km^2/point. We have replaced a couple of dozen points which were destroyed during the last twenty years and have integrated them into MGH-80. The new network contains 430 2nd order points.

5. Division of network points according to function

The use of network points is shown in the vertical column of Fig. 1. Not only do zero order points make it possible to deduce the scale of the national network, but they also have a similar role in the establishment of continental and regional networks (UEGN, UGN) too [BOEDECKER et al. 1994].

Testing of the scale is done by repeated absolute measurements. The calibration factor of the relative gravimeters has to be determined and to be checked from time to time on calibration lines. It means that before and after the field work (normally in spring and in autumn), high precision gravity measurements are carried out at the points of a calibration line regarded as standard.

The fundamental points of regional gravimetric lines (i.e. Carpathian polygon) established to observe the non-tidal changes of gravity field and tie points between the gravimetric networks of neighbouring countries are as far as possible absolute stations. As shown in Fig. 2, absolute points along the borders (border region points) served this purpose during the measurements of MGH-2000 and UGN.

1st order points are meant to create ties between absolute stations placed at greater distances. Their high accuracy enables them to be used as absolute points during gravimetric measurements as well.

2nd order bases are used more extensively than high grade network points because they are directly used for detailed gravity surveys. These bases are used as eccentric high order points, vertical bases, and as local points of deformation study lines (Debrecen) in Hungary.

The geodynamic points of the National GPS Network have a special status within Hungary's gravimetric base network because gravimetry aspects did not enjoy any priority in the selection of their location. The applicability

of GPS technique and some geological considerations have to be reckoned with. For this reason, it has been made possible by applying an appropriate survey mark to carry out relative gravimetric measurements on them too and they were integrated into the MGH-2000 network. With regard to the fact that these points were generally placed on sparsely surveyed highland regions, they can serve as local bases for detailed gravity surveys.

6. National gravimetric calibration line

An outline of the National Gravimetric Calibration Line is given in *Fig. 3*; its location within the country can be seen in Fig. 2. Its present form has been undergoing development by ELGI since 1969. There are four absolute stations within the 210 mGal range of the line (the highest Δg value is 250 mGal between the base points of the country). The other points of the calibration line are 1st and 2nd order points at an average distance of 30 km from each other. Δg values between the points were previously determined with Askania Gs-12, GAG-2, Sharpe, Worden, then LCR gravimeter groups. The vertical gradients of the points were determined by a group of LCR instruments with an accuracy of 4–7 μGal, and with standard A-B-A-B-A-B observation sequence [CSAPÓ 1987]. The relative accuracy of each point is 7–12 μGal. The calibration line is part of UGN, and the section of

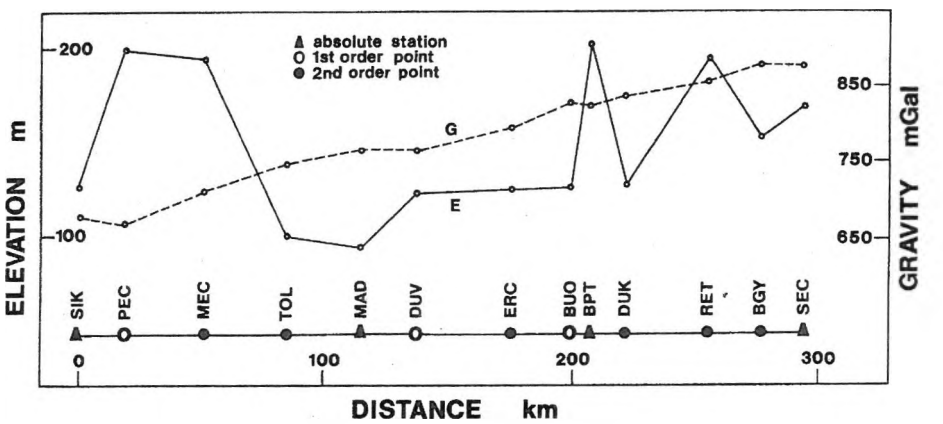


Fig. 3. Sketch of national gravimetric calibration line

3. ábra. A nemzeti gravimetriai hitelesítő alapvonal vázlata

Siklós–Budapest is the southern part of the 'Carpathian Polygon'. The Carpathian Polygon was established by ELGI in collaboration with Czech, Polish, and Slovak partners in 1973 to monitor the non–tidal variations of the gravity field in the Carpathian Basin (the line starts from Siklós absolute point then goes via Budapest, Zilina, Zakopane to Cracow). It was reobserved in 1978–79 and 1988–89, but the latest (planned) observation campaign was not launched on account of ELGI's financial difficulties.

Calibration of gravimeters

The present form of the calibration line shown in Fig. 3 was finalized by ELGI in 1985. The former calibration line was based on the 'Potsdam Gravimetric System'. The first absolute station was established on the former line at Budapest in 1980. It is also part of the new calibration line. ELGI carries out measurements twice a year on the line. The scale of gravimeters can also be checked by adjustment of the measurements carried out between two absolute points. Such a version was also prepared for adjusting MGH–2000.

ELGI has built laboratory calibration equipment [CSAPÓ, SZATMÁRI 1995] for the calibration of LCR gravimeters with feedback electronics. With the help of this equipment the possible time variations of the calibration factor of LCR gravimeters can be studied as well [MEURERS 1994].

7. Data preparation

For the formerly mentioned reasons the observed data of MGH–2000 contain data obtained from outside the country (absolute measurements, polygons, tie measurements along the borders, etc.) as well. Three databases have been compiled for the processing of field records.

Catalogue of points

This contains the numbering of points, names, geographical coordinates according to a Gauss-Krüger projection system and Baltic heights, preliminary gravity values and their reliability obtained from adjustments in the eighty character records. The number of the 1:10000 scale topographic maps

of the Unified National Mapping System and a four digits information code are given as well. This data file contains 600 records.

Catalogue of instruments

Information is given on the relative gravimeters applied (9 Sharpe gravimeters, 3 Worden, and 10 LCR gravimeters). The catalogue contains the make, type, registration number, original calibration tables and coefficients, as well as the corrections needed on account of the periodic errors of the measuring system, including other data (barometric correction factors, the operating temperature of thermostat, etc.) which were taken into account in the processing of observations.

Field records

The field records contain the observed data in analogue form. These records are stored in computer files consisting of eighty character records for each instrument. The field records contain both numeric and alphabetical data, such as name of the tie, catalogue number of the points, date of observation, time zone, dial and voltmeter readings, local time of readings, daily conversion factor of galvanometer reading to μGal (mVolt/μGal), instrument heights above the bench mark or ground, meteorological data and method of gravimeter reading (optical, CPI, FB), the name of the observer, and data on the transporting vehicle.

Data files are arranged in a library so that the field records observed in different years can easily be identified, and complemented (i.e. BASELINE\SIKLÓS\1994\LCR-963).

8. Data processing

Absolute measurements

Table I shows that during the last one and half decades absolute measurements were carried out in Hungary with three different types of instruments. It was noted that some differences appeared during the course of the data processing as well as during the application of corrections. In view of this, it was necessary to reprocess all observations on the basis of a unified approach. Instrument corrections as given by the observers were

accepted. It has recently turned out that the observed values are dependent on the light intensity of laser fringes. Should the (old ADM 686) fringe detector be replaced in an equipment by a new AD 9696, different gravity data can be obtained in a given light intensity range (D. RUESS, personal communication). This is what happened when the AXIS FG5 No. 107 instrument was used, DMA is due to reprocess all its data obtained after 1993. The possible corrections are not included in the present adjustment.

The calculation of tide correction was based 505 tidal waves. The parameters used in the calculations are given in *Table II*. The parameters were compiled on the basis of the data of Pecný (Czech Republic) tide registration station [HOLUB et al. 1986]. Ocean tide was disregarded.

WAVE	DELTA	KAPPA	AMPL. CORR.	PHASE CORR.
Q1	1.1525	-0.28	1	0
O1	1.1486	-0.03	1	0
M1	1.1391	0.45	1	0
K1	1.1362	0.10	1	0
J1	1.1542	-0.20	1	0
OO1	1.1565	0.08	1	0
2N2	1.1586	0.85	1	0
N2	1.1781	1.43	1	0
M2	1.1848	1.03	1	0
L2	1.1554	-0.16	1	0
S2	1.1861	0.30	1	0
M - S0	1.0000	0.00	1	0
SSA	1.1600	0.00	1	0
MM	1.1600	0.00	1	0
MF	1.1600	0.00	1	0
MTM	1.1600	0.00	1	0
M3	1.0775	-0.22	1	0

Table II. Parameters of tidal waves used for the processing of MGH-2000
II. táblázat. Az MGH-2000 mérési eredményeinek feldolgozásánál figyelembe vett árapály hullámok paraméterei

Correction due to polar motion. The observed gravity is corrected for changes in the centrifugal acceleration due to the variation of the distance of the earth's rotational axis from the gravity station. The following formula was used in the calculations:

$$\Delta g_{pm} = 1.164\ \omega^2 R \sin 2\varphi\ (x \cos\lambda - y \sin\lambda)\ \text{ms}^{-2}$$

where:

$\omega = 7.292.\ 10^{-5}$ rad/s = angular velocity of the Earth's rotation

$R = 6.371.\ 10^{6}$ m = mean radius of the Earth

φ, λ = geographical coordinates of the absolute station

The value of correction in μGal:

$\Delta g_{pm} = -19.1 \cdot \sin 2\varphi\ (x''\cos\lambda - y''\sin\lambda)$

the actual values for x and y can be found in the Annals of the International Earth Rotation Service (IERS).

Correction due to the changing atmospheric masses. Only the local part of the atmospheric effect was taken into consideration. The relation is given by the standard of DIN 5450 of 1968 and with the following empirical coefficient:

$$c = 0.30\ \mu\text{Gal/hPa}$$

The local part of the correction compensates for 80% of the total atmospheric effect.

Reduction of gravity value to the bench mark. In the adjustment the gravity values relevant to bench marks were used. Therefore, the results of the absolute measurements which are related to a certain height depending on the type of the equipment (850-1300 mm) were referred to the bench marks. The vertical gradients (i.e. the Δg between two points by 1 m vertically apart) were determined by a group of gravimeters. The vertical gradients of Hungary's absolute stations can be found in *Table III*. The vertical gradients measured between 1979 and 1994 at the Budapest absolute point are presented in *Fig. 4*. The vertical gradients as a function of height — determined at the absolute point of Budapest — are presented in *Fig. 5*.

Relative measurements

The computer stored field records were processed by a program developed in ELGI. Processing steps:

— conversion of the readings to mGal

— correction calculations (tidal, height, barometric and periodic error corrections)

— calculation of the corrected relative gravity values

— drift calculations

NUMBER	STATION	LAST VALUE (mGal)	N_{VG}	VERTICAL GRADIENT (eötvös)
81	SIKLÓS	980678.327	22	3407 ± 16.0
82	BUDAPEST	980824.294	50	2509 ± 9.7
85	KŐSZEG	980784.713	22	2661 ± 23.9
86	SZERENCS	980872.789	51	2969 ± 11.4
88	NAGYVÁZSONY	980765.818	18	2565 ± 12.3
89	GYULA	980766.404	29	2913 ± 10.5
90	SZÉCSÉNY	980873.111	15	3059 ± 17.8
91	KENDERES	980810.283	12	2662 ± 23.7
92	MADOCSA	980761.777	8	2560 ± 17.1
93	IHAROSBERÉNY	980699.028	19	2805 ± 9.7
94	ÖTTÖMÖS	980725.926	17	2634 ± 10.0
95	TARPA	980880.426	15	2712 ± 20.0

Table III. Vertical gradients of the Hungarian absolute stations (N_{VG} — number of determinations)

III. táblázat. A magyarországi abszolút állomások vertikális gradiens értékei (N_{VG} = mérések száma)

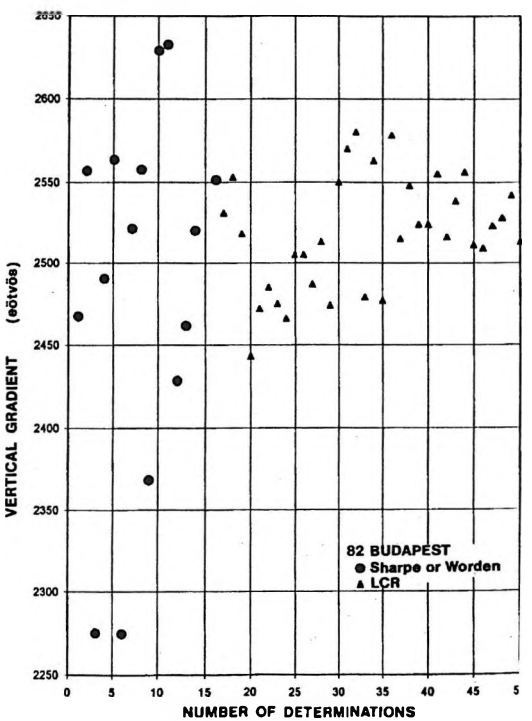


Fig. 4. Vertical gradients obtained between 1979 and 1994 at the Budapest absolute station

4. ábra. A budapesti abszolút állomáson 1979–94 között végzett vertikális gradiens mérések eredménye

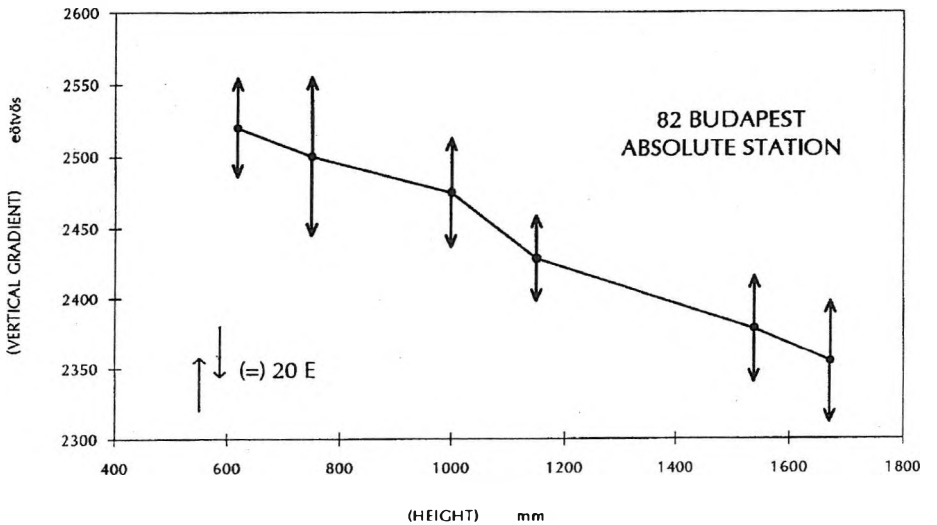


Fig. 5. Vertical gradient as a function of height — at the absolute station of Budapest
5. ábra. A vertikális gradiens értékének magasságfüggése — Budapest abszolút állomáson

— calculations of drift corrected relative gravity values
— Δg calculation
— error calculations

A total of 9100 measuring sections were processed for adjusting MGH-2000 (and UGN).

9. Adjustment of measurements

The observed data were adjusted by the least squares method as a constrained network. The fixed points of the network were the latest results of absolute determinations. The mean values of the observed gravity differences (Δg) observed in the A-B-A-B-A system by one gravimeter (which means the average of the four observed differences) were taken as one individual measurement. The adjustment procedure was reviewed in detail in a former paper [CSAPÓ, SÁRHIDAI 1990b]; here only the weight determination is discussed. To decrease the effect of relatively large errors in the adjustment their weight should be decreased, but before the adjustment the errors are not known. This contradiction can be solved by an iteration

('Danish method'): in the first step ($j=1$) all observed data have equal weight ($p=1$), in the further steps the weight will be:

$$p_{ij}=1/1+a_k \cdot v^2_{j-1}$$

where j represents the actual iteration step. The a_k coefficient is correct if $p\leq 0.25$ for the erroneous measurement [SOHA 1986]. The threshold of errors can be taken as the function of the errors of unit weight, then:

$$a_k=3/v^2_k$$

where

$$v_k=3\mu_0 \quad \text{if} \quad v_{max}>3\mu_0$$

$$v_k=2\mu_0 \quad \text{if} \quad 2\mu_0<v_{max}<3\mu_0$$

$$v_k=\mu_0 \quad \text{if} \quad \mu_0<v_{max}<2\mu_0$$

The erroneous measurements will get less weight with each subsequent iteration step. The iteration should be continued until the error of unit weight decreases to a considerable extent. In the adjustment of MGH-2000, two iteration steps proved to be sufficient. Adjustments were carried out on two different data sets and both of them in two versions.

The first data set consisted of the measurements of EGH and all Hungarian data except the measurements of the 2nd order net of MGH-80 (9400 ties, 621 points, 19 absolute points as fixed values).

The second data set consisted of all the measurements of MGH-2000, absolute points near to the border in the neighbouring countries, and connecting ties across the borders; all together 8464 individual observed gravity differences and their corrections, 561 unknowns (point values and scale factors of gravimeters) 16 absolute points, 539 gravimeter points including 8 Austrian and 42 Slovakian ones.

The adjustments provided a possibility for comparing the results obtained from two data sets having different homogeneity, and a comparison of adjustments carried out by ELGI and Czech and Slovak partners with basically different methods.

In the first version of adjustment the yearly scale factor was determined for each gravimeter separately. Then the obtained scale factors were included in the final adjustment.

In the second version of adjustment one scale factor was determined for each gravimeter for the whole time interval (1969–95) of measurements so that the scale factors were taken as unknowns in the adjustment.

Iterations were carried out in both versions and with both data sets. The best results were obtained in all cases when one scale factor was determined for each gravimeter for the whole time interval of the observation campaign. This version was accepted as the final adjustment of MGH–2000.

10. Conclusions

Table IV. contains the scale factors and the mean errors of the gravimeters. Comparison of the standard errors gives information about the quality

I	Gravimeters	Calibration factor	Number of observations	Mean error (mGal)
1	LCR–D 9	1.000181	24	0.009
2	LCR–G 176	0.996145	2	0.001
3	LCR–G 220	0.996926	116	0.017
4	LCR–G 625	0.999996	24	0.014
5	LCR–G 779	1.001211	16	0.014
6	LCR–G 821	1.000654	85	0.017
7	LCR–G 919	1.000916	234	0.020
8	LCR–G 963	1.000724	449	0.019
9	LCR–G 1011	1.003474	6	0.010
10	LCR–G 1919	1.000370	2198	0.015
11	Sharpe 174	1.000306	189	0.020
12	Sharpe 176	1.000867	44	0.019
13	Sharpe 181	0.999468	1887	0.020
14	Sharpe 184	0.999510	3	0.009
15	Sharpe 197	0.999682	117	0.024
16	Sharpe 228	0.999932	27	0.014
17	Sharpe 253	1.025750	42	0.017
18	Sharpe 256	1.000104	1931	0.017
19	Sharpe 280	0.999739	200	0.018
20	Worden 961	0.999628	135	0.026
21	Worden 971	0.997728	557	0.031
22	Worden 978	0.998789	178	0.021
Σ			8464	

Table IV. Calibration factors and RMS errors of adjustment for gravimeters applied in the MGH–2000 network

IV. táblázat. Az MGH–2000 méréseinél alkalmazott graviméterek kiegyenlítésből származó méretarány-tényezői és átlagos mérési hibái

of instruments but it must be taken into account that the numbers of observations carried out with the individual gravimeters are significantly different. *Figure 6* represents the histogram of the corrections; it can be seen that 90 % of them are less than 40 μGal and only 2 % higher than 100 μGal, but the weight of this latter 2 % is so small that it hardly influenced the

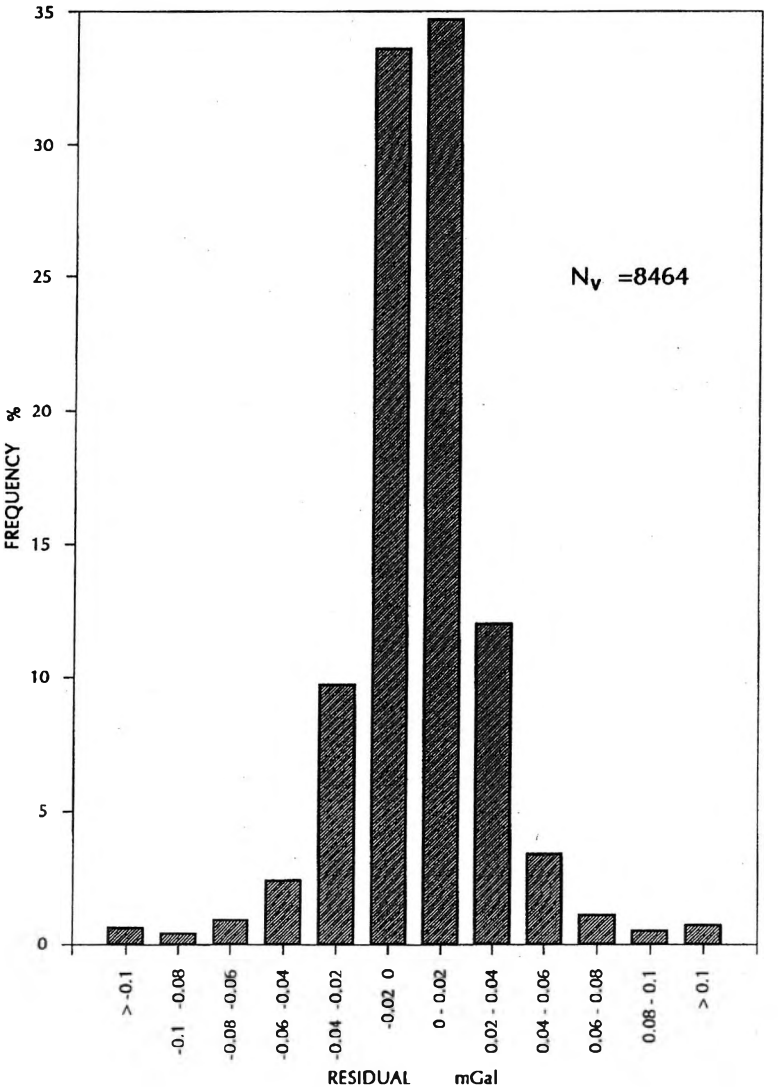


Fig. 6. Histogram of residuals

6. ábra. A mérési javítások hisztogramja

adjusted values. In *Fig. 7* the errors of the adjusted gravity are presented. Most of the points are in the 6–10 μGal error interval, but for those points where the number of measurements are small greater errors are obtained. The error of unit weight of the adjusted network is ±20 μGal.

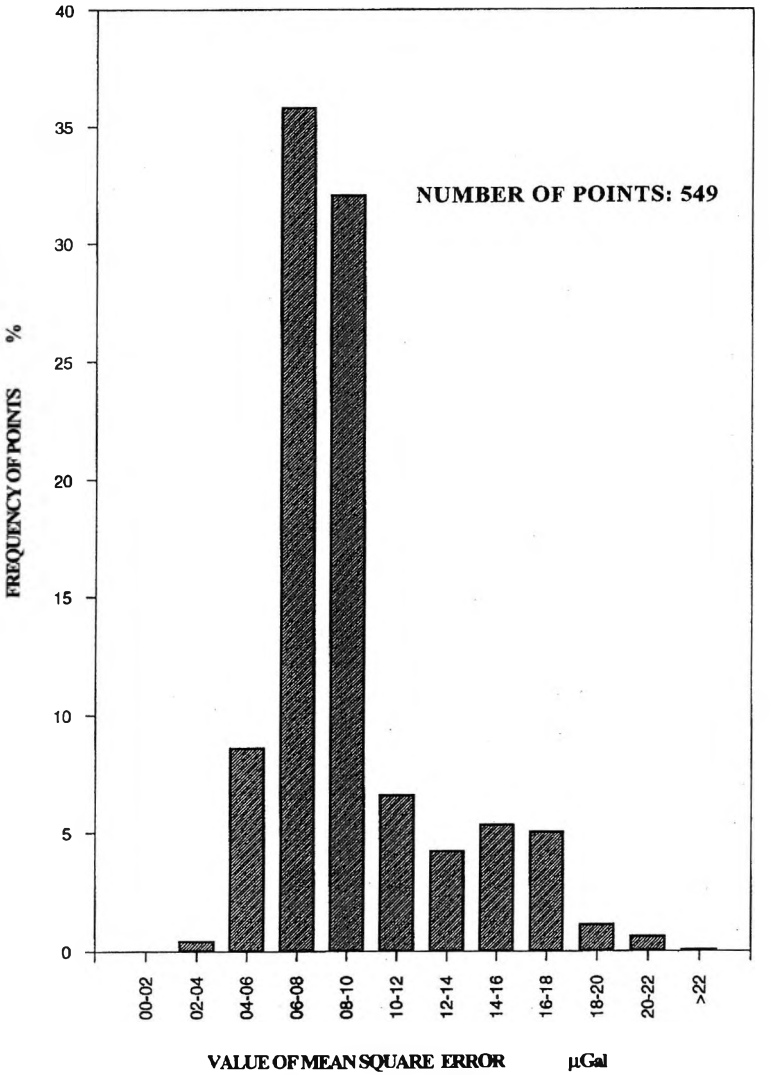


Fig. 7. Histogram of mean square error frequency of adjusted gravity
7. ábra. A kiegyenlített nehézségi értékek középhibáinak hisztogramja

To check the quality of the adjustments two comparisons were carried out. **In the first** one we compared the adjusted values of identical points obtained from the two different data sets (EGH and MGH) calculated by the second version of ELGI adjustment, the deviations between the adjusted values were from 1–3 μGal only, in spite of the different number of observations, different extent of area involved and different number of gravimeters in the calculations. The small values indicate that the national gravity scale depends mainly on the absolute points, and relative measurements of the country and far away points do not influence the results in a perceptible way.

In the second comparison the results of the Czech and Slovakian free adjustments and the Hungarian constrained adjustment were involved. The deviations on the Hungarian base points between the three adjustments were in the 0–20 μGal interval. The result is very satisfying, taking into account the different size of data sets involved and the different methods ('free' and 'constrained') of adjustments. Significant deviation was not experienced between the datum levels. From the 105 compared points only three (Esztergom, Kunhegyes, Nyiregyháza) differed by more than 30 μGal.

A similar comparison was carried out between the Austrian and Hungarian gravity datum. Based on 8 points the Hungarian datum proved to be higher than the Austrian by 18 μGal.

In the adjustment of the 'Unified European Gravity Network' 94' five Hungarian base points were involved [BOEDECKER et al. 1994]. Based on the five points the Hungarian datum is higher by 14 μGal than the international one (see *Table. V*).

UEGN' 94	MGH–2000	UEGN'94	MGH–2000	difference in microgal
number and	name of point	gravity value	in microgal	
1835 FRTOD	4111 Fertőd	980824222 ± 8.0	980824234 ± 4.9	12
1836 HGYEHAL	4122 Hegyeshalom	980844449 ± 12.0	980844460 ± 7.0	11
1837 KESZG	85 Kőszeg	980784705 ± 15.0	980784713 ± 5.0	8
1838 SPRO	4105 Sopron	980808350 ± 14.0	980808375 ± 5.4	25
1839 VELCJ	4112 Völcsej	980802189 ± 14.0	980802203 ± 4.1	14

Table V. Comparison of UEGN' 94 and MGH–2000 networks based on identical Hungarian points

V. táblázat. Az UEGN'94 és az MGH–2000 hálózat összehasonlítása a közös magyarországi pontok alapján

11. Plans for the near future

— Based on the plans of the International Gravity Commission two more absolute gravity points will be measured in Hungary during 1996–97 (Debrecen és Zalalövő).
— The next adjustments of the Unified European Gravity Network will be carried out in 1998 [BOEDECKER et al. 1994]. Since Hungary's intention is to join the European Network the relative gravity measurements between the absolute points should be concluded in 1997 (see Fig. 2).
— A further task is to connect the new and the old networks in order to have the transformation equation between them.
— It is important to provide maintenance and restoration of the gravity base points, and to ensure the necessary funds for the three yearly reobservation of the absolute points.

Acknowledgement

The author wishes to express his gratitude to the National Committee for Technological Development (OMFB) and the U.S.–Hungarian Science and Technology Joint Fund (MAKA) which organizations supported the measurements of MGH–2000 (MEC–94–0508 and JF–369 contracts). Their financial assistance rendered possible the completion of the project. Special gratitude is due to DMA for their support of ELGI. He is indebted to D. RUESS for carrying out the absolute measurements with the JILAG–6 gravimeter, to A. SÁRHIDAI for the adjustments, and to all local and foreign partners who assisted in his work with advice and discussions.

REFERENCES

ÁDÁM J., KENYERES A., BORZA T., CSAPÓ G., LÉVAI P., NÉMETH Zs., TÓTH L. 1994: Some GPS Activities in Hungary Related to the Use of EUREF — National Report of Hungary. Veröffentlichungen der Bayerischen Komm. für die Internat. Erdmessung der Bayerischen Akademie der Wissenschaften, Astro.- Geod. Arbeiten, 54 p.

BOEDECKER G. 1993: Ein Einheitliches Schweregrundnetz für Europa: Unified European Gravity Network (UEGN). Zeits.f. Verm.wesen **9**, 9, pp. 422–428

BOEDECKER G., MARSON I., WENZEL H. G. 1994: The adjustment of the Unified European Gravity Network (UEGN' 94). Internat. Association of Geodesy Symposia, 113, Gravity and Geoid, Springer, pp. 82–91

CSAPÓ G. 1981: Determination of the value of gravity acceleration using ballistic laser gravity meter in Hungary (in Hungarian). Geodézia és Kartográfia **33**, 3, pp. 176–180

CSAPÓ G., SÁRHIDAI A. 1985: Planning of gravity networks by simulation methods (in Hungarian). Geodézia és Kartográfia **37**, 4, pp. 251–255

CSAPÓ G. 1987: Some practical problems of absolute gravity determination (in Hungarian). Geodézia és Kartográfia 1987/2

CSAPÓ G., SÁRHIDAI A. 1990a: The new Gravity Base Network of Hungary (MGH-80) (in Hungarian). Geodézia és Kartográfia **42**, 2, pp. 110–116

CSAPÓ G., SÁRHIDAI A. 1990b: Adjustment of the New Hungarian Gravimetric Network (MGH-80) (in Hungarian). Geodézia és Kartográfia **42**, 3, pp.181–190

CSAPÓ G., MEURERS B., RUESS D., SZATMÁRI G. 1993: Interconnecting gravity measurements between the Austrian and the Hungarian network. Geophysical Transactions **38**, 4 , pp. 251–259

CSAPÓ G., SZATMÁRI G., KLOBUSIAK M., KOVACIK J., OLEJNIK S., TRÄGER L. 1994: Unified Gravity Network of the Czech Republic, Slovakia and Hungary. International Association of Geodesy Symposia, 113, pp. 72–81

CSAPÓ G., SZATMÁRI G. 1995: Apparatus for moving mass calibration of LaCoste–Romberg feedback gravity meters. Metrologia, iss. Gravimetry **32**, 3 , pp. 225–230

FRIEDERICH J. 1993: Absolute Gravity Campaign, Hungary, Publ. GGB-94-001 DMA Aerospace Center, USA

HOLUB S., SIMON Z., BROZ J. 1986: Tidal Observations with Gravity Meter Gs15 No. 228 at Station Pecny. Travaux Geoph. **XXXIV**, pp. 584–593

KRAUTERBLUTH K. 1995: Absolute Gravity Survey, Hungary, 1995. Publ. SMWD3–95–033, DMA Aerospace Center, USA

MEURERS B. 1994: Problems of gravimeter calibration in high precision gravimetry. Internat. Association of Geodesy Symposia 113, pp. 19–26

RENNER J., SZILÁRD J. 1959: Gravity Network of Hungary. Acta Techn. Academiae Scientiarum Hungaricae **XXIII**, 4, pp. 365–395

SÁRHIDAI A. 1986: Design of fundamental gravity networks based on the approximation of the given variant-covariant matrix. Bull. Geod. **60**, pp. 355–376

SOHA G. 1986: A robust adjustment with weights depending from measurement corrections 1986. Geodézia és Kartográfia, 1986/4

SPITA W. 1994: Absolute Gravity Campaign, Hungary, Publ. GGB-94-056 DMA, Aerospace Center, USA

Az új magyarországi gravimetriai alaphálózat (MGH-2000)

CSAPÓ Géza

A közelmúltban fejeződtek be a közép-európai országok közös gravimetriai hálózatának kialakítására végzett munkák, amelyek keretében a magyarországi alaphálózatot is korszerűsítették. Ismertetésre kerülnek a korábbi hálózatok, azok feladata és a fejlesztés szükségszerűségének okai. Az új hálózat méretaránya az SI mértékrendszerben adott; ezt nagyszámú abszolút állomás telepítésével és mérésével biztosították. A szerző bemutatja az alkamazott gravimetriai módszereket, az adatfeldolgozás és kiegyenlítés folyamatát, majd összehasonlítást végez a cseh, a szlovák, az osztrák és az „Egységes Európai Gravimetriai Bázishálózat", valamint a magyar hálózat között.

GEOPHYSICAL TRANSACTIONS 1997
Vol. 40. No. 3-4. pp. 145-153

Group traveltime estimation by wavelet transform with linear chirp as the basic wavelet

Péter SCHOLTZ*

A method for group traveltime estimation is analysed in this paper. The method is based on the wavelet transform of a seismic signal when the basic wavelet of the wavelet transform is a linear chirp modulated Gaussian. The technique is applied to a synthetic dispersive seismic trace which is calculated for a simple wave guiding sequence. Several tests were carried out on synthetic Love channel waves to compare the result of the above method with the performance of the wavelet transform with a Morlet wavelet as the basic wavelet. Where the frequency-group traveltime curve of the seismic signal can be linearly approximated the wavelet transform with a linear chirp basic wavelet gives a superior result compared to the wavelet transform with a Morlet wavelet as the basic wavelet. It is shown that peaks of the moduli of the wavelet coefficients are sharper, noisy arrivals can be selected easily, and inseparable dispersive arrivals are well resolved.

Keywords: wavelet transform, channel wave, traveltime, dispersion pattern

1. Introduction

There are several methods which are used to gain geological information by the use of dispersive waves. Examples of these are the in-seam reflection technique where guided waves map faults or discontinuities with the help of recompressive filtering [BAKI et al. 1988]; the static correction computation, which is based on velocity dispersion analysis of surface waves [MARI 1984]; or the study of near surface inhomogeneities by group traveltime tomography

* Eötvös Loránd Geophysical Institute of Hungary (ELGI),
H-1145 Budapest, Kolumbusz u. 17-23, Hungary

Manuscript received: 26 September, 1996

of Rayleigh and Love waves [DOMBROWSKI, DRESEN 1995]. All of these methods and several others require a non-stationary signal analysing tool.

The wavelet transform (WT) has been proved to be a useful way of getting the frequency–group traveltime (or group velocity) function of dispersive waves [PYRAK-NOLTE, NOLTE 1995]. A good analysing method should make the identification of group arrivals easier even in noisy conditions and separate different waves at close arrival times. To achieve these a linear chirp modulated Gaussian is chosen as the basic wavelet for the WT [GROSSMANNet al. 1989].

2. Method

The wavelet transform of a seismic signal $s(t)$ can be given as

$$W(a,b) = \frac{1}{\sqrt{a}} \int g^*\left((t-b)/a\right) s(t) \mathrm{d}t$$

where g^* is the complex conjugate of the scaled (denoted by a) and delayed (denoted by b) version of the basic wavelet. The basic wavelet should meet the admissibility conditions: $g(t)$ is absolutely integrable, square integrable, band limited and has zero mean [SHENSA 1992]. A very simple and commonly used version of $g(t)$ is the Morlet wavelet [MORLET et al. 1982] viz.

$$g(t) = e^{-\frac{t^2}{2}} e^{i2\pi\nu t}$$

where $i^2 = -1$, ν is the central frequency, and $2\pi\nu \geq 5$. The last of these is needed to meet the admissibility conditions. If the data to be analysed can be approximated locally by linear chirp a better choice could be a linear chirp modulated Gaussian as the basic wavelet for the wavelet transform as

$$g(t) = e^{-\frac{t^2}{2}} e^{i2\pi\left(c+\frac{k}{2}t\right)t}$$

where c is the central frequency, k is the rate of change of the instantaneous frequency $c+kt$. Though a new unknown parameter is introduced (k) into the analysis, it can be used to gain additional information and also a better resolution is expected where the linear chirp approximation is valid. These are achieved by repeating the wavelet transform with a set of k values to form a 3D volume, where at each scale the maximum peaks of the moduli of the wavelet coefficients

are selected at the best fit k rate. The additional parameter, which is provided by the choice of the basic wavelet, is the slope of the frequency- group traveltime function derived from that specific k rate at a given a scale.

3. Examples

To analyse the performance of the above method we created synthetic seismograms for a simple model. Two symmetric half spaces separated by a lower velocity layer form a wave guiding sequence. The rocks are homogeneous and nondissipative. We calculated Love arrivals for the following model parameters:

$\rho_0 = 1500$ kg/m^3,
$\beta_0 = 1000$ m/s,
$\rho = 2500$ kg/m3,
$\beta = 2000$ m/s,
$d = 2$ m

where ρ_0 is the density and β_0 is the shear velocity of the seam; ρ is the density and β is the shear velocity of the surrounding rocks; and d is the thickness of the seam. The source and the receiver were placed in the middle of the seam at an offset of 200 m.

4. Group traveltime estimation by WT

Figure 1 shows a noise free dispersive arrival (left) and its WT (middle) where the moduli of the wavelet coefficients are displayed on the c/a–b (frequency and traveltime) plane at $k = 0$ in the basic wavelet (the result is equivalent to a WT with a Morlet wavelet). The figure shows the k–b slice of the 3D volume at a given a scale ($c/a = 240$ Hz), too (right). This a scale was chosen from the linear section of the frequency–group traveltime curve. $k = 0$ for the first displayed channel and k increases channel by channel. The WT has the sharpest peak at about $k = 800$ Hz/s chirp rate. This occurs when the seismic signal is more closely approximated locally by a linear chirp with the given k rate than with $k = 0$. It makes the group traveltime estimation more accurate and also provides additional information since the k rate gives the slope of the frequency–group traveltime curve at the a scale.

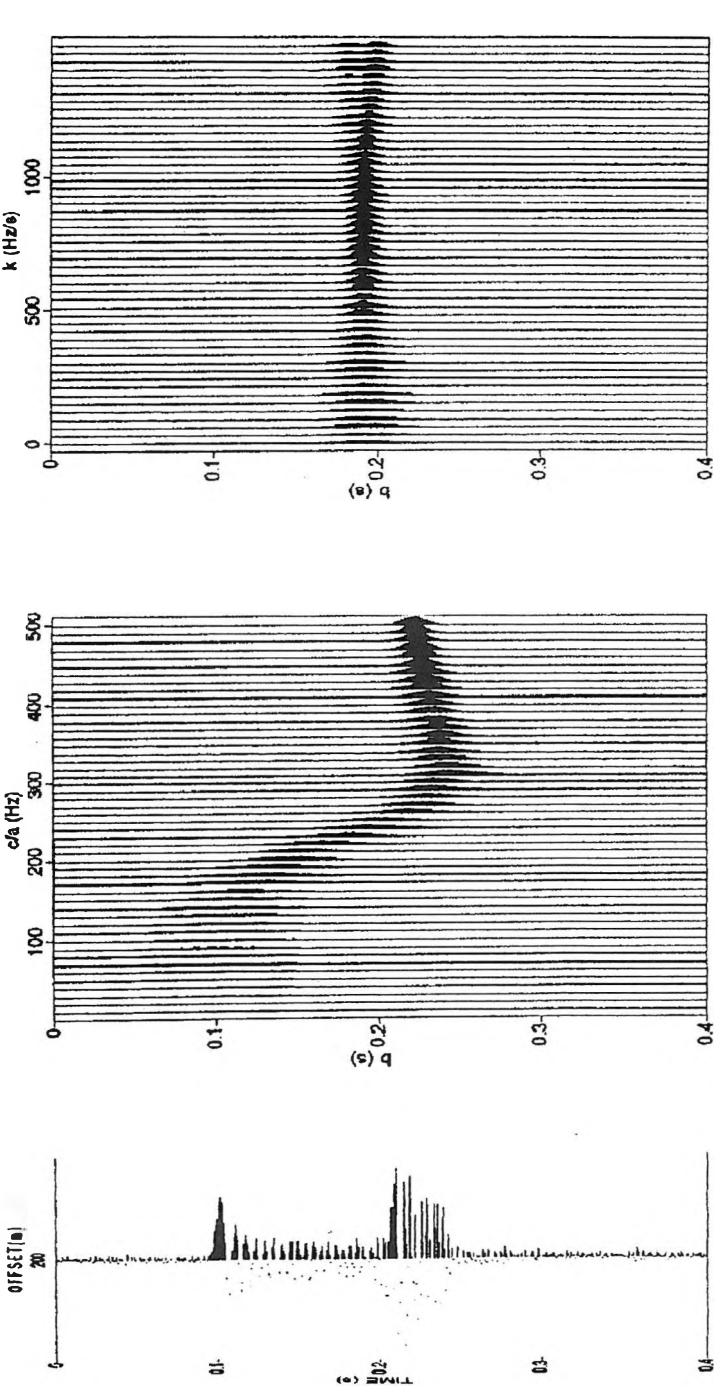


Fig. 1. Noise free synthetic Love channel wave (left), its wavelet transform (middle) and the *k*–*b* slice (right)
1. ábra. Zajmentes szintetikus Love-típusú csatornahullám (balra), wavelet transzformáltja (középen) és a *k*–*b* metszet (jobbra)

Figure 2 shows a noisy dispersive arrival (left) and its WT at $k = 0$ in the basic wavelet (middle), on the *c/a–b* plane (similarly to Figure 1.). The noise is random and its distribution is Gaussian. The level of the noise was chosen to cover the dispersive arrival. Based on the wavelet transform with a Morlet wavelet as the basic wavelet the theoretical frequency–group traveltime curve is hardly recognizable due to the added strong random noise, but the *k–b* slice of the 3D volume at the same *a* scale as previously (right) still enables us to select the right group traveltime at a specific *k* rate accurately enough.

Figure 3 shows two dispersive arrivals when a 200 m offset and a 250 m offset trace were added (left). With this simple example we demonstrate the resolving power of the method. The WT (middle) at $k = 0$ in the basic wavelet, on the *c/a–b* plane (similarly to the previous figures) were able to differentiate between the two separate arrivals only at the higher frequency range. The *k–b* slice of the 3D volume at the same *a* scale as previously (right) where the WT with the Morlet wavelet showed no separate arrivals (middle) the WT with a linear chirp as the basic wavelet at different *k* rates was able to produce distinct arrivals.

5. Conclusions

The group traveltime of a seismic signal can be analysed by the wavelet transform with linear chirp as the basic wavelet. Repeating the WT at different chirp rates a 3D volume can be created where the amplitude peaks at each scale are sharper at the frequency rates which correspond to the linear approximation of the frequency–group traveltime curve at the given frequencies. In this way we get the slope of that curve, the selection is more accurate on noisy data and different arrivals are better separated even at close arrival times than by the WT with a Morlet wavelet as the basic wavelet. Though the examples were calculated on dispersive seismic traces the method can be used whenever the linear approximation of the frequency–group traveltime function of a signal is valid.

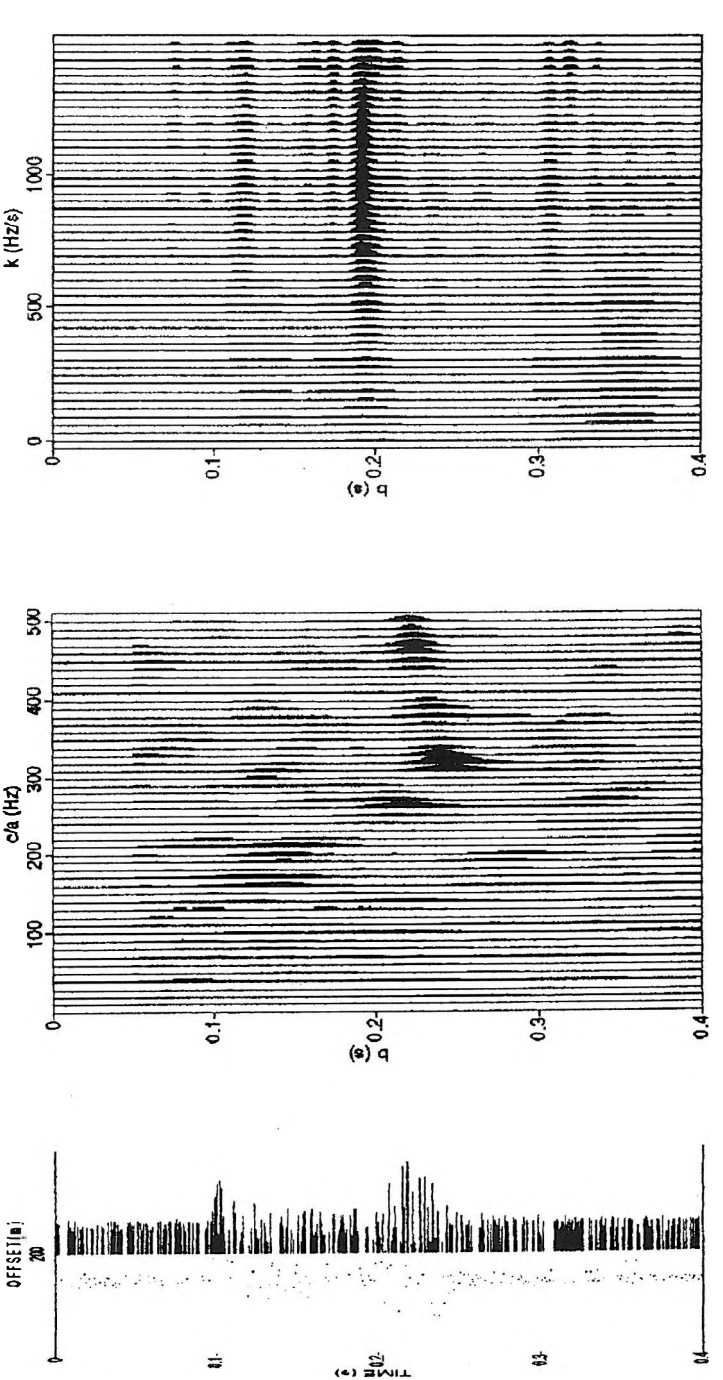


Fig. 2. Noisy synthetic Love channel wave (left), its wavelet transform (middle) and the *k–b* slice (right)

2. *ábra.* Zajos szintetikus Love-típusú csatornahullám (balra), wavelet transzformáltja (középen) és a *k–b* metszet (jobbra)

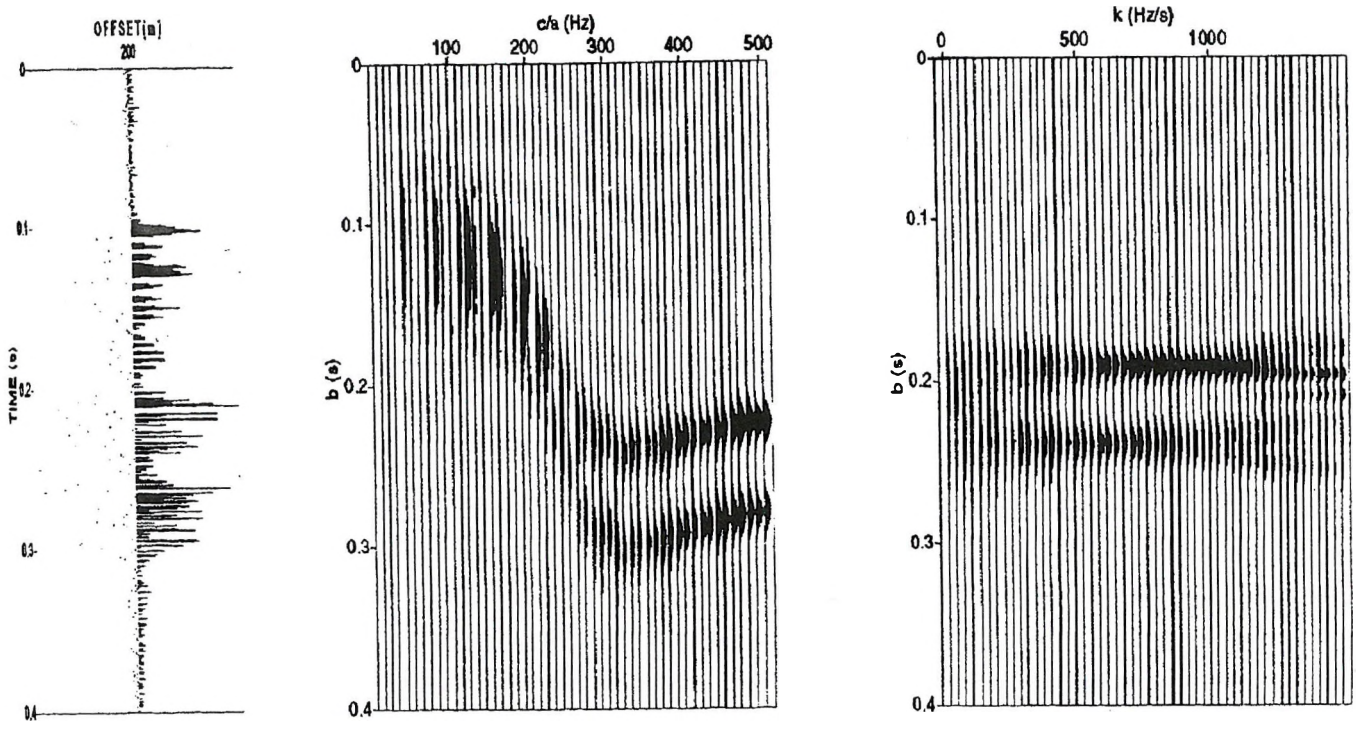


Fig. 3. Added synthetic Love channel waves (left), the wavelet transform (middle) and the *k–b* slice (right)
3. ábra. Szuperponált Love-típusú csatornahullámok (balra), wavelet transzformáltjuk (középen) és a *k–b* metszet (jobbra)

Acknowledgements

In preparing this paper I used the results of OTKA project (No. F 014492) supported by the National Scientific Research Fund.

REFERENCES

BAKI GY., BODOKY T., CZILLER E. and SCHOLTZ P. 1988: Possibilities and limitations of recompressive filtering in the processing of seam-wave seismic surveys. Geophysical Transactions **33**, pp. 221-236

DOMBROWSKI B. A., DRESEN L. 1995: Localization of near surface inhomogeneities with surface waves. EAGE 57th Conference and Technical Exhibition-Glasgow, Extended Abstracts

GROSSMANN A., KRONLAND-MARTINET R., MORLET J. 1989: Reading and understanding continuous wavelet transforms, in Wavelets Time-Frequency Methods and Phase Space. Springer-Verlag pp. 2-20

MARI J. L. 1984: Estimation of static corrections for shear-wave profiling using the dispersion properties of Love waves. Geophysics **49**, pp. 1169-1179

MORLET J., ARENS G., FOURGEAU E., GIARD D. 1982: Wave propagation and sampling theory—Part II: Sampling theory and complex waves. Geophysics **47**, pp. 222-236

PYRAK-NOLTE L. J., NOLTE D. D. 1995: Wavelet analysis of velocity dispersion of elastic interface waves propagating along a fracture. Geophysical Research Letters **22**, pp.1329-1332

SHENSA M. J. 1992: The discrete wavelet transform: wedding the a trous and mallat algorithms. IEEE Trans. Signal Processing **40**, pp. 2464-2482

Csoport futási idő meghatározás wavelet transzformált segítségével lineárisan változó frekvenciájú alap wavelet alkalmazásával

SCHOLTZ Péter

A dolgozat wavelet transzformáció segítségével történő futási idő meghatározást mutat be. A diszperzív szeizmikus hullám analízisét lineárisan változó frekvenciájú jellel modulált Gauss-típusú alap wavelet alkalmazásával végzi. A módszert egy egyszerű hullámvezető modell alapján számítható szintetikus diszperzív szeizmikus csatornán vizsgálja. Többféle tesztet végez Love-típusú csatornahullámokon a közölt módszer és a már régebben is alkalmazott wavelet transzformációs eljárás összehasonlítására, melynél az alap wavelet a Morlet-típusú wavelet. Bemutatja,

hogy ahol a szeizmikus jel frekvencia- csoport futási idő függvénye egyenessel közelíthető, ott a lineárisan változó frekvenciájú jellel modulált Gauss-típusú alap wavelet használata jobb eredményt nyújt, mint a Morlet-típusú wavelet: a wavelet transzformáció eredményeként kapott wavelet koefficiensek képe élesebb, zajos beérkezések könnyebben azonosíthatók és az időben közeli beérkezések jól felbonthatók.

GEOPHYSICAL TRANSACTIONS 1997
Vol. 40. No. 3-4. pp. 155-174

VSP inversion — A new method using edge detection

Zoltán WÉBER*

A novel approach for obtaining interval velocity estimates from vertical seismic profiles (VSPs) is described. The edge detection technique, well known in image processing, can supply large amount of gradient data from the VSP image. Statistical evaluation of several gradient values obtained for the same depth interval yields reliable velocity estimates and error limits.

The preprocessing sequence contains only wavelet filtering and wavefield separation. True amplitude processing and definition of initial velocity model are not necessary. The method can handle offset (OVSP) data as well. In order to obtain the interval velocities from OVSP data, the gradient values are multiplied by a correction factor, which is determined by ray trace modelling.

Several tests on synthetic and real data have proved that the method is efficient and relatively independent of random noise and model errors. The procedure can be carried out economically on personal computers.

Keywords: VSP inversion, edge detection, most frequent value

1. Introduction

Vertical seismic profiling (VSP) is becoming increasingly recognized as a valuable technique in seismic exploration [GALPERIN 1974, HARDAGE 1985, BALCH and LEE 1984]. Perhaps the most vital information inferred from VSP is the seismic velocity. A major processing task is to determine the velocity-depth function, which becomes especially difficult when the source is offset

* formerly: Geophysical Research Group of the Hung. Acad. Sci. Department of Geophysics, Eötvös University H-1083 Budapest, Ludovika tér 2., Hungary
presently: Seismological Observatory of the Hung. Acad. Sci., H-1112 Budapest, Meredek u. 18., Hungary

Manuscript recevied: 21 August, 1995

some distance from the wellhead. Source offset, however, is often necessary to avoid or reduce the generation of tube waves and to sample the subsurface away from the borehole in order to obtain two-dimensional information.

Several methods are currently in use to compute interval velocities from offset VSP (OVSP) data. The simplest method assumes vertical propagation and computes an apparent velocity $\Delta z/\Delta t$, where Δz is the difference in geophone depths and Δt is the corresponding difference in the arrival times. Better approximation is achieved by assuming slanted straight rays [LASH 1980]. The next step towards more accurate velocities is based on a flat layered model using Snell's law to predict the direction of ray propagation [GRANT and WEST 1965].

An even more sophisticated method is proposed by STEWART[1984], who solves the problem using a generalized linear inversion (GLI) technique. He assumes an initial velocity structure and perturbs it iteratively until the computed traveltimes match the observed ones in a least squares sense. At each iteration a ray tracing technique is used to compute traveltimes. LINES et al. [1984] use a different but related method to determine layer dips from VSP arrival times and sonic log velocities. PUJOL et al. [1985] propose a layer stripping technique and a generalized inversion somewhat similar to that of STEWART [1984].

The methods described so far have inherent limitations. The first two are reasonable approximations only for small offsets. Since traveltimes cannot be determined with sufficient accuracy neither interactively nor automatically, all methods based on arrival times are prone to error [STEWART 1984, BALCH and LEE 1984].

Specifically for zero-offset VSP inversion GRIVELET [1985], MACE and LAILLY [1986], URSIN [1986] and others have proposed GLI methods, which use the whole VSP wavefield. Since the problem is nonlinear, the initial estimate of the acoustic impedance function must be close enough to the real one in order to get the global, and not the local, minimum of the objective function. Because the GLI method utilizes the dynamic features of the VSP data instead of the velocity function, it is the acoustic impedance function which can be estimated and, on the other hand, the amplitudes of the VSP data have to be correctly reconstructed during data processing. Another disadvantage of the GLI technique follows from the huge amount of computations. The calculation of realistic synthetic VSPs and the solution of the large sets of linear equations require extremely fast computers.

The interval velocity estimation technique described in this paper uses a much simpler approach. The gradient of the traveltime curve of the first arrivals is estimated at each geophone depth by means of edge detection, a standard procedure of image processing [e.g. ROSENFELD and KAK 1982]. In order to obtain the interval velocities from OVSP data, the gradient values are multiplied by a correction factor, which is determined by ray trace modelling, assuming a horizontally layered earth model. The preprocessing sequence of the VSP data contains only wavelet filtering and wavefield separation. By virtue of the statistical processing of the gradient data calculated by edge detection, the error of the estimation can be given in addition to the estimated velocities. The definition of an initial velocity model is not necessary.

Several tests on synthetic and real data have proved that the method is efficient and relatively insensitive to random noise and model errors. The procedure needs so little computer time that it can be realized economically even on personal computers.

2. The gradient method based on edge detection

The series of first breaks on a VSP defines the traveltime curve of the first arrivals. As is well known, the gradient of this traveltime curve (i.e. $\Delta z/\Delta t$) is closely connected with the wave velocity: in the case of zero-offset VSP and horizontal layering the gradient is equal to the interval velocity.

The gradient of the traveltime curve as a function of depth can be estimated by edge detection. On the VSP, which is considered a two-dimensional image, the gray levels change rapidly in the vicinity of the first arrivals because of the large amplitudes. The dominant direction of these rapid changes should be found at each geophone level. The Sobel operator, which is a 3_3 point two-dimensional edge detection operator, determines edge directions at each data point of the VSP. At a given geophone depth the most probable value of the gradient, characterizing the traveltime curve of the first arrival at this depth, can be estimated by a statistical method using several gradient values in an appropriately chosen time gate around the first break. The width of this time gate is usually between about 100 and 200 ms. Such a time gate includes 50-100 gradient values if the sampling interval is 2 ms. The statistical processing of the large number of gradient data involves that in addition to the estimated velocity values the error of the estimation can

also be given. In this paper the most frequent value procedure is used because of its robustness and resistance [STEINER 1988].

Experience gained during the tests with the gradient method suggests that edge detection works more effectively on the instantaneous phase section calculated from the VSP than on the original VSP data. (For the computation of instantaneous phase see e.g. TANER et al. [1979]). The reason for this is that the instantaneous phase section contains sharper 'jumps' than the original section. As a consequence, the gradient values obtained from the instantaneous phase are more accurate than those calculated from the original VSP data.

3. Application of the gradient method to OVSP data

In the case of zero-offset VSP and horizontally layered media, gradients provide good estimates of the velocities, but this is not true for gradients computed from OVSP data.

In order to find the relation between the gradient values and the velocities in the latter case, let us assume that the traveltimes of the first arrivals to the depths z_i and z_{i+1} are t_i and t_{i+1}, respectively, and the wave arriving to the geophone at depth z_{i+1} travels the path length Δs_i in the time interval $\Delta t_i = t_{i+1} - t_i$ (*Fig. 1*). Because, by time t_i, this wave is already in the ith layer, the velocity in this depth interval is $\Delta s_i / \Delta t_i$ to a good approximation. (Here the depth interval $[z_i, z_{i+1}]$ is considered as the ith layer.) Moreover, let us assume that the velocity is equal to the gradient $\Delta z_i / \Delta t_i$ multiplied by a correction factor C_i, i.e.

$$\frac{\Delta s_i}{\Delta t_i} = C_i \frac{\Delta z_i}{\Delta t_i},$$

where Δz_i denotes the thickness of the ith layer. It is obvious that the correction factor is

$$C_i = \frac{\Delta s_i}{\Delta z_i}$$

In the case of zero-offset VSP, Δs_i is equal to Δz_i, i.e. the correction factor C_i is equal to 1. Now the question is how to determine the correction factor C_i if the offset is not zero.

Let us assume that the interval velocities are known to depth z_i, and the calculated traveltime to this depth is t_i. The traveltime t_{i+1} to depth z_{i+1} can also be calculated if an arbitrary velocity value v_i is assumed in the ith

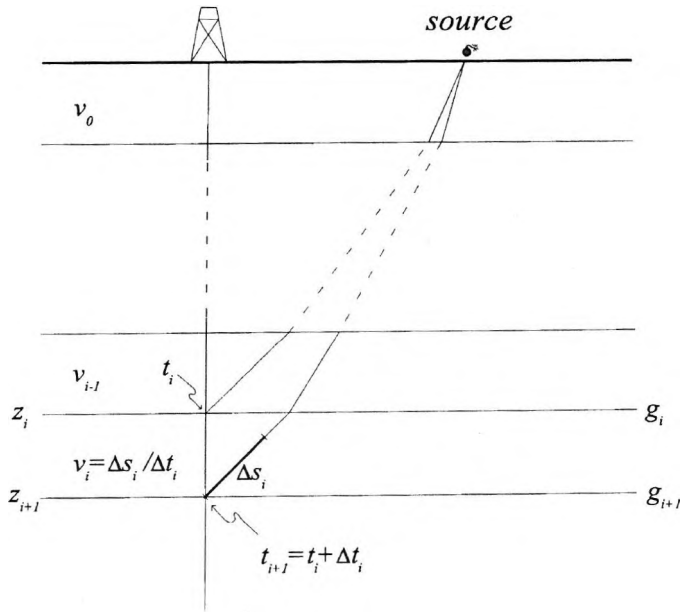


Fig. 1. Offset VSP geometry for velocity estimation in a horizontally layered earth model using the gradient data g_i

1. ábra. Offszetes VSP mérés geometriája horizontális rétegzettség mellett a g_i irányadatok alapján végzett sebességbecsléshez

layer. In the time interval $[t_i, t_{i+1}(v_i)]$ the wave travels the path length $\Delta s_i = v_i\,[t_{i+1}(v_i)-t_i]$, i.e. the correction factor is

$$C_i = \frac{v_i\left[t_{i+1}(v_i) - t_i\right]}{\Delta z_i}. \tag{1}$$

Velocity v_i has been chosen correctly if it is equal to gradient g_i of the traveltime curve of the first arrival multiplied by the correction factor C_i:

$$C_i g_i = v_i\,. \tag{2}$$

Substituting Eq. (1) into Eq. (2):

$$\frac{t_{i+1}(v_i) - t_i}{\Delta z_i}\, g_i = 1 \tag{3}$$

If Eq. (3) is not satisfied, velocity v_i should be modified until the equality becomes valid within an acceptable tolerance.

In order to calculate the raypath and the traveltime of the first arrival, a system of nonlinear equations should be solved. The Newton-Raphson

method has been used here (see Appendix). This iterative method requires an initial approximation. The iteration is convergent only if this initial approximation is close enough to the actual solution. In our problem the simplest way to define an initial guess is to assume a straight raypath between the source and the geophone. However, this approximation is close enough to the actual solution only if the unknown velocity v_i is close enough to the velocity in the *(i*-1*)*st layer.

Based on the foregoing discussion, the following procedure is proposed for finding the correct velocity v_i. In the first step let the unknown velocity v_i be equal to the (known) velocity v_{i-1}. Under this condition the traveltime t_{i+1} can be calculated. If the left hand side of Eq. (3) is greater than 1, then t_{i+1} is too large, i.e. velocity v_i should be increased, otherwise v_i should be decreased. According to the result of the first step, v_i should be altered by a preset velocity increment Δv. The raypath calculated in the previous step can be used as an initial approximation in the calculation of the new raypath and traveltime t_{i+1} of the first arrival. If the left hand side of Eq. (3) is still not close enough to 1, then velocity v_i should again be altered by Δv. This procedure is continued until Eq.(3) holds to an acceptable tolerance.

4. Synthetic examples

The features of the inversion method described in the previous sections have been studied on synthetic data. By means of the examples we examine how the method is influenced by random noise and model errors.

The velocity and density functions used in synthetic zero-offset VSP calculation consist of 1000 layers. Each layer is 2 metres thick. The calculated synthetic VSP consists of 139 traces, the distance between the equally spaced geophones (traces) is 10 metres, the sampling rate is 2 ms. A Klauder wavelet with peak frequency of 30 Hz is assumed. The forward method used here to compute the synthetic VSP takes into consideration the effects of absorption and dispersion [GANLEY 1981].

The downgoing wavefield used in the velocity estimation and the corresponding instantaneous phase section are shown in *Fig. 2*. Since the instantaneous phase section contains much sharper 'jumps' than the original VSP, the gradient data were calculated from the instantaneous phase section. For the velocity estimation 200 ms wide time gates were used around the first break at each geophone level. The final results are shown in *Fig. 3*. (The

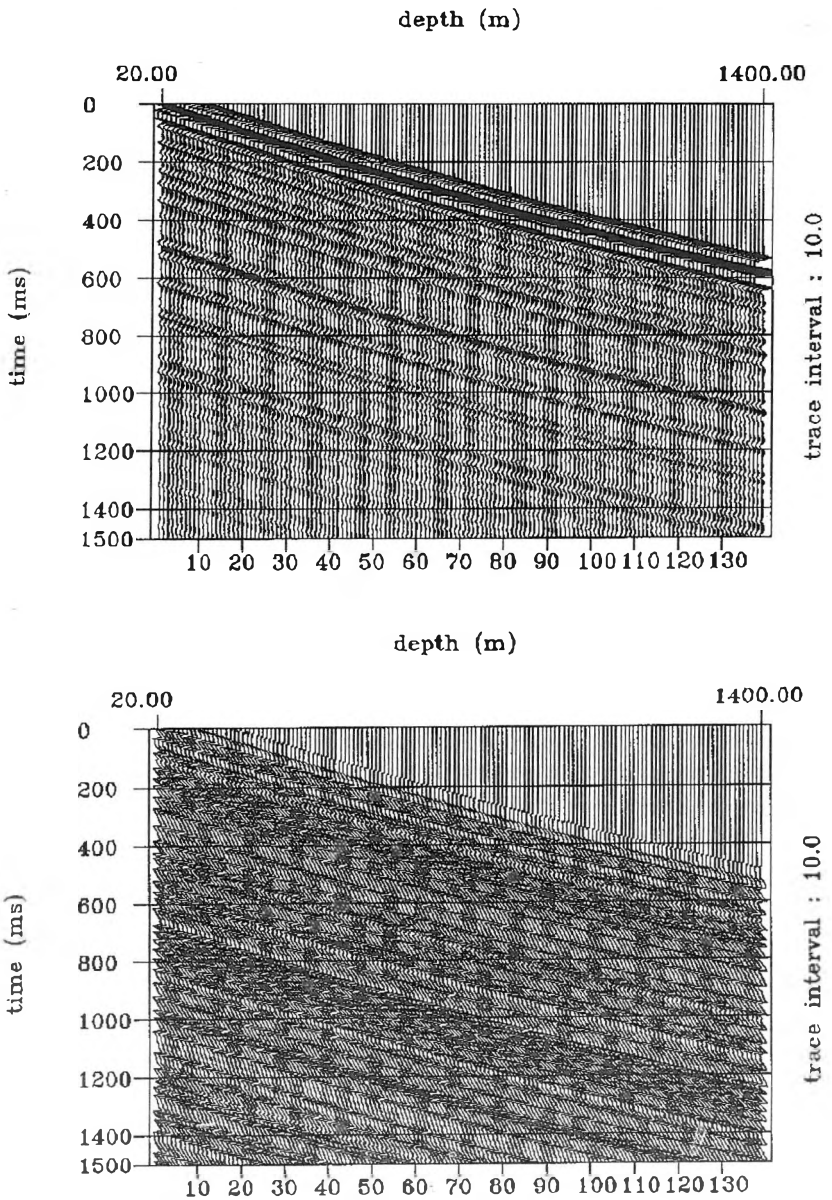


Fig. 2. above: Synthetic zero-offset downgoing wavefield. (wavelet: Klauder with peak frequency of 30 Hz; source depth 10 m). *below*: The corresponding instantaneous phase section

2. ábra. felül: Szintetikus zérus offszetű VSP lefelé haladó hullámtere (wavelet: 30 Hz-es Klauder wavelet; a forrás mélysége 10 m). *alul*: A megfelelő pillanatnyi fázis szelvény

estimation uncertainty shown in *Fig. 3b* corresponds to about 66 per cent confidence level, i.e. the true velocity values are between the lower and the upper limits with a probability of about 66 per cent.)

The results show that the velocities estimated from a noise-free synthetic VSP are very close to the true velocities. The mean difference between the two velocity functions is only 1.2 per cent and the estimation uncertainty is less than $\pm$5-6 per cent of the true values. (Throughout this paper *difference* is meant in $\mathcal{L}_1$ norm.)

The most frequent value procedure, which is used here as a tool of statistical processing, maximizes the quotient $n^2_{eff}(\varepsilon)/\varepsilon$, where $n_{eff}(\varepsilon)$ is the so-called number of effective data (the sum of the weights used in the calculation of the most frequent value) and ε denotes the so-called dihesion, expressing the degree of gathering of the data around the most frequent value [STEINER 1988]. *Fig. 3c* illustrates this quotient normalized by n^2, where n is the number of the gradient data used in the course of the statistical processing. This curve informs us about the quality of the VSP data: if $n^2_{eff}(\varepsilon)/\varepsilon$ is very small (e.g. at depths 120 and 330 m), the deviation of the gradient data is large and, as a consequence, the estimation uncertainty increases. On the other hand, when the quality of the data is good, i.e. when the gradient values are close to each other, $n^2_{eff}(\varepsilon)/\varepsilon$ is fairly large and the estimation uncertainty becomes small.

The effect of additive random noise has been investigated by several synthetic VSPs. The standard deviation of the noise was calculated for each trace from the energy of the noise-free VSP trace, random numbers were added to each sample of the synthetic impulse response and the sum was convolved by a Klauder wavelet with peak frequency of 30 Hz. In most cases the inversion procedure used 200 ms wide time gates around the first breaks (i.e. 100 data).

When the additive noise is less than about 6-7 per cent, the differences between the estimated and the true velocities are negligible. In the case of 10 per cent noise the deviation of the estimated velocities from the true velocities becomes visible with a mean difference of 2.6 per cent and the estimation uncertainty is between $\pm$10 and $\pm$20 per cent in most depth intervals (*Fig. 4*). Comparison of the $n^2_{eff}(\varepsilon)/\varepsilon$ plots in Figs. 3 and 4 shows that the quality of the noise-corrupted data is lower than that of the noise-free data, thereby resulting in the increased estimation error and uncertainty. The estimates can be improved if shorter time gates are considered around the first breaks, where the signal to noise ratio remains sufficiently high. If this time interval

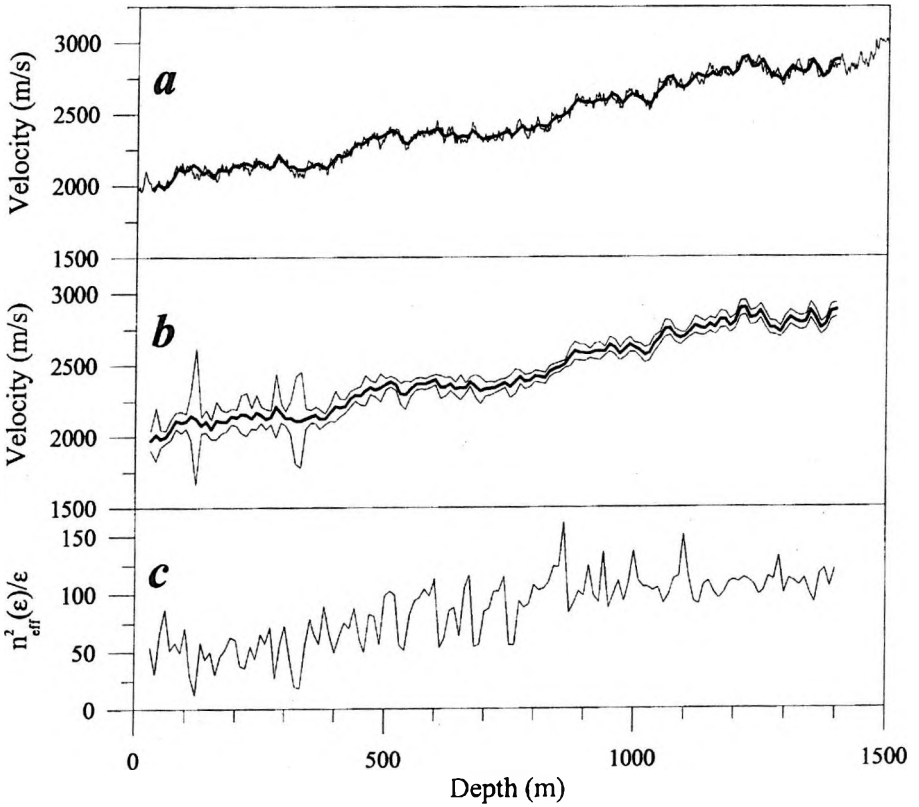


Fig. 3. (a) True velocities (thin line) and estimated velocities (thick line) in case of no noise. *(b)* Upper and lower limits (thin lines) of estimated velocities (thick line) illustrating estimation uncertainties. The true velocities are between the two limits with a probability of about 66 per cent. *(c)* $n^2_{eff}(\varepsilon)/\varepsilon$ plot illustrating the quality of the processed VSP data

3. ábra. (*a*) A valódi (vékony vonal) és a becsült (vastag vonal) sebességek zajmentes esetben. (*b*) A becsült sebességek (vastag vonal) bizonytalanságát illusztráló alsó és felső hibahatárok (vékony vonalak). A valódi sebességek kb. 66%-os valószínűséggel esnek a két hibahatár közé. (*c*) Az $n^2_{eff}(\varepsilon)/\varepsilon$ görbe, amely a feldolgozott VSP adatok minőségét jellemzi

is only 50 ms (25 data), the inversion of the VSP corrupted by 10 per cent additive noise again gives fairly good final results: the estimation uncertainty considerably decreases and the mean difference between the estimated and the true velocity functions diminishes to 1.3 per cent.

The gradient of the traveltime curve of the first arrivals is equal to the wave velocity only if the medium is horizontally stratified, even for zero-off-

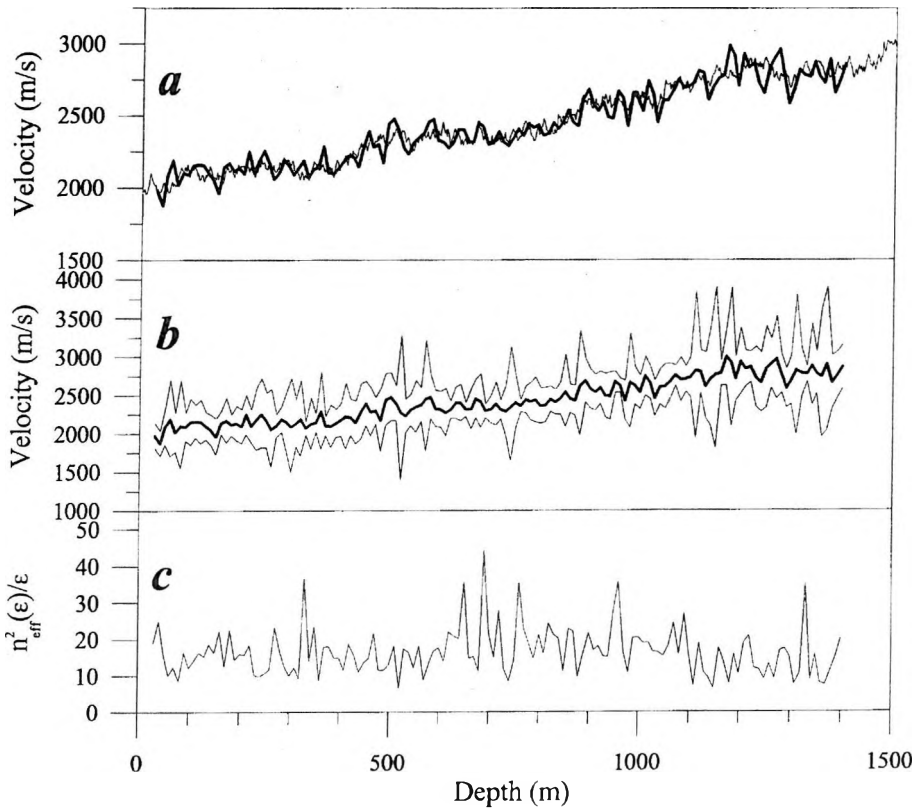


Fig. 4. (a) True velocities (thin line) and estimated velocities (thick line) in case of 10 per cent coloured additive noise. *(b)* Upper and lower limits (thin lines) of estimated velocities (thick line) illustrating estimation uncertainties. The true velocities are between the two limits with a probability of about 66 per cent. *(c)* $n^2_{eff}(\varepsilon)/\varepsilon$ plot illustrating the quality of the processed VSP data

4. ábra. (*a*) A valódi (vékony vonal) és a becsült (vastag vonal) sebességek 10%-os additív színes zaj esetén. (*b*) A becsült sebességek (vastag vonal) bizonytalanságát illusztráló alsó és felső hibahatárok (vékony vonalak). A valódi sebességek kb. 66%-os valószínűséggel esnek a két hibahatár közé. (*c*) Az $n^2_{eff}(\varepsilon)/\varepsilon$ görbe, amely a feldolgozott VSP adatok minőségét jellemzi

set VSPs. Thus it is worth investigating how the estimated velocity values differ from the true velocities if the layer boundaries are not horizontal.

In order to carry out this investigation, several two-dimensional geological models were defined with parallel but not horizontal layers, and then noise-free synthetic VSPs were calculated for these models by ray tracing with the following parameters: offset 10 metres; depth of the uppermost

geophone 100 metres; geophone spacing 10 metres; and the wavelet is a damped sine function with a peak frequency of 30 Hz. The density is assumed to be constant throughout the models. For the inversion of the synthetic VSPs, a 200 ms wide time gate was used around the first break at each geophone level.

When the layer dips are smaller than about 10-15 degrees, the estimated velocities are practically equal to the true ones (*Fig. 5a*). In the case of layers with a dip of 20 degrees, the differences between the estimated and the true velocity values are somewhat greater, but their mean value is still only 1.3 per cent (*Fig. 5b*). This estimation error systematically increases with depth; this can be explained only by the dip of the layers, which dip was not taken into account during the inversion procedure. In the final analysis, the inversion method can give reliable results even if the dip of the layers is - from a practical point of view - large.

When inverting OVSPs, in the course of the calculation of the correction factor C, a horizontally layered earth model is assumed. In order to investigate how the accuracy of the proposed inversion technique depends on model

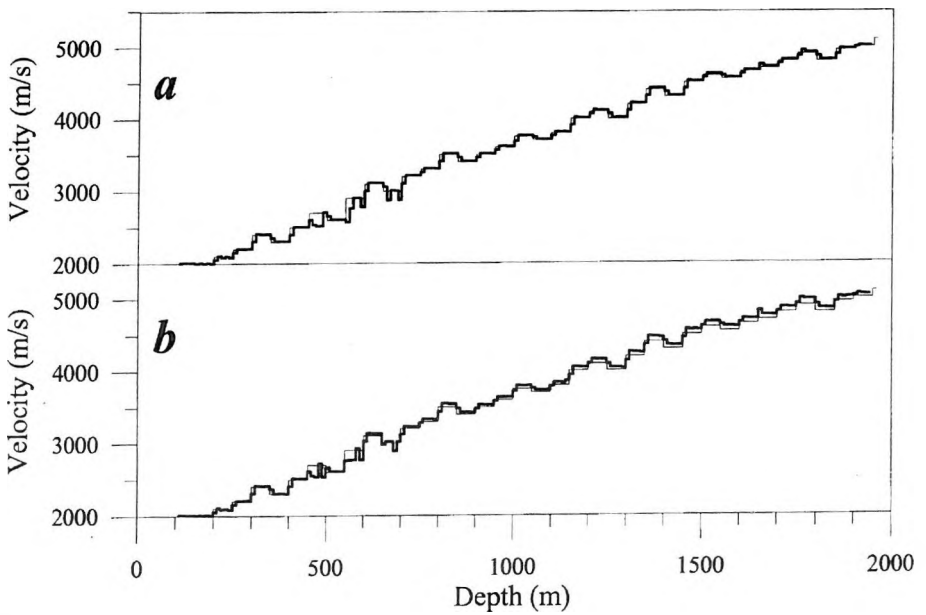


Fig. 5. True velocities (thin line) and estimated velocities (thick line) in case of *(a)* 10-degree and *(b)* 20-degree dipping layers

5. ábra. A valódi (vékony vonal) és a becsült (vastag vonal) sebességek (*a*) 10 fokos és (*b*) 20 fokos dőlt rétegek esetén

errors, several noise-free synthetic OVSPs were calculated and then inverted with different source offsets and different layer dips.

The inversion results show that in the case of horizontal layering the estimated velocities are very good approximations to the true ones, independently of the source offset: e.g. for a source offset of 300 m the mean difference between the estimated and the true velocity values is 0.5 per cent only (*Fig. 6a*).

Inversion results for 10-degree dipping layers and an updip source with offset of 300 m are shown in *Fig. 6b*. For a downdip source with offset of 300 m, the results are shown in *Fig. 6c*. The systematic errors between the estimated and the true velocities are due to the nonzero dips of the layers (model errors).

It is also clear from the inversion results that, for a given dip, the estimation error increases with the source offset and if the source is updip, the estimated velocities are greater than the true ones, and if the source is downdip, the estimated velocities are below the true velocities. However, even the greatest systematic errors do not exceed a few per cent.

It should be noted that if the true velocity function consists of relatively thick layers, the estimated velocity function follows the true velocity changes (velocity 'jumps') by insertion of one or two additional steps, even if the estimation is very good. This is due to the fact that the 3×3 point Sobel edge detection operator, during the calculation of the necessary differences, determines the weighted average of the adjacent data [ROSENFELD and KAK 1982]: whereas the traveltime curve of the first arrivals is broken, i. e. where the velocity changes rapidly, the edge detection operator rounds off the traveltime curve. Because of this rounding off, the gradient data do not change rapidly, so the estimated velocity function cannot change rapidly, either.

5. Application to field data

In this section the proposed inversion method is illustrated with a zero-offset VSP and a finite offset VSP measured in different boreholes in Hungary. The source offset for the OVSP is 735 metres. The separated downgoing and upgoing wavefields, the acoustic logs, and the velocity function estimated from the first break times of the zero-offset VSP have been made available for test purposes.

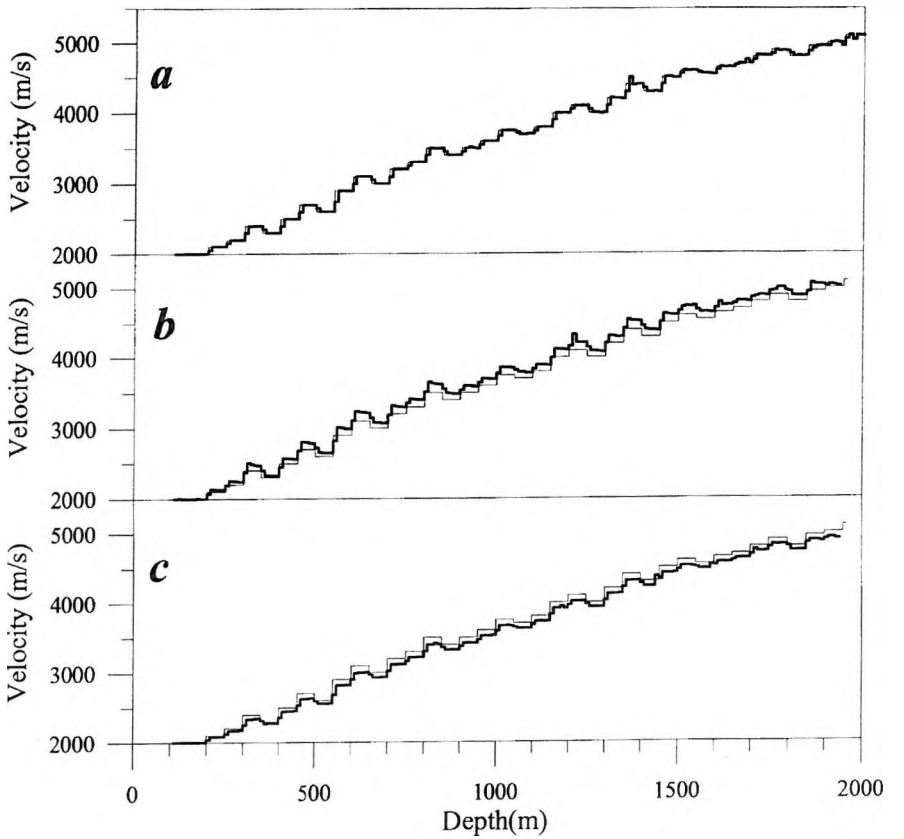


Fig. 6. True velocities (thin line) and estimated velocities (thick line) in case of 300 m source offset: *(a)* horizontal layers, *(b)* 10-degree dipping layers with source updip and *(c)* 10-degree dipping layers with source downdip

6. ábra. A valódi (vékony vonal) és a becsült (vastag vonal) sebességek 300 m-es offszet esetén: (*a*) vízszintes réteghatárok, (*b*) 10 fokos dőlt rétegek, forrás emelkedés irányban, (*c*) 10 fokos dőlt rétegek, forrás lejtés irányban

The downgoing wavefields separated from the two measured VSPs are illustrated in *Fig. 7.* Time gates applied to the OVSP and zero-offset VSP inversions were 100 ms and 150 ms wide, respectively, around the first breaks.

The results of the zero-offset inversion are illustrated in *Fig. 8.* The velocity function calculated from the acoustic log (CVL) and that estimated from the first break times are shown in *Fig. 8a.* The estimated velocity values deviate considerably from the CVL at many depths due to errors in estimates

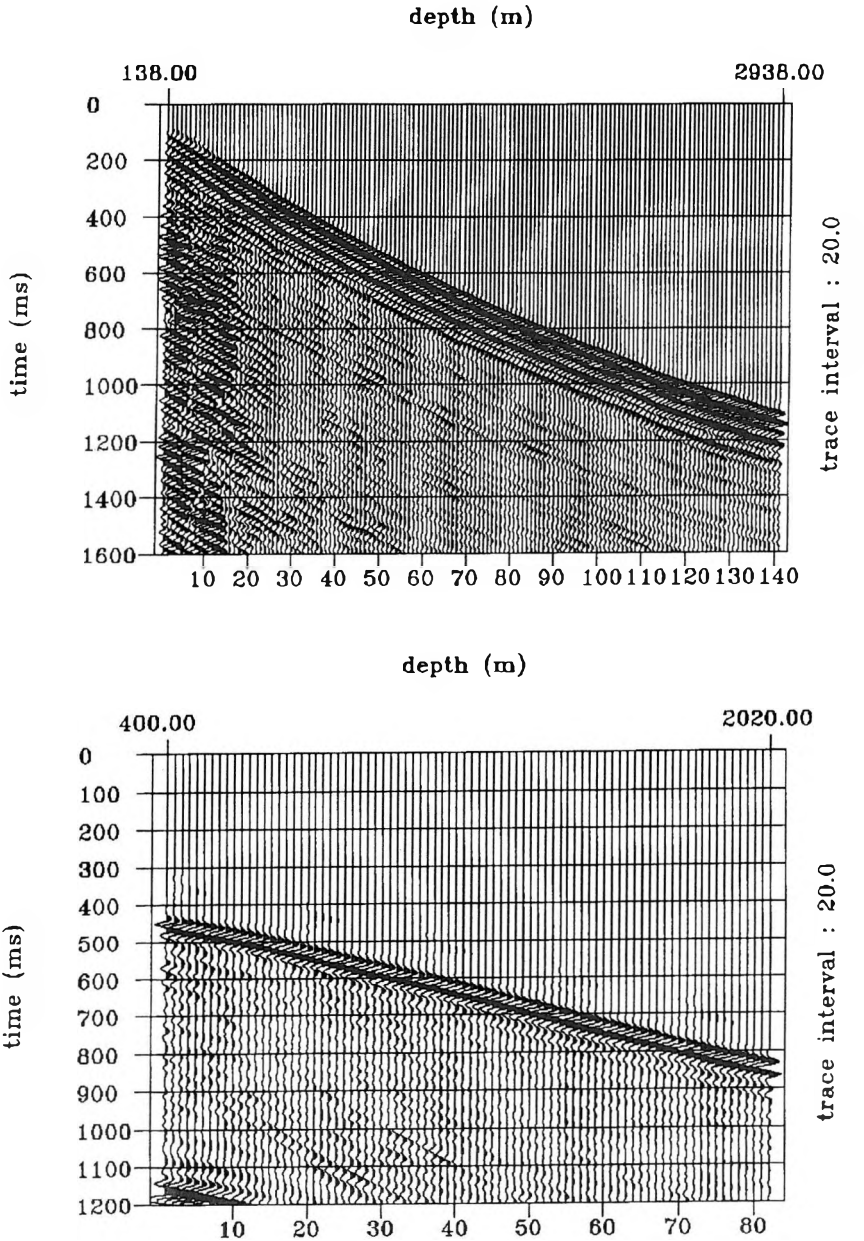


Fig. 7. above: Downgoing wavefield separated from a measured VSP. *below*: Downgoing wavefield separated from a measured OVSP. Source offset 735 m

7. ábra. felül: Egy mért zérus offszetű VSP szelvény lefelé haladó hullámtere. *alul*: Egy mért OVSP szelvény lefelé haladó hullámtere. Offszet: 735 m

of the first break times. The mean difference between the two velocity functions is 11.7 per cent, the largest difference is 104.1 per cent.

In *Fig. 8b* the velocities estimated by the proposed method are compared to the CVL: the mean difference between the two sets of velocities is 6.7 per cent, the largest difference is 33 per cent. In most depth intervals the estimation uncertainty does not exceed the value of ± 10 per cent of the estimated velocities. These results prove that the proposed inversion method gives more accurate results than the velocity estimation based on the first break times.

On the basis of the figure it can be observed that the low frequency part of the velocities obtained by the proposed inversion and of that estimated from the first break times are very similar. However, in the depth intervals of 2300-2400 m and 2700-2800 m the results of the inversion deviate considerably from the measured velocity values. These deviations cannot be explained by the uncertainties of the estimates (*Fig. 8c*). In other words, the two velocity functions estimated by two different methods from the same VSP are similar, but - at least in some depth intervals - they are not in accordance with the acoustic log.

Fig. 9 compares the CVL and the velocities calculated from the OVSP data. The mean difference between the two velocity functions is 6.5 per cent, the largest difference is 20.3 per cent. Comparison of *Figs. 8d* and *9c* shows that the quality of the OVSP is lower than that of the zero-offset VSP, resulting in a larger estimation uncertainty (*Fig. 9b*).

The fairly good results illustrated above are partly due to the fact that the geological layers around the boreholes are nearly horizontal, i.e. model errors are small.

6. Conclusions

The inversion technique presented in this paper uses a new method to estimate velocities from zero-offset and finite offset VSPs. After wavelet filtering and wavefield separation, the instantaneous phase section is calculated from the downgoing wavefield. The gradient of the traveltime curve of the first arrivals as a function of depth is estimated from the instantaneous phase section by means of edge detection. In order to obtain the interval velocity from OVSP data, the gradient is multiplied by a correction factor determined by ray trace modelling. The procedure does not need any a priori

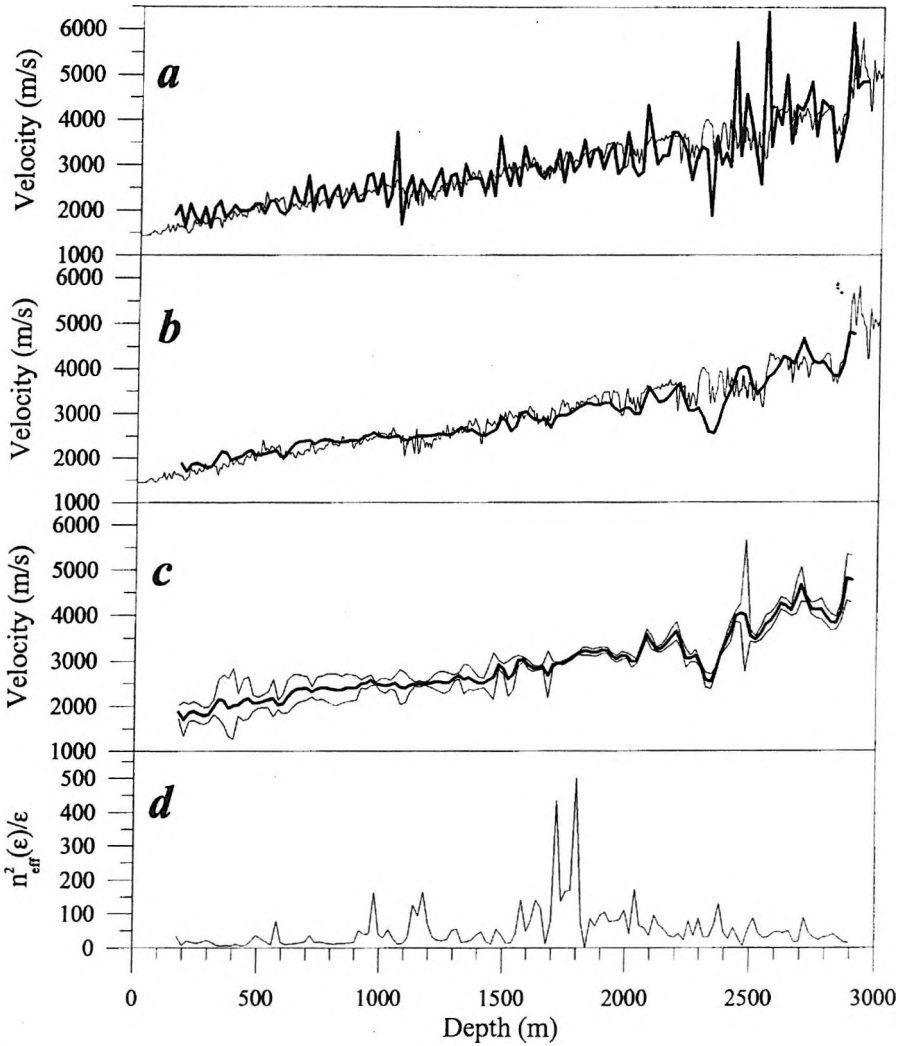


Fig. 8. (a) Velocity log (thin line) and velocities estimated from first break times of the zero-offset VSP (thick line). *(b)* Velocity log (thin line) and velocities estimated by the proposed method (thick line). *(c)* Upper and lower limits (thin lines) of estimated velocities (thick line) illustrating estimation uncertainties. The true velocities are between the two limits with a probability of about 66 per cent. *(d)* $n^2_{eff}(\varepsilon)/\varepsilon$ plot illustrating the quality of the processed VSP data

8. ábra.(a) A sebesség log (vékony vonal) és a zérus offszetű VSP első beérkezési időiből becsült sebességek (vastag vonal). *(b)* A sebesség log (vékony vonal) és a dolgozatban ismertetett eljárással becsült sebességek (vastag vonal). *(c)* A becsült sebességek (vastag vonal) bizonytalanságát illusztráló alsó és felső hibahatárok (vékony vonalak). A valódi sebességek kb. 66%-os valószínűséggel esnek a két hibahatár közé. *(d)* Az $n^2_{eff}(\varepsilon)/\varepsilon$ görbe, amely a feldolgozott VSP adatok minőségét jellemzi

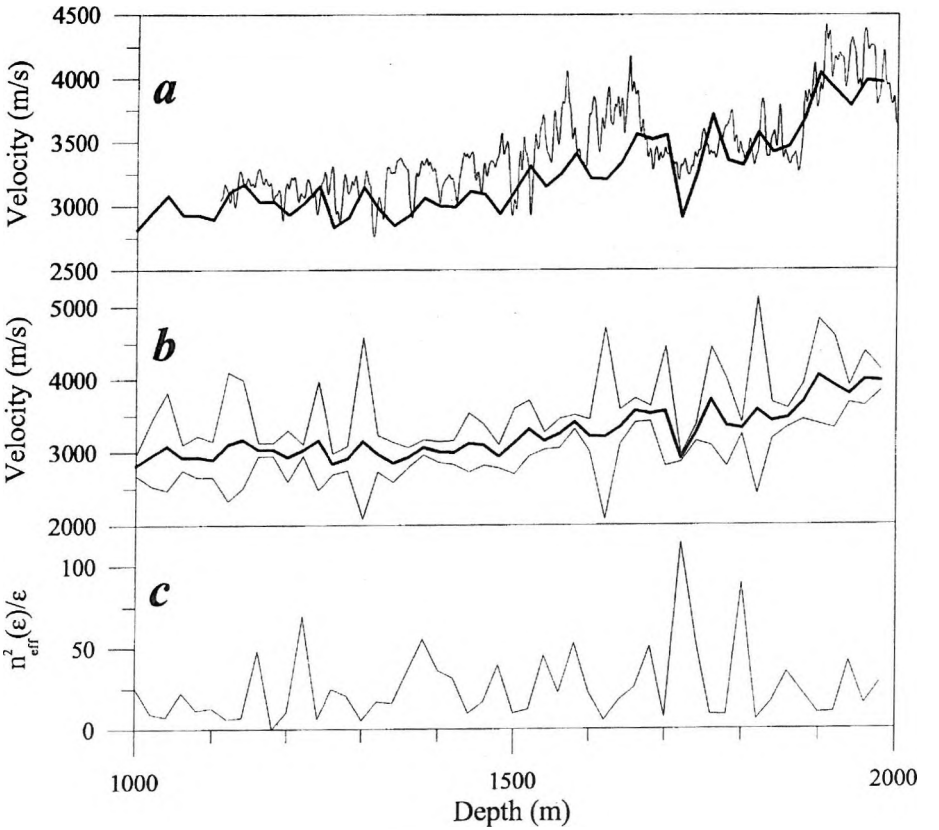


Fig. 9.(a) Velocity log (thin line) and velocities estimated by the proposed method from the finite-offset VSP (thick line). *(b)* Upper and lower limits (thin lines) of estimated velocities (thick line) illustrating estimation uncertainties. The true velocities are between the two limits with a probability of about 66 per cent. *(c)* $n^2_{eff}(\varepsilon)/\varepsilon$ plot illustrating the quality of the processed VSP data

9. ábra.(*a*) A sebesség log (vékony vonal) és a dolgozatban ismertetett eljárással becsült sebességek (vastag vonal). (*b*) A becsült sebességek (vastag vonal) bizonytalanságát illusztráló alsó és felső hibahatárok (vékony vonalak). A valódi sebességek kb. 66%-os valószínűséggel esnek a két hibahatár közé. (*c*) Az $n^2_{eff}(\varepsilon)/\varepsilon$ görbe, amely a feldolgozott VSP adatok minőségét jellemzi

information. True amplitude processing and initial velocity model are also not required. In addition to the estimated velocity values the estimation uncertainty can also be given. In our implementation the robust and resistant most frequent value procedure was used as a statistical tool. Because of the

simple algorithm the inversion procedure is very fast: it can be carried out economically even on personal computers.

Tests of this new technique on synthetic VSPs indicate that the method is very efficient and relatively insensitive to random noise and model errors. Inversions of real zero-offset VSP data prove that the resulting velocities are more accurate than those estimated from the first break times. In the case of OVSPs there is no other method that estimates velocities so accurately and rapidly. However, in complicated geological structures (faults, large layer dips, etc.) model errors can be significant and more testing should be done with field data in areas with good well control.

Appendix

On tracing direct rays with specified end points in layers of constant velocity and horizontal interfaces

When estimating the interval velocities calculation of the raypaths of the first arrivals should be carried out several times. Here, the tracing of direct rays with specified end points is discussed, assuming a horizontally layered earth model, where the horizontal layer boundaries coincide with consecutive geophone positions.

The traveltime t of a direct wave is given by

$$t = \sum_{i=0}^{N} \frac{1}{v_i} \sqrt{(x_{i+1} - x_i)^2 + d_i^2} \ ,$$

where x_i is the x-coordinate of the intersection point of the i-th interface and the raypath d_i denotes the thickness of the i-th layer and v_i is the wave velocity in the i-th layer. Since the summation starts with $i=0$, the raypath is defined - disregarding the end points - by N breakpoints. If the borehole is at the zero horizontal coordinate, then x_0 is equal to the offset and x_{N+1} is zero.

Fermat's principle states that traveltime t is stationary for small variations of the raypath. Accordingly, for the actual raypath the following system of nonlinear equations holds:

$$\frac{\partial t}{\partial x_i} = 0\,, \quad i.e.$$

$$\frac{1}{v_{i-1}} \frac{x_i - x_{i-1}}{\sqrt{\left(x_i - x_{i-1}\right)^2 + d_{i-1}^{\,2}}} + \frac{1}{v_i} \frac{x_{i+1} - x_i}{\sqrt{\left(x_{i+1} - x_i\right)^2 + d_i^2}} = 0 \tag{A-1}$$

$$\left(i = 1,2,\ldots,N\right)$$

These nonlinear equations can be solved by the Newton-Raphson iteration procedure [e.g. YAKOWITZ and SZIDAROVSZKY 1986.] If $f_i(x_1, x_2,\ldots, x_N)$ denotes the left hand side of Eq. (A-1), then the Jacobian matrix of the above system of non-linear equations is

$$J_{ij}\left(x_1, x_2, \ldots x_N\right) = \frac{\partial f_i\left(x_1, x_2, \ldots x_N\right)}{\partial x_j} \tag{A-2}$$

Taking into consideration that according to Eq. (A-1) $f_i(x_1\ ,x_2\ ,..,x_N\ )$ actually depends on x_{i-1} , x_i and x_{i+1} only, the Jacobian matrix of (A-1) is tridiagonal. This involves that the interval velocity estimation method described in this paper is a very fast and efficient procedure and can be carried out economically on personal computers.

REFERENCES

BALCH A. H., LEE M. W. 1984: Vertical Seismic Profiling: Techniques, Applications and Case Histories, IHRDC, 488 p.

GALPERIN E. I. 1974: Vertical seismic profiling: SEG Special Publication No. 12, Tulsa

GANLEY D. C. 1981: A method for calculating synthetic seismograms which include the effects of absorption and dispersion. Geophysics **46**, pp. 1100-1107

GRANT F. S., WEST G. F. 1965: Interpretation Theory in Applied Geophysics. McGraw Hill, New York, 583 p.

GRIVELET P. A. 1985: Inversion of vertical seismic profiles by iterative modeling. Geophysics **50**, pp. 924-930

HARDAGE B. A. 1985: Vertical Seismic Profiling. Part A: Principles. Handbook of Geophysical Exploration, Section I, Vol. 14A, Geophysical Press, London - Amsterdam

LASH C.C. 1980: Shear waves, multiple reflections and converted waves found by deep vertical wave test (vertical seismic profile). Geophysics **45**, pp. 1373-1411

LINES L.R., BOURGEOIS A., COVEY J. D. 1984: Traveltime inversion of offset vertical seismic profiles. A feasibility study. Geophysics **49**, pp. 250-264

MACE D., LAILLY P. 1986: Solution of the VSP one-dimensional inverse problem. Geophysical Prospecting **34**, pp. 1002-1021

PUJOL J., BURRIDGE R., SMITHSON S. B. 1985: Velocity determination from offset vertical seismic profiling data. J. Geophys. Res. **90 (B2)**, pp. 1871-1880

ROSENFELD A., KAK A. C. 1982: Digital Picture Processing. Academic Press, New York, 457 p.

STEINER F. 1988: Most frequent value procedures. Geophysical Transactions **34**, pp. 139-260

STEWART R. R. 1984: VSP interval velocities from traveltime inversion. Geophysical Prospecting **32**, pp. 608-628

TANER M. T., KOEHLER F., SHERIFF R. E. 1979: Complex seismic trace analysis. Geophysics **44**, pp. 1041-1063

URSIN B. 1986: Complete inversion of zero-offset seismic data. Geophysical Prospecting **34**, pp. 1213-1218

YAKOWITZ S., SZIDAROVSZKY F. 1986: An Introduction to Numerical Computations, Macmillan Publishing Company, New York

Képfeldolgozási algoritmusok alkalmazása VSP szelvények inverziójában

WÉBER Zoltán

Jelen dolgozatban az intervallumsebesség vertikális szeizmikus szelvények (VSP) alapján történo meghatározásának egy újszerű megközelítését ismertetjük. A VSP szelvényt kétváltozós képnek tekintve és rajta a képfeldolgozásból jól ismert éldetektáló alforitmust alkalmazva nagy mennyiségú irányadathoz juthatunk. Az azonos mélységhez tartozó irányadatok statisztikus feldolgozásával mind a sebességértékek, mind azok hibái megbecsülhetok.

A javasolt eljárás alkalmazása elott csupán a rutin feldolgozásnak számító jelalak szűrés és hullámtér szétválasztás műveleteit kell végrehajtanunk. Kezdeti sebességmodell definiálására és a valódi amplitúdók visszaállítására nincs szükség. A módszer segítségével offszetes VSP (OVSP) szelvények inverziójára is lehetoség nyílik. Ennek érdekében az OVSP adatok feldolgozása során nyert irányadatokat egy korrekciós faktorral kell megszoroznunk, melynek értékét sugárkövetéses modellezéssel határozzuk meg.

A szintetikus és mért adatokon végzett vizsgálatok azt mutatják, hogy a javasolt módszer nagyon hatékony és viszonylag érzéketlen a véletlen zajokra és a modellhibákra. Az eljárás személyi számítógépen is gazdaságosan megvalósítható.

GEOPHYSICAL TRANSACTIONS 1997
Vol 40. No. 3-4. pp. 175-189

Vertical steel casing as a monopole transmitter antenna for electromagnetic prospecting

Ernő TAKÁCS*

A vertical steel casing excited as a monopole antenna could potentially be used as a source for electromagnetic investigation of the surroundings of a borehole. In this respect the behaviour of the current distribution along the casing and the surface electric field due to it is treated. Special emphasis is laid on the frequency dependence, which seems to be of good use in interpretation. Numerical models developed to simulate the casing response — although considerably simplified — match the experimental data very well.

Keywords: frequency sounding, E-field excited casing, casing as antenna, surface electric field of a casing

1. Introduction

The steel casings of two drill-holes used as vertical input electrodes — line sources — produce a higher current density with increasing depth than point sources located at the ground surface [ROCROI, KOULIKOV 1985]. This procedure (called Tubel) was developed to delineate a partially known hydrocarbon deposit overlain by resistive layers. HAGREY [1994] presented a case history for investigating electric anisotropy in a granite basement by applying the mise-à-la-masse technique. One of the two fixed current electrodes was connected to a fracture via the metal casing of a borehole. Similar measurements were carried out in Hungary to determine the main directions

* University of Miskolc, Department of Geophysics, H-3515 Miskolc-Egyetemváros, Hungary

Manuscript received (revised version): 30 August, 1995

of large karstic water-filled cavities inside the limestone bedrock of the coal seam in Lencsehegy MiningWorks [TAKÁCS 1989].

Otherwise, in recent years several papers have been published in the geophysical literature in connection with borehole casing, discussing questions such as

— the effect of well casings on electrical surveys [e.g. JOHNSTON et al. 1992],
— communication capability of a casing or a drillstring for measurement-while-drilling (MWD) [e.g. XIA et al. 1993],
— electrical resistivity measurement through metal casing [e.g. SCHENKEL et al. 1994],
— DC resistivity imaging from a steel cased well [SCHENKEL 1994].

This paper investigates whether the electromagnetic field of the vertical steel casing of a single well excited by an alternating voltage applied to an isolation gap that divides the casing into two parts near the top or bottom end could be used for the exploration of the surroundings of the well. Frequency domain measurements offer an opportunity to govern the penetration depth and transform the measured voltages into apparent resistivities.

2. Surface radial electric field of an E-field excited casing in a homogeneous half-space

If a voltage is applied to an isolation gap near the top or bottom of an E-field excited casing a complex $I(z)$ current-distribution will arise along its length. The maximum value of $I(z)$ will be at the gap, and at the ends of the casing its value becomes zero. With increasing frequency $I(z)$ concentrates itself towards the gap. Taking the casing as an "antenna" the current distribution along its axis can be determined by procedures known from antenna theory [KING, SMITH 1981; XIA et al. 1993].

In practice, E-field excitation can be more or less accomplished if the top of the casing and surface electrodes around it, or an electrode located beneath the casing shoe are taken as points of current supply.

A computer program was developed to determine the current distribution by the method of moments if the casing has perfect electrical contact with the homogeneous half-space [VARGA 1994].

In *Fig. 1* current distributions along a casing can be seen. The casing is in a homogeneous half-space with specific conductivities of 0.04 and 0.07 S, and the frequencies are 3.5 and 3500 Hz. *Im I(z)* values at 3.5 Hz are negligibly small. They could not be plotted on this scale. The casing emits radial currents towards the formation in the vicinity that are proportional to the partial derivative of *I(z)* with respect to the vertical direction. Elemental currents along the casing can be considered as elementary vertical electric dipoles with real and imaginary currents. Their superposed field can be measured at the earth's surface. The contribution of an elementary section of the casing to the surface field depends on both the current and the relative positions of the elementary sections with respect to the observation point.

Amplitude and phase frequency sounding curves are presented in *Fig. 2* for the half-spaces considered above. The frequency dependence of the surface radial electric field is similar to the field of a vertical electric dipole placed in the middle part of the casing.

3. Geometrical sounding over a horizontally layered half-space

It is well known that the electromagnetic field generated by a vertical electric dipole is weaker by a factor of at least 100 for a given separation and frequency, and that this attenuates more rapidly with the separation than those fields due to a horizontal dipole with the same source strength [GOLDMAN 1990]. At the same time, the rock volume between the source and observation point has a larger effect on the field than the section below the source [TAKÁCS 1988].

Owing to the similiarity between the electromagnetic fields of the casing and a buried vertical electric dipole the rules mentioned above will probably apply to our case. It is therefore advisable to examine the attenuation of the surface electric field and the effect of the layering below the casing.

In *Figs. 3* and *4* surface amplitude and phase values of the radial electric field are shown along a profile above a homogeneous and a layered half-space for 5 Hz and for the top and bottom excitation respectively. In the layered section a resistive layer —1000 Ωm — is embedded into the homogeneous half-space below the casing.

The surface field is rather sensitive to the presence of the 1000 Ωm layer — which lies close to the termination of the casing — already at relatively small offset. At $R < 3.5h_1$ — h_1 is the thickness of the first layer — the field

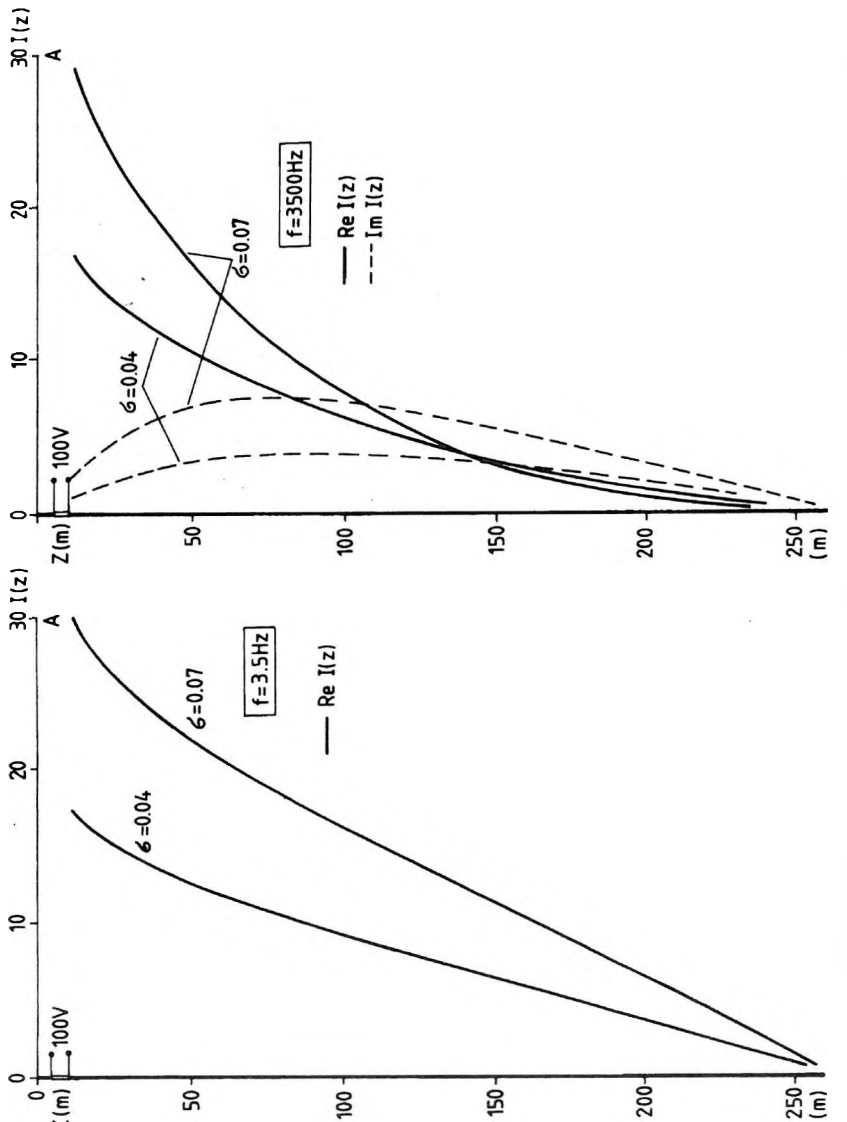


Fig. 1. Real and imaginary parts of current distributions along a vertical casing in homogeneous half-space
1. ábra. A valós és képzetes áramok eloszlása a homogén féltérben levő függőleges béléscső mentén

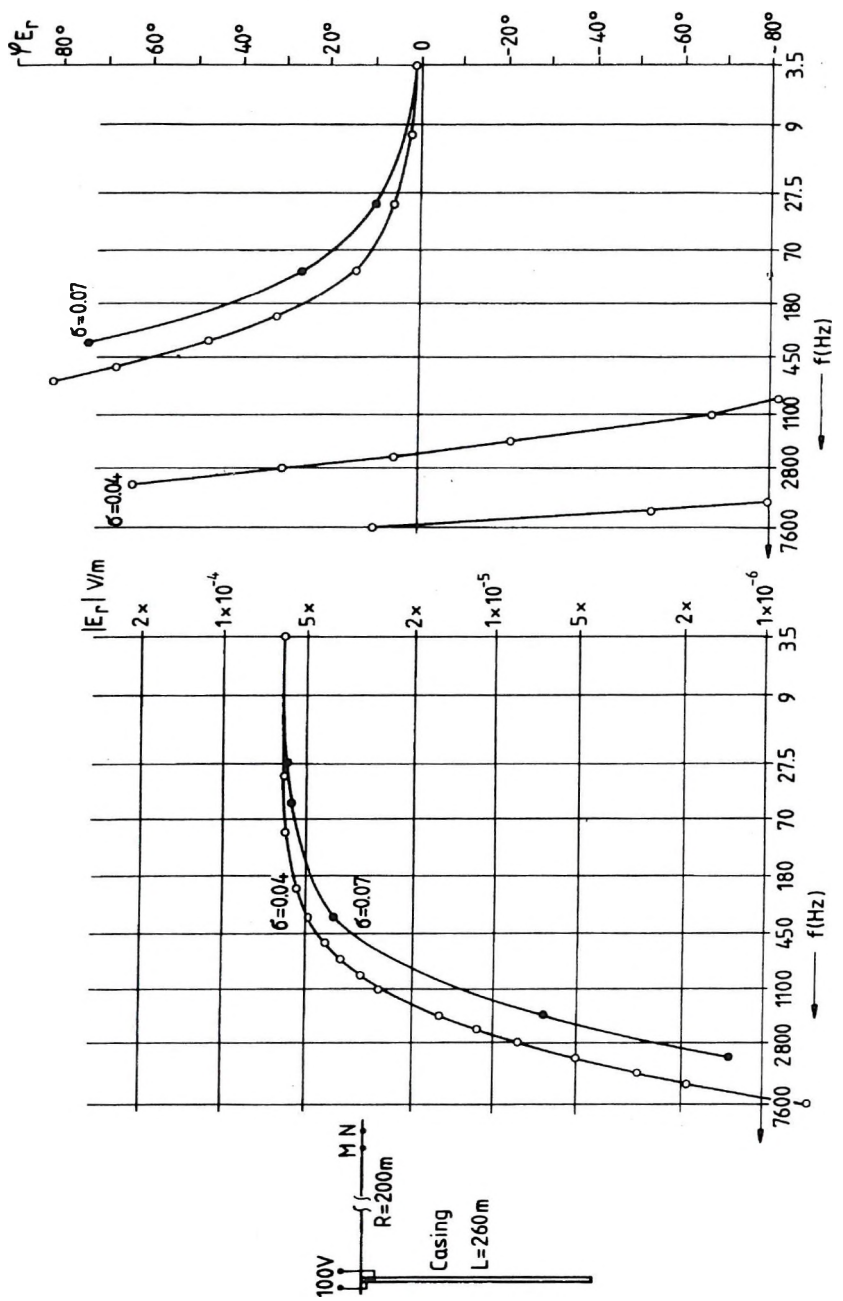


Fig. 2. Amplitude and phase frequency sounding curves of the radial electric field on the surface of homogeneous half-space
2. *ábra.* A radiális elektromos térerősség amplitúdó és fázis frekvenciaszondázási görbéi homogén féltér felszínén

is lower than for the homogeneous case, at $R > 3.5h_1$, on the other hand, the values are higher for the layered model. Even at this low frequency — 5 Hz — the phase grows dynamically with offset and the effect of the 1000 Ωm layer is an unambiguous decrease at $R > h_1$.

The difference is more pronounced in the case of the bottom excitations, and the amplitudes are larger owing to the geometry and to the fact that from the bottom end of the casing more current dissipates at a given voltage than it does at top excitation.

The effect of the resistive layer is more clearly indicated by the ρ_a apparent resistivity curve calculated from the ratio of the amplitude in the layered earth to the amplitude in a half-space having the resistivity of the uppermost layer. The ρ_a apparent resistivity curves have a false minimum. At the same time the practical application of this apparent resistivity definition is not so straightforward as for point electrodes.

Another apparent resistivity determination — ρ_a^* — can be made based on the homogeneous half-space electric number phase characteristic, i.e. phase curve versus offset normalized to the skin-depth [TAKÁCS 1988]. ρ_a^* values are also given in Figs. 3 and 4. It can be seen, that they have a direct relation to the section without any "undershoot".

Geometrical sounding data for large offset are needed to obtain reliable information about the section. It should not be forgotten, however, that the low signal sets limits to the separation.

Owing to the fact that the effect of the lower 1000 Ωm layer appears at rather small separation it is expected that frequency domain soundings could have advantages in the exploration of the immediate vicinity of the well.

4. Frequency sounding over a layered half-space

Amplitude and phase frequency sounding curves are presented in *Figs. 5* and *6* for the homogeneous and layered half-spaces in the case of a top and a bottom excitation. The separations are 250 and 500 m, both in the range where amplitudes respond with a decrease due to the presence of the resistive layer. The effect of the 1000 Ωm layer for both excitations increases with the offset. However, it is more pronounced in the case of bottom excitation.

So far as the amplitudes are concerned the difference between the curves for layered and homogeneous half-space is the gratest at the lowest frequency as a result of the larger penetration depth. On the other hand, the phase

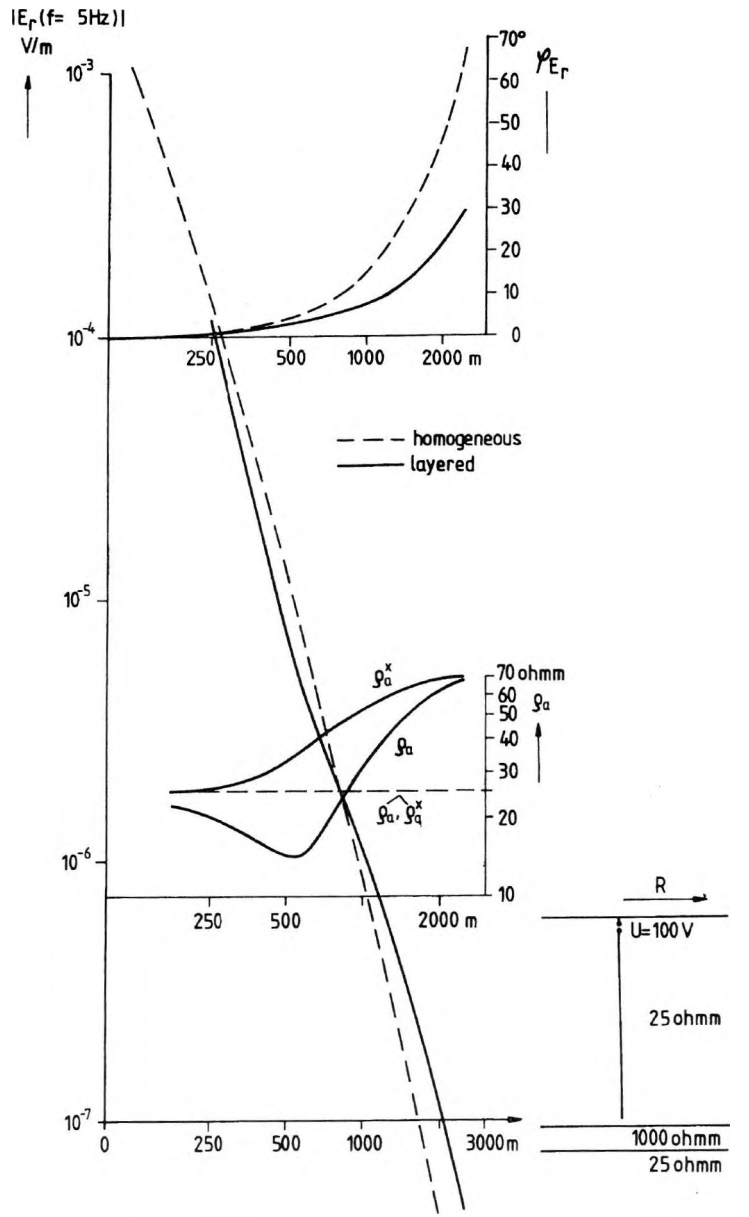


Fig. 3. Amplitudes, phases of the 5 Hz radial electric field and apparent resistivities (ρ_a, ρ_a^*) on the surface of a homogeneous — 25 Ωm — and a layered — 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm — half-space (top excitation)

3. ábra. Az 5 Hz-es radiális elektromos térerősség amplitúdó-, fázis-, és látszólagos fajlagos ellenállás (ρ_a, ρ_a^*) szelvényei a 25 Ωm-es homogén és a 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm paraméterű rétegzett féltér felszínén (felszíni árambevezetés)

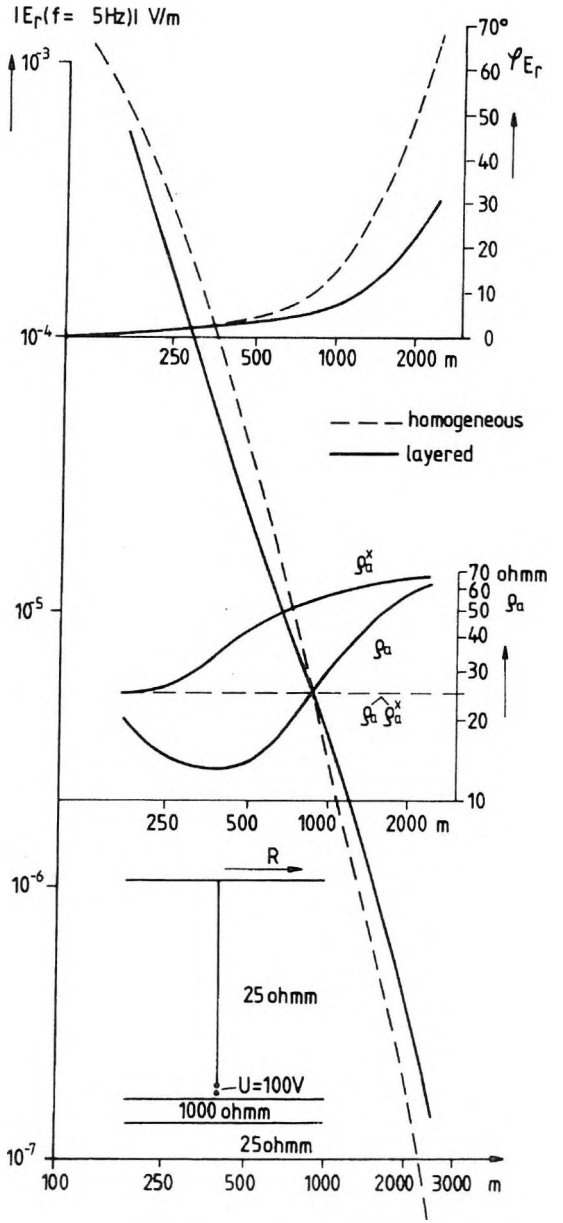


Fig. 4. Amplitudes, phases of the 5 Hz radial electric field and apparent resistivities (ρ_a, ρ_a^*) on the surface of a homogeneous — 25 Ωm — and a layered — 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm — half-space (bottom excitation)

4. ábra. Az 5 Hz-es radiális elektromos térerősség amplitúdó-, fázis- és látszólagos fajlagos ellenállás (ρ_a, ρ_a^*) szelvényei a 25 Ωm-es homogén és a 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm paraméterű rétegzett féltér felszínén (alsó árambevezetés)

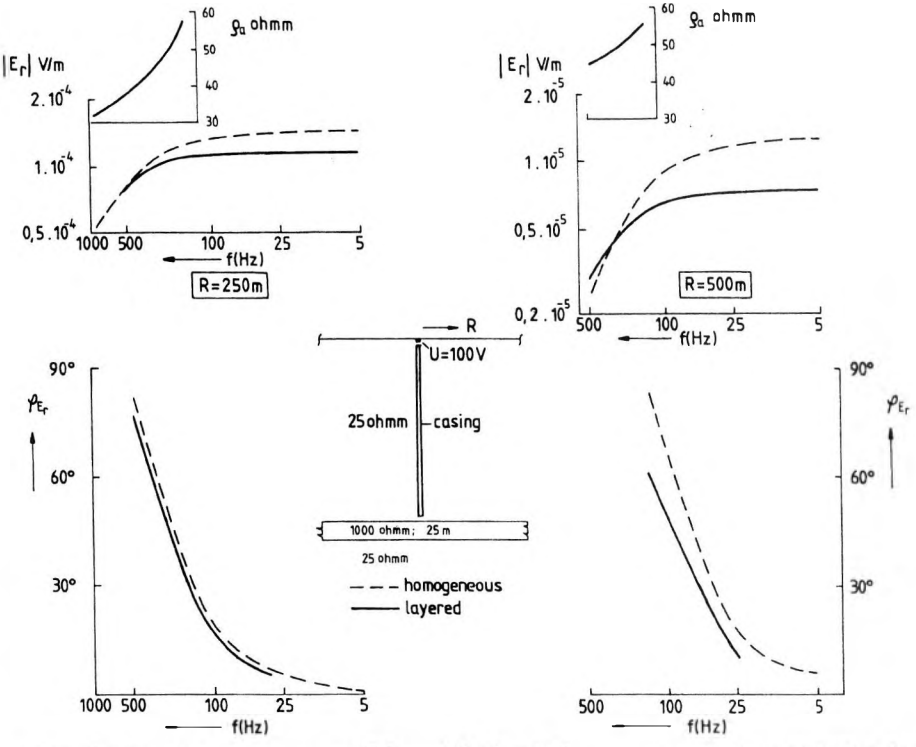


Fig. 5. Amplitude and phase frequency sounding curves of the radial electric field and ρ_a apparent resistivities derived from amplitudes on the surface of a homogeneous — 25 Ωm — and a layered — 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm — half-space at separations $R = 250$ m (*a*) and $R = 500$ m (*b*) (top excitation)

5. ábra. A radiális elektromos térerősség amplitúdó és fázis frekvenciaszondázási görbéi és az amplitúdóból származtatott ρ_a látszólagos fajlagos ellenállások a 25 Ωm-es homogén és a 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm paraméterű rétegzett féltér felszínén $R = 250$ m (*a*) és $R = 500$ m (*b*) távolságban (felső árambevezetés)

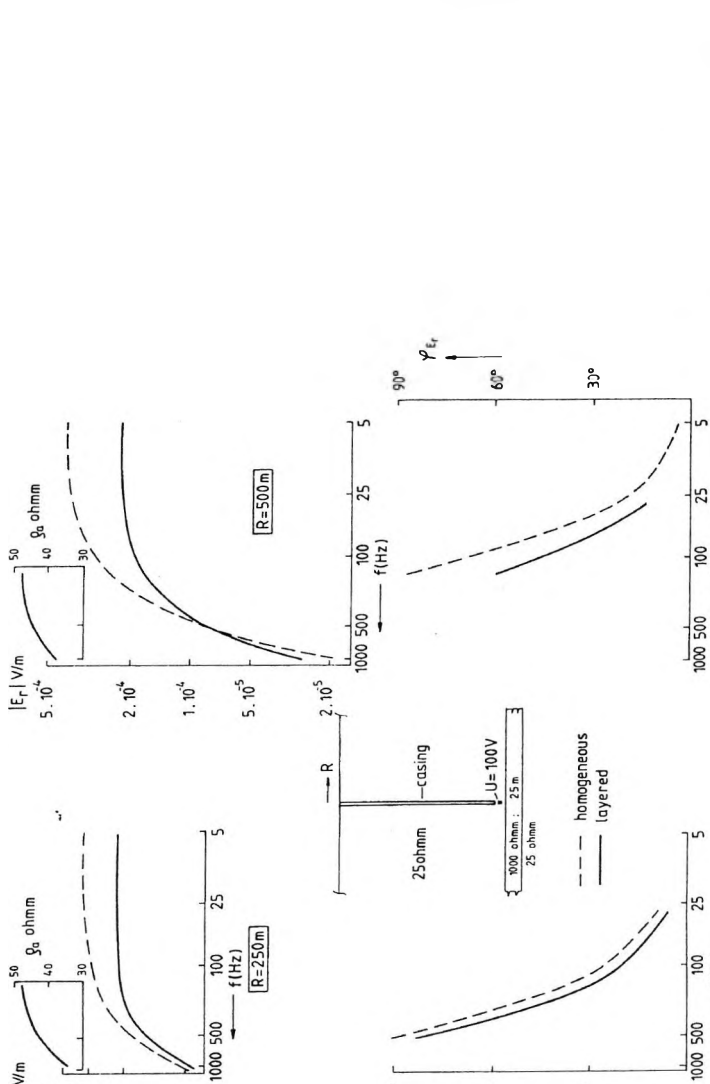


Fig. 6. Amplitude and phase frequency sounding curves of the radial electric field and ρ_a apparent resistivities derived from amplitudes on the surface of a homogeneous — 25 Ωm — and a layered — 1000 Ωm, 25 m; 25 Ωm — half space at separations R = 250 m (*a*) and R = 500 m (*b*) (bottom excitation)

6. ábra. A radiális elektromos térerősség amplitúdó és fázis frekvenciaszondázási görbéi és az amplitúdóból származtatott ρ_a látszólagos fajlagos ellenállások a 25 Ωm-es homogén és a 25 Ωm, 250 m; 1000 Ωm, 25 m; 25 Ωm paraméterű rétegzett féltér felszínén R = 250 m (*a*) és R = 500 m (*b*) távolságban (alsó árambevezetés)

difference has an opposite tendency which is connected mainly with the changes of the lateral extension of the current systems of changing sign around the casing. Otherwise, the phase values are represented only for the inner current system around the casing. After the first sign-change they rapidly change.

The monotonic character of the frequency sounding curves conceals the information about the layering. However, it will clearly appear in the apparent resistivity curves calculated by the use of the characteristics of the electric number over a homogeneous half-space [TAKÁCS 1988]. Such characteristics can be seen in *Fig. 7* for top excitation, where phases and normalized amplitudes to the DC field are represented as a function of the separation normalized to the skin-depth (δ). Similar characteristics can be given for the bottom excitation.

For given separation and frequency, knowing the measured phase — or the amplitudes at the frequency considered and at DC — ρ_a values can be calculated.

ρ_a values determined for the descending branches of the amplitude curves — where the rate of change is well defined — clearly show the presence of the 1000 Ωm layer near the bottom of the casing. The apparent resistivity decreases significantly with increasing frequency from 150 Hz to 1000 Hz, this means that it does not remain constant as it does in the case of a homogeneous half-space.

5. Field experiment

Numerical models developed above to simulate the casing response match the experimental data very well.

There is a cased borehole at the Miskolc University campus, which has offered us a possibility to check the theoretical considerations. The length of the steel casing is 261 m. Immediately below the casing shoe — from 267 m downwards — highly resistive Mesozoic limestone basement can be found. Overlying formations containing the casing are mostly argillaceous sediments of low resistivity. Apparent resistivities measured in the borehole exceed 20 Ωm only in the depth interval 70–132 m.

Phase curve and apparent resistivities deduced from it measured at a separation $R = 250$ m are shown in *Fig. 8*.

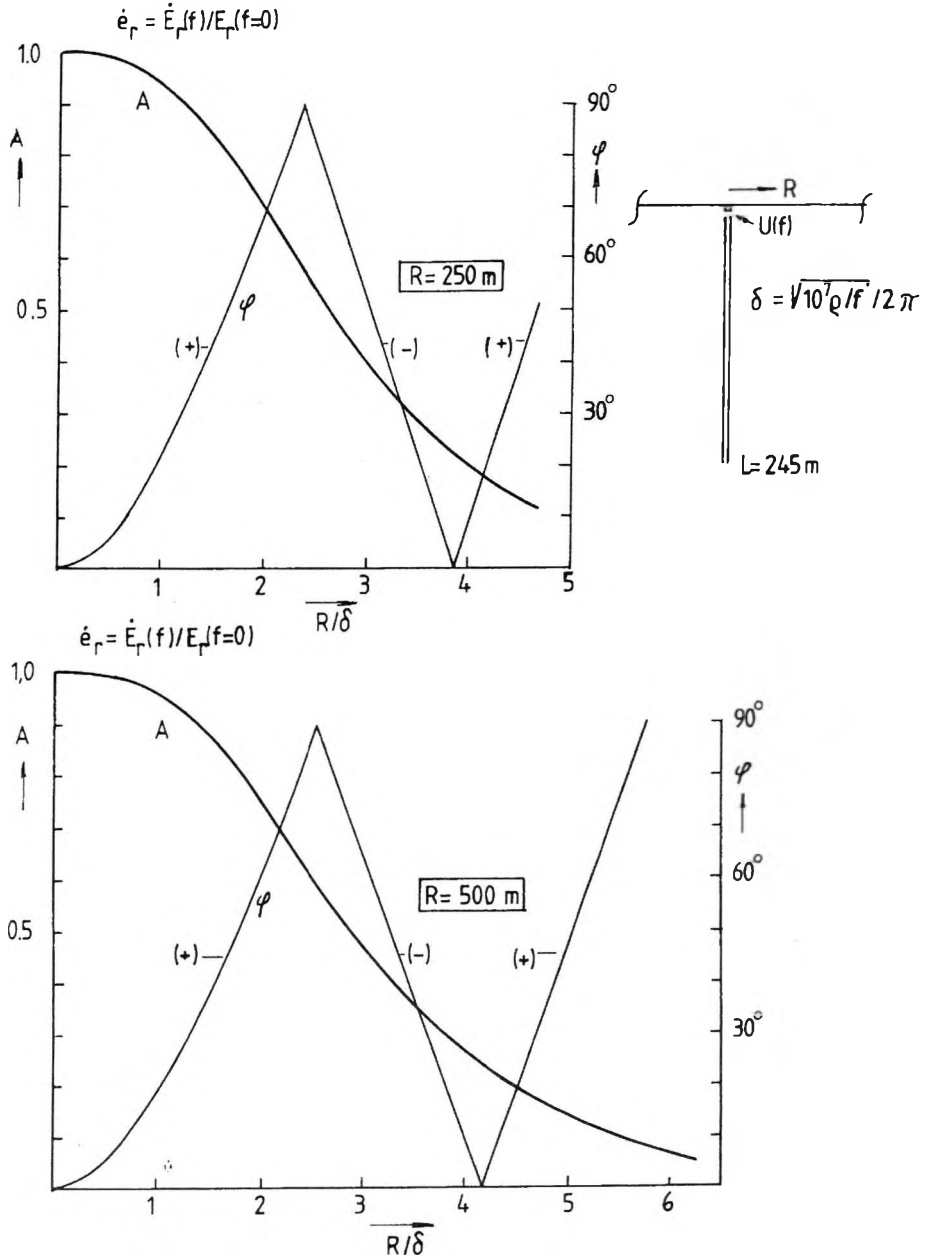


Fig. 7. Amplitudes and phases of the radial electric number ($\dot{e}_r$) as a function of the induction number R/δ (top excitation)

7. ábra. A radiális elektromos szám ($\dot{e}_r$) amplitúdó és fázisgörbéje az R/δ indukciós szám függvényében (felső árambevezetés)

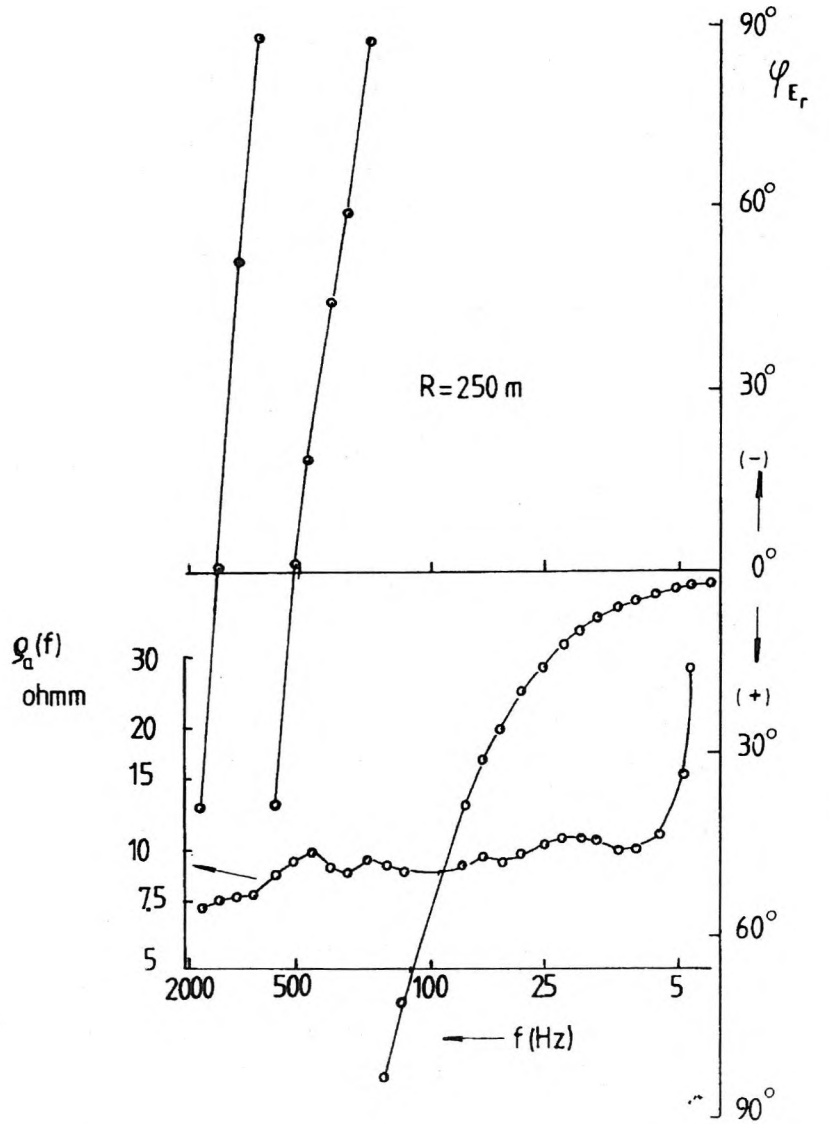


Fig. 8. Results of a field test obtained by the top excitation of a casing 261 m in length. Frequency dependence of the phase and apparent resistivities (ρ_a) derived from it are shown. 50 V voltage was applied between the casing and short-circuited electrodes around a circle with 10 m radius

8. ábra. Egy 261 m hosszúságú béléscső felső megtáplálásával kapott kísérleti mérés eredményei. A fázis frekvenciafüggése és a belőle levezetett ρ_a látszólagos fajlagos ellenállások láthatók. Az 50 V-os feszültséget a béléscső teteje és a körülötte 10 m sugarú körön levő rövidre zárt elektródákhoz csatlakoztatták

The author's opinion is that the aim of such measurements is not so much the quantitative interpretation for layering, but the location of the possible lateral inhomogeneities in the neighbourhood of the borehole. For that purpose several measurements have to be made around the borehole on the circumference of a series of concentric circles.

In our case the signs of human activity prevented us from making more measurements. Therefore our single test measurement proves only that the $\rho_a(f)$ curve has such characteristic features, that can be used for correlation between stations with the same separation around the borehole to detect lateral inhomogeneities.

6. Conclusions

This study has shown that the steel casing of a borehole may be used as a transmitter antenna. The *E*-excitation mode — applying a voltage to an isolation gap that divides the casing into two parts — was selected. Based on numerical simulations and a test measurement it is to be expected that the changes of the surface radial electric field around the circumference of a circle whose centre is the borehole will reflect inhomogeneities in the neighbourhood of the borehole. Additional work needs to be done since in reality there are other factors — e.g. contact impedance, dip — that could bias the results.

Acknowledgement

This work was supported by the Hungarian Scientific Research Fund (OTKA) — contract number 2383 and CO216.

REFERENCES

GOLDMAN M. M. 1990: Non-conventional methods in geoelectrical prospecting. Ellis Horwood

HAGREY S. A. 1994: Electric study of fracture anisotropy at Falkenberg, Germany. Geophysics **59**, 6, pp. 881–888

JOHNSTON R. H., TROFIMENKOFF F. N., HASLETT I. W. 1992: The complex resistivity response of a homogeneous earth with a finite-length contained vertical conductor. IEEE Transactions on Geoscience and Remote Sensing **30**, 1, pp. 46–54

KING R. W. P., SMITH G. S. 1981: Antennas in matter. The MIT Press

ROCROI I. P., KOULIKOV A. V. 1985: The use of vertical line sources in electrical prospecting for hydrocarbon. Geophysical Prospecting **33**, pp. 138–152

SCHENKEL C. I., MORRISON H. F. 1994: Electrical resistivity measurement through metal casing. Geophysics **59**, 7, pp. 1072–1082

SCHENKEL C. I. 1994: DC resistivity imaging using a steel cased well. 64th Annual International SEG Meeting Expanded Abstracts pp. 403–406

TAKÁCS E. 1988: In-mine frequency sounding with a buried grounded vertical dipole source. Geophysical Transactions **34**, 4, pp. 343–359

TAKÁCS E. 1989: Mapping voids in the triassic limestone basement — recovered by boreholes — with application of a casing-surface geoelectric circular profiling. Report, Dorog Coal Mines

VARGA M. 1995: Program 'Casing' for the determination of current distribution along an E-field excited casing. OTKA-report, contract no 2383

XIA M. Y., CHEN Z. Y. 1993: Attenuation predictions at extremely low frequencies for measurement — while — drilling electromagnetic Telemetry System. IEEE Transactions on Geoscience and remote sensing **31**, 6, pp. 1222–1228

A béléscsőnek, mint adó- és vevőantennának használata az elektromágneses kutatásban

TAKÁCS Ernő

A monopol-antennaként megtáplált acél béléscső elvileg az elektromágneses tér forrása lehet a fúrás közvetlen környezetének kutatásához. A tanulmány ilyen célzattal vizsgálja a béléscsőmenti árameloszlást és az általa létrehozott felszíni elektromos térerősséget. Különös figyelmet szentel a frekvenciafüggésnek, ami az értelmezésben jól használható. A béléscső elektromágneses terét leíró modell — bár sok közelítést tartalmaz — a kísérleti méréssel jól egyező eredményt ad.

GEOPHYSICAL TRANSACTIONS 1997
Vol. 40. No. 3–4. pp. 191–216

MFV-corrected variances

Ferenc STEINER*, Béla HAJAGOS*, Gábor HURSÁN*

The first part of the paper shows and quantitatively characterizes the probability distortions of the 'corrected empirical variance' (denoted by $\sigma^2_{emp,corr}$) which can be considerable for some, in practice often occurring, parent distribution types and sample sizes, respectively. Even for the pure Gaussian distribution (as the parent one) such probability distortions of the conventionally corrected empirical variance are present.

The second part of the paper deals with the proper correction method called 'MFV-correction' or 'M_k-correction' (M_k means the most frequent value, the actual values of k being 1/2; 1; 2 and 3). The correction constants $C(F,k;n)$ are tabulated for three types of parent distributions (see Table IIIa). Table IIIb gives the coefficient-triplets $A_{F,k}$, $B_{F,k}$ and $C_{F,k}$ to calculate with very good approximation $C(F,k;n)$ for an arbitrary value of $n < 1000$.

Eight tables (Table IVa to Table IVh) give detailed information about the most important probability characteristics of the 'MFV-corrected empirical variances' $\sigma^2_{M_k corr}$ as random variables. The values are convincing in demonstrating that if one single $\sigma^2_{M_k corr}$ is calculated (this is the ordinary case in practice), then with great probability this estimate is near to the true value of the variance.

In the overwhelming majority of the discussed cases $c(F,k;n) = C(F,k;n)(n\text{-}1)/n$ is significantly greater than unity, showing that though using the simple correction factor $n/(n\text{-}1)$, $E(\sigma^2_{emp,corr}) = \sigma^2$ really holds, this requirement does not, however, result in such a distribution of the estimates which is acceptable in practice; in asymmetrical cases the fulfilment of the equation '*E(estimate) = true value*' is rather misleading in respect of the occurrence probabilities of the bias. It seems better to speak about 'unbiased estimate' if the equation *MFV(estimate) = true value* holds.

Keywords: variance, estimates of the variance, corrected empirical variances, unbiased estimate, most frequent value

* University of Miskolc, H-3515 Miskolc-Egyetemváros

Manuscript received: 20 November, 1995

1. Introduction

In practice, based on the sample x_1, x_2, ..., x_i, ... ,x_n, generally the expression

$$\sigma^2_{emp} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \tag{1}$$

is used as an estimate of the true value of the variance (i.e., of the square of the scatter) defined by

$$VAR(\xi) \equiv \sigma^2 = \int_{-\infty}^{\infty}\left(x - E\right)^2 \cdot f(x); \tag{2}$$

where $\bar{x}$ is the arithmetic mean of the data x_i, ξ is the random variable from where the sample was taken, $f(x)$ is the probability density function of ξ, and E is its 'expected value':

$$E(\xi) = \int_{-\infty}^{\infty} x \cdot f(x)\mathrm{d}x\,. \tag{3}$$

$VAR(\xi)$ plays a fundamental role in classical geostatistics for calculating the variograms [see e.g. MATHERON 1965], and their square root, i.e., σ itself is widely used as an error-characteristic. The question arises as to whether or not Eq. (1) satisfactorily estimates the true value of variance σ^2 (and its square root the scatter σ)?

In every standard book of probability theory and mathematical statistics [see e.g. CRAMÉR 1945] it is proven that

$$E\left(\sigma^2_{emp}\right) \neq \sigma^2 \tag{4}$$

and therefore we speak about distortion. The correction is, however, very simple: in the same books it is proven that the 'corrected empirical variance' $\sigma^2_{emp,corr}$ defined by

$$\sigma^2_{emp,corr} = \frac{n}{n-1}\sigma^2_{emp} \tag{5}$$

fulfils the equation

$$E\left(\sigma^2_{emp,corr}\right) = \sigma^2 \tag{6}$$

and therefore $\sigma^2_{emp,corr}$ is usually known as an 'unbiased' estimate of σ^2, i.e., of the true value of the variance.

HAJAGOS and STEINER [1994], however, pointed out that even $\sigma^2_{emp,corr}$ shows distortions but of other type: these estimates are less (or even significantly less) with greater (eventually significantly greater) probability than 0.5 than the true value of σ^2. Such distortions may be called 'probability distortions' and are extremely important in the practice. A proposal for their quantitative characterization is given in STEINER [1991] (see page 247). More intensive investigations of these probability distortions are reported in KIS and HURSÁN [1995]. In the next point a fairly complete characterization of this type of distortion of the $\sigma^2_{emp,corr}$ estimates will be given for three parent distribution types: geostatistical, Jeffreys, and Gaussian (for the formulae of their probability density functions see in Appendix C/ of the present paper, where the distribution functions are also given). Comments are given on the choice of two parent distribution types: a/ DUTTER [1986/87] states that in geostatistic practice, geostatistical distribution generally serves as an appropriate model (this type is not only characteristic of geostatistics, see e.g. the practical: on one hand geodetical on the other hand astronomical examples given in HAJAGOS and STEINER [1994]); b/ after Jeffreys [cited by KERÉKFY 1978] shorter flanks than those of the Jeffreys-distribution never or very rarely occur in practice.

All investigations of the present paper were made with the Monte Carlo method characterized by the repetition number $N=10000$ for both of the above- mentioned parent distributions. Although the Gaussian case as a parent distribution is far from the realities of the geosciences, for comparison purposes this classical case is also treated in this paper. The Monte Carlo method is not needed for investigating the Gaussian case (only integrating or differentiating is neccessary) as the probability density function of $\sigma^2_{emp,corr}$ can be given in this case also analytically:

$$f\left(\sigma^2_{emp,corr}\right)=\frac{n-1}{\sigma^2}\,\frac{y^{\frac{n-1}{2}-1}\cdot e^{-\frac{y}{2}}}{2^{\frac{n-1}{2}}\cdot\Gamma\left(\frac{n-1}{2}\right)}, \tag{7}$$

where $y=\sigma^2_{emp,corr}$ [see e.g. CRAMÉR 1945].

2. Characterization of the probability distortions

Figure 1 shows the frequency function of $\sigma^2_{emp,corr}/\sigma^2$ for the sample size $n=9$ and for the geostatistical parent distribution. The already cited proposal of STEINER [1991] for measuring the probability distortion will be denoted by *pd*; the definition is very simple:

$$pd = \frac{P\left(\sigma^2_{emp,corr} < \sigma^2\right)}{P\left(\sigma^2_{emp,corr} > \sigma^2\right)} - 1 . \tag{8}$$

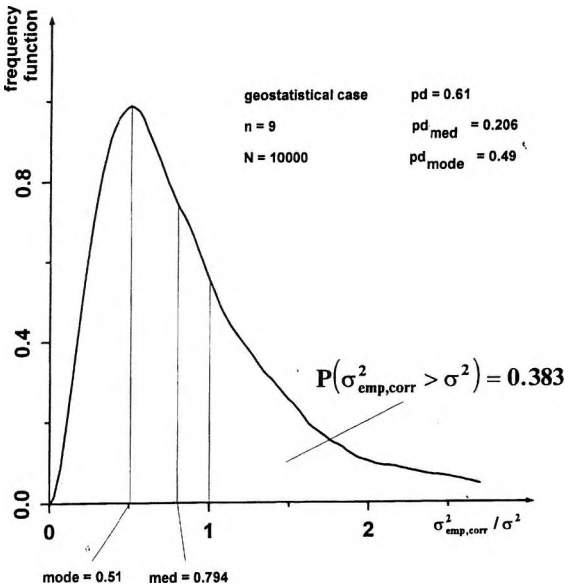


Fig. 1. Three alternatives for characterizing 'probability distortions' of the so-called 'corrected variance' ($\sigma^2_{emp,corr}$): the value of the mode and that of the median of the estimations are significantly less than the true value of the variance (denoted by σ^2), and the probability of the event $\sigma^2_{emp,corr} > \sigma^2$ is significantly less than 0.5. The parent probability distribution calculating this curve was of geostatistical type but these probability distortions occur also if the parent distribution is pure Gaussian, nevertheless in the latter case they are not so considerably (even shockingly) large as those presented here

1. ábra. A „korrigált empirikus variancia" ($\sigma^2_{emp,corr}$) háromféleképpen lehet 'valószínűségi értelemben torzított': a módusz és a medián értékei szignifikánsan kisebbek lehetnek a variancia elméletileg helyes (σ^2-tel jelölt) értékénél, valamint $\sigma^2_{emp,corr} > \sigma^2$ a reláció előfordulásának valószínűsége szignifikánsan kisebb lehet 1/2-nél. Az ábra esetében geostatisztikus típusú volt az anyaeloszlás, de ezek a „valószínűségi torzulások" fellépnek akkor is, ha steril Gauss típusú az anyaeloszlás, csak utóbbi esetben kisebb mértékűek ezek a torzulások az ábrán bemutatottaknál

The value of the denominator is given in Fig. 1. (=0.383) and the other probability value figuring in Eq. (8) is clearly 1–0.383=0.617; therefore *pd* equals 0.61.

It can be more informative in respect of the probability distortion to determine the difference between the median of the estimates (i.e., of the value $med(\sigma^2_{emp,corr})$) and the true σ^2 value. Denoting this difference by pd_{med}, this other characteristic of the probability distortion can be calculated simply as

$$pd_{med} = 1 - med\left(\sigma^2_{emp,corr}\right)/\sigma^2 . \tag{9}$$

The third characteristic of the probability distribution expresses our spontaneous wish: it would be desirable that the mode (maximum) of the $\sigma^2_{emp,corr}$-curve should coincide with the true value σ^2. In this respect it is understandable to define this third characteristic of the probability distortion as

$$pd_{mode} = 1 - mode\left(\sigma^2_{emp,corr}\right)/\sigma^2 . \tag{10}$$

Table I presents the values of all three probability distortions for the geostatistical, Jeffreys and Gaussian parent types — and for the sample sizes n=4; 9; 16; 25; 49; 100; 400; 900. *Figure 2* shows the results as curves; naturally all probability distortions increase with decreasing sample size. The distortion values themselves are considerable even if the sample sizes are not extremely small. Neglecting this distortion is eventually justified at the Gaussian but only at very great sample sizes.

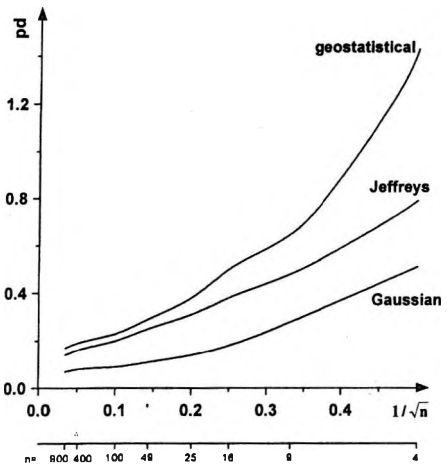
3. Most frequent values (M_k-s) of the estimates $\sigma^2_{emp,corr}$

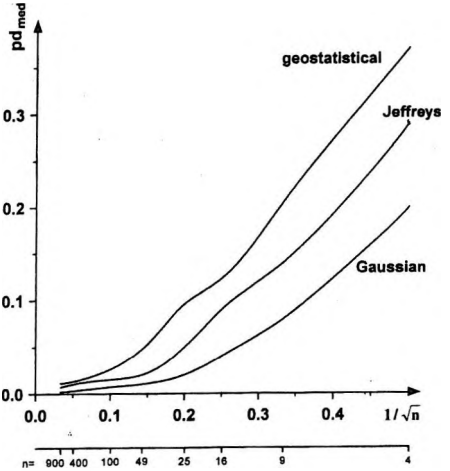
The formulae for determining the so-called 'most frequent values' (MFVs) denoted by M_k simultaneously with the corresponding dispersion-characteristics ε_k called dihesions (k=1/2; 1; 2 and 3) are given, for example, in STEINER [1991], and in STEINER [1989]. These formulae are reproduced in Appendix A/ both for samples and for given density functions.

Figure 3 shows the probability density curve for the same case as was dealt with in Fig. 1. In Fig. 3, however, besides the frequency function the M_k-value of the random variable for k=3 is also given as well as the interval $(M_k-\varepsilon_k, M_k+\varepsilon_k)$ with the corresponding value of the probability of the event

parent distribution type	Values of the probability distortions for different parent distribution types and various sample sizes							
	characteristics of the probability distortions (see Eqs.8., 9. and 10)	size of the sample (n)							
		n=4	n=9	n=16	n=25	n=49	n=100	n=400	n=900
geostatistical	pd=	1.43	0.65	0.50	0.38	0.29	0.23	0.19	0.17
	pd_{med}=	0.371	0.206	0.123	0.096	0.048	0.026	0.013	0.011
	pd_{mode}=	0.80	0.49	0.24	0.16	0.12	0.07	0.03	0.015
Jeffreys	pd=	0.79	0.48	0.38	0.31	0.25	0.20	0.16	0.14
	pd_{med}=	0.29	0.14	0.09	0.05	0.02	0.015	0.01	0.007
	pd_{mode}=	0.78	0.34	0.21	0.14	0.09	0.04	0.02	0.01
Gaussian	pd=	0.51	0.28	0.18	0.14	0.11	0.09	0.08	0.07
	pd_{med}=	0.2	0.08	0.04	0.02	0.01	0.007	0.003	0.002
	pd_{mode}=	0.61	0.32	0.12	0.07	0.03	0.01	0.005	0

Table I. Probability distortions of the variances $\sigma^2_{emp,corr}$ (see Eq. 5)
I. táblázat. A $\sigma^2_{emp,corr}$ varianciák valószínűségi torzítottságai (ld. (5) egyenlet)





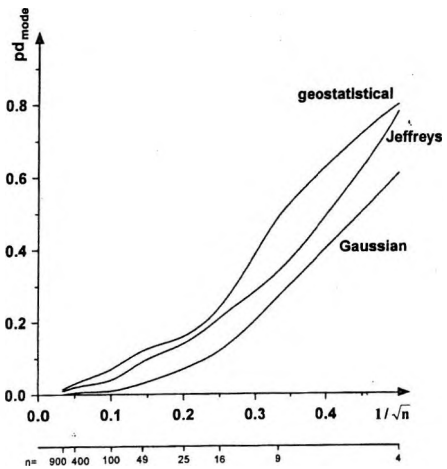


Fig. 2. Curves of the probability distortions *pd*, pd_{med} and pd_{mode} (see Eqs. 8, 9 and 10) versus $1/\sqrt{n}$ where n is the size of the sample. Note that at some frequently occurring parent distributions these distortions cannot be neglected even if the sample size is 400 or 900 (see also Table I)

2. ábra. A (8), (9) és (10) egyenletekkel definiált és *pd*-vel, pd_{med}-del valamint pd_{mode}-dal jelölt valószínűségi torzulások görbéi az $1/\sqrt{n}$ függvényében, ahol n a mintaelemszámot jelenti. Látható, hogy e torzulások viszonylag gyakran előforduló eloszlástípusokra még elég nagy n-eknél sem hanyagolhatók el: a *pd* esetében például még $n=400$, sőt $n=900$ esetén sem

$M_k-\varepsilon_k<\sigma^2_{emp,corr}/\sigma^2<M_k+\varepsilon_k$. It is by no means a surprise that the values of a random variable concentrare around a most frequent value (MFV), i.e. around $M_{1/2}$, M_1, M_2 or M_3 with different but considerable probability. In respect of the appropriate estimate of the true value of the variance (σ^2) naturally MFV = 1 would be desirable — but the estimate $\sigma^2_{emp,corr}/M_k$ trivially fulfils this demand (*Table II* gives the corresponding M_k-values for a great variety of cases). In *Fig. 4*, M_k means the most frequent value of this newly corrected estimate if the $E(estimate)\neq\sigma^2$ holds: according to the traditional definition of unbiasedness this new estimate should be considered as a biased one. However, Fig. 4 clearly shows that this once more corrected estimate

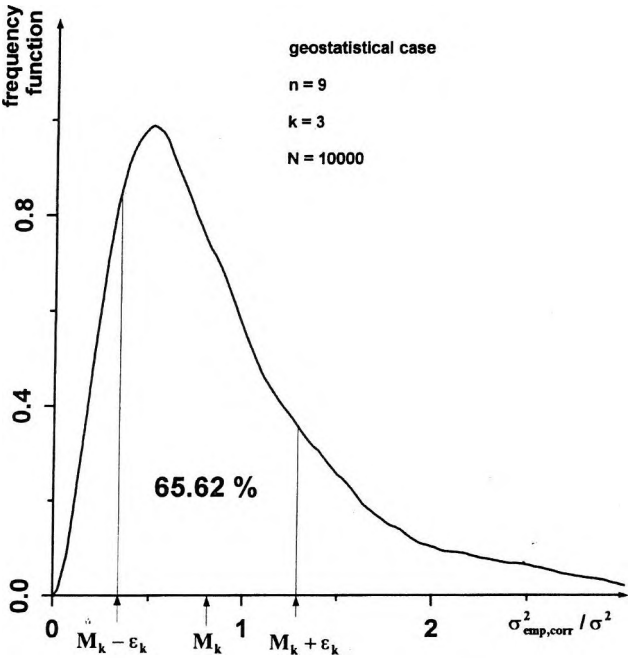


Fig. 3. For just the same case as Fig. 1, symmetrical interval of $2\varepsilon_3$ length (i.e., twice that of the dihesion) marked around the most frequent value M_3. The probability that $\sigma^2_{emp,corr}/\sigma^2$ occurs in this interval is also given in per cent

3. ábra. Az M_3 leggyakoribb érték körüli, $2\varepsilon_3$ (azaz kétszeres dihéziónyi) hosszúságú intervallum az 1. ábrán már megismert esetben; százalékosan adott annak a valószínűsége, hogy a $\sigma^2_{emp,corr}/\sigma^2$ arány ebbe az intervallumba esik

fully meets our practical demands: a considerable percentage of the estimates are concentrated around the true value. Consequently, it is indifferent in practice if our estimates are 'unbiased' in the conventional use of Eq. (8), or generally written: in all asymmetrical cases it is completely indifferent from the point of view of practice if

$$E(estimate) = true\ value \tag{6a}$$

holds or not. The classical notion of 'unbiased estimate' is of no use: it has no practical meaning according to our present investigations. It is logical to demand

$$MFV(estimate) = true\ value \tag{11}$$

and if for the estimate Eq. (11) is fulfilled, the estimate can be called 'unbiased' in the new sense and this adequately corresponds to the demands of the practice.

parent distribution type	M_k and ε_k value-pairs of $\sigma^2_{emp,corr}/\sigma^2$ for different parent distribution types and various sample sizes (see Eqs.A1,...,A4 in the Appendix)							
		size of the sample (n)						
		$n=4$	$n=9$	$n=16$	$n=25$	$n=49$	$n=100$	$n=400$	$n=900$
geostatistical	$M_{1/4}=$	0.433	0.633	0.852	0.882	0.950	0.965	0.986	0.988
	$\varepsilon_{1/4}=$	0.372	0.371	0.346	0.37	0.27	0.187	0.085	0.06
	$M_1=$	0.513	0.684	0.874	0.903	0.958	0.970	0.987	0.988
	$\varepsilon_1=$	0.428	0.398	0.353	0.304	0.236	0.170	0.085	0.06
	$M_2=$	0.631	0.758	0.889	0.918	0.964	0.974	0.987	0.988
	$\varepsilon_2=$	0.519	0.445	0.366	0.320	0.240	0.171	0.085	0.06
	$M_3=$	0.704	0.821	0.906	0.931	0.973	0.978	0.988	0.989
	$\varepsilon_3=$	0.638	0.476	0.375	0.329	0.243	0.172	0.085	0.06
Jeffreys	$M_{1/4}=$	0.568	0.798	0.874	0.909	0.965	0.973	0.988	0.992
	$\varepsilon_{1/4}=$	0.4787	0.424	0.343	0.289	0.222	0.158	0.082	0.054
	$M_1=$	0.663	0.842	0.898	0.928	0.977	0.978	0.989	0.993
	$\varepsilon_1=$	0.5404	0.443	0.350	0.294	0.225	0.159	0.082	0.054
	$M_2=$	0.781	0.900	0.933	0.952	0.990	0.983	0.991	0.993
	$\varepsilon_2=$	0.623	0.471	0.363	0.303	0.229	0.159	0.083	0.054
	$M_3=$	0.847	0.932	0.953	0.965	0.996	0.986	0.991	0.994
	$\varepsilon_3=$	0.673	0.489	0.372	0.308	0.231	0.160	0.083	0.054
Gaussian	$M_{1/4}=$	0.691	0.876	0.940	0.959	0.983	0.986	0.996	0.997
	$\varepsilon_{1/4}=$	0.550	0.428	0.323	0.258	0.184	0.132	0.064	0.043
	$M_1=$	0.769	0.908	0.954	0.968	0.986	0.988	0.997	0.998
	$\varepsilon_1=$	0.590	0.438	0.325	0.260	0.184	0.132	0.064	0.043
	$M_2=$	0.866	0.946	0.973	0.980	0.991	0.992	0.998	0.998
	$\varepsilon_2=$	0.646	0.451	0.329	0.261	0.184	0.132	0.064	0.043
	$M_3=$	0.866	0.965	0.983	0.987	0.994	0.994	0.998	0.998
	$\varepsilon_3=$	0.679	0.458	0.332	0.263	0.185	0.133	0.064	0.043

Table II. MFV (most frequent value) of the variances $\sigma^2_{emp,corr}$ (see Eq. 5)
II. táblázat. A $\sigma^2_{emp,corr}$ varianciák legyakoribb értékei (lásd (5) egyenlet)

4. MFV-corrected estimates of the variance ($\sigma^2_{M_k corr}$)

Empirical variances are defined both in Eq. (1) and Eq. (5); in the first case the estimate is denoted by σ^2_{emp}, in the second one by $\sigma^2_{emp,corr}$ To have M_k-corrected estimates, the new correction factors in both cases depend equally upon F (the parent distribution, characterized here by its distribution function), upon k (which was chosen from the values 1/2; 1; 2 and 3) and naturally from the sample size n. For the M_k-values that are known (see

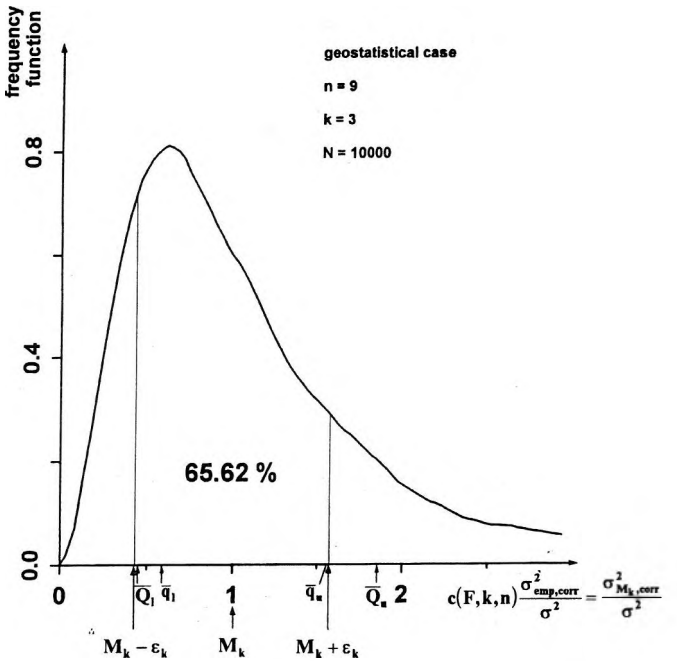


Fig. 4. MFV-corrected variances ($\sigma^2_{emp,cor}$) estimate the true value of the variance (σ^2) in the interval ($M_k-\varepsilon_k$, $M_k+\varepsilon_k$) with considerable probability

4. ábra. Az MFV-korrigált varianciák ($\sigma^2_{emp,cor}$) az ($M_k-\varepsilon_k$, $M_k+\varepsilon_k$) intervallumban jelentős valószínűséggel adnak becslést a variancia σ^2-tel jelölt, elméletileg helyes értékére

Table II), *Table IIIa* presents the coefficients of this new correction in the sense that c and C values are given for calculating in two ways:

$$\sigma^2_{M_k corr} = c(F,k;n) \cdot \sigma^2_{emp,corr} \tag{12}$$

or

$$\sigma^2_{M_k corr} = C(F,k;n) \cdot \sigma^2_{emp}\,. \tag{13}$$

To be able to execute this correction not only for the sample sizes $n=4$, 9, 16, 25, 49, 100, 400 or 900 given in Table IIIa, *Table IIIb* gives the coefficients of simple n-dependent polynomials both for $c(F,k;n)$ and $C(F,k;n)$. Calculations for the coefficients of a simple approximation were made on the basis of the supposition that a polynomial of the form

$$c(F,k;n) = 1 + A_{F,k}/n^{1/2} + B_{F,k}/n + C_{F,k}/n^{3/2} \tag{14}$$

parent distribution type	k	sample sizes (n)							
		$n=4$	$n=9$	$n=16$	$n=25$	$n=49$	$n=100$	$n=400$	$n=900$
geostatistical	½	3.077	1.776	1.320	1.175	1.066	1.046	1.016	1.012
	1	2.596	1.643	1.218	1.153	1.062	1.041	1.015	1.012
	2	2.114	1.484	1.181	1.133	1.058	1.036	1.014	1.012
	3	1.893	1.370	1.177	1.117	1.049	1.032	1.013	1.011
Jeffreys	½	2.346	1.432	1.220	1.145	1.058	1.037	1.014	1.008
	1	2.010	1.356	1.188	1.122	1.045	1.033	1.013	1.008
	2	1.706	1.270	1.143	1.093	1.031	1.027	1.012	1.008
	3	1.573	1.226	1.119	1.078	1.023	1.020	1.011	1.008
Gaussian	½	1.928	1.303	1.134	1.086	1.037	1.024	1.006	1.003
	1	1.733	1.258	1.117	1.076	1.035	1.021	1.005	1.003
	2	1.539	1.208	1.096	1.062	1.021	1.018	1.005	1.003
	3	1.453	1.184	1.085	1.055	1.027	1.016	1.004	1.002

Table IIIa. Correction factors $C(F,k;n)$ to calculate MFV-corrected variances denoted by $\sigma^2_{M_k corr}$ as $C(F,k;n)\cdot\sigma^2_{emp}$

IIIa. táblázat. A $\sigma^2_{M_k corr}$-al jelölt MFV korrigált varianciák σ^2_{emp} alapján történő kiszámításához szükséges $C(F,k;n)$ korrekciós faktorok

is able to give satisfactory approximations of the n-dependency of the MFV-correction factor if $\sigma^2_{emp,corr}$ is to be corrected: that $A_{F,k}$, $B_{F,k}$, $C_{F,k}$ coefficient-triplets which are the best for the actual k-value and for the F parent distribution in question (in the sense of the L_2-norm,) can be easily calculated using the simple least squares techniques, based on the c-values given in Table IIIa.

Obviously the C correction constant, also defines the same value-triplet, as

$$C(F,k;n) = \frac{n}{n-1}\,c(F,k;n) \tag{15}$$

can easily be proven on the basis of Eqs. (1), (5) and (14). The $A_{F,k}$, $B_{F,k}$, $C_{F,k}$ value-triplets obtained are presented in Table IIIb, and *Fig. 5* shows the curves of the C-approximations (some numerical differences from the true C-values are given — in per cent — in the last column of Table IIIb).

parent distribution type	k	$A_{F,k}$	$B_{F,k}$	$C_{F,k}$	maximum of the difference between the true value of C and its approximation in the sample size interval $4 \leq n \leq 900$
geostatistical	½	-0.75	40.86	-37.41	2.3
	1	-0.3	34.79	-35.89	1.9
	2	0.253	18.97	-21.12	1.7
	3	1.74	6.49	-5.68	1.1
Jeffreys	½	3.944	-6.157	18.07	1.2
	1	2.98	-2.08	8.41	1.3
	2	2.03	0.490	2.198	1.2
	3	1.586	1.247	0.334	1.2
Gaussian	½	1.491	1.966	4.958	0.8
	1	1.368	2.02	2.22	0.8
	2	1.03	2.79	-1.09	0.7
	3	0.83	3.17	-2.43	0.7

Table IIIb. $A_{F,k}$, $B_{F,k}$ and $C_{F,k}$ coefficients to have $C(F,k;n)$ for arbitrary n (between $4 \leq n \leq 900$) using Eqs. (14) and (15)

IIIb. táblázat. A $C(F,k;n)$ meghatározásához szükséges $A_{F,k}$, $B_{F,k}$, $C_{F,k}$ koefficiensek tetszőleges ($4 \leq n \leq 900$) n-ekre a (14) és (15) egyenletekben

5. Detailed numerical characterization of the random variable $\sigma^2_{M_k corr}$

The purpose of this point is to characterize in a sufficiently detailed way the measure of concentration of the estimates $\sigma^2_{M_k corr}$ around the true value of the variance (σ^2). For each studied sample size n (see *Table IVa* to *Table IVh*) the corresponding occurrence probabilities are given for 5 lengths of the interval around the correct σ^2-value . Also given are the $\overline{q}_l$ and $\overline{q}_u$ values defined by the following probabilities (denoted for short $\sigma^2_{M_k corr}/\sigma^2$ by y): $P(\overline{q}_l \leq y \leq 1) = 1/4$ and $P(1 \leq y \leq \overline{q}_u) = 1/4$. The differences $\overline{q}_u - \overline{q}_l$ are also given (being more or less independent of k). These latter intervals belonging to the probability of 50 % are enhanced with each density curve of *Figs. 6a, 6b.* and *6c.* (In Figs. 6a, 6b, and 6c, for the discussed three parent distributions and for all eight studied sample sizes the density curve for that k is shown where the difference $\overline{q}_u - \overline{q}_l$ is minimum, see the last column in each part of Table IV.)

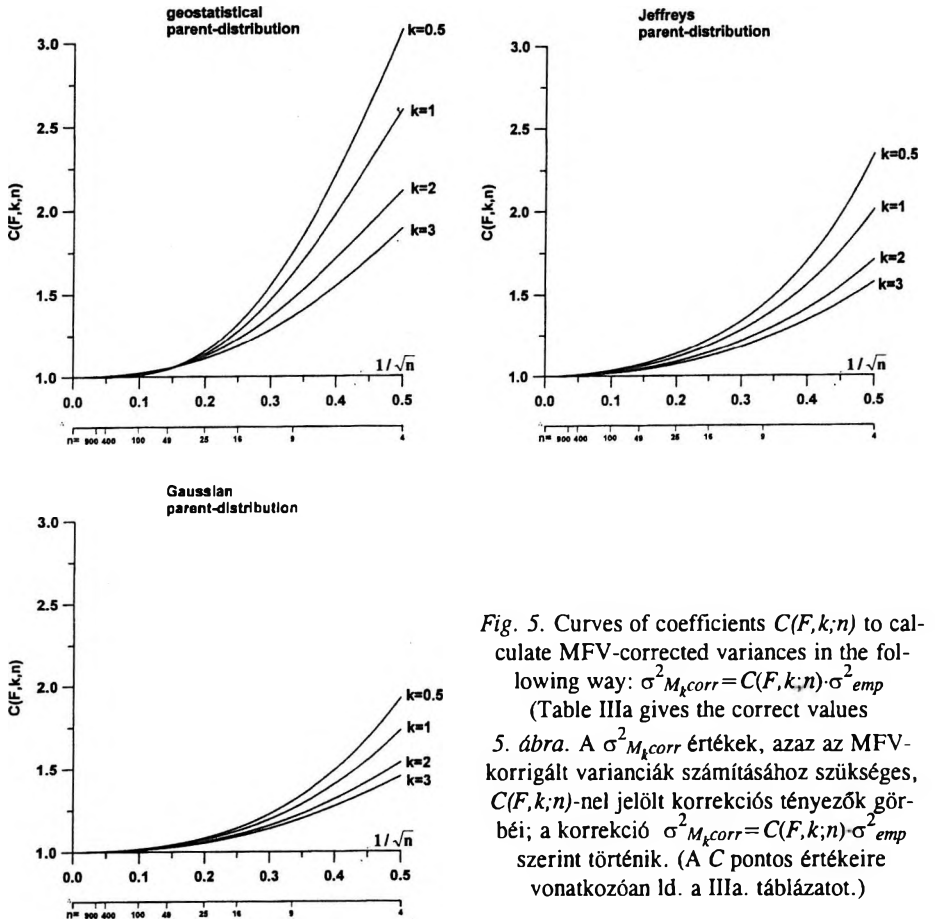


Fig. 5. Curves of coefficients $C(F,k;n)$ to calculate MFV-corrected variances in the following way: $\sigma^2_{M_k corr} = C(F,k;n) \cdot \sigma^2_{emp}$ (Table IIIa gives the correct values

5. ábra. A $\sigma^2_{M_k corr}$ értékek, azaz az MFV-korrigált varianciák számításához szükséges, $C(F,k;n)$-nel jelölt korrekciós tényezők görbéi; a korrekció $\sigma^2_{M_k corr} = C(F,k;n) \cdot \sigma^2_{emp}$ szerint történik. (A C pontos értékeire vonatkozóan ld. a IIIa. táblázatot.)

* * *

It is hoped that our results are found to be convincing: working on the basis of the MFV-concept, i.e., using the P_k-norms instead of the L_2 one, all conventional definitions and statements must be checked. This is very laborious work but unavoidable. It is also hoped, that by introducing the MFV-corrected estimate of the variance, the variograms in the classical geosciences and the scatters (as error-characteristics) in many fields of science and its applications can be determined much more reliably than with the conventional formulae.

Concentration of the MFV corrected variances around the true value of the variance for the sample size n=4							
parent distribution type		probability of the event $1-\Delta < \sigma^2_{M_k corr} < 1+\Delta$					$\bar{q}_u - \bar{q}_l$
	k	Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.1385	0.2836	0.4247	0.5632	0.6584	1.5447
	1	0.1525	0.295	0.4547	0.6108	0.7218	1.4974
	2	0.1562	0.3101	0.4828	0.6571	0.7877	1.5411
	3	0.152	0.3181	0.4885	0.6737	0.8192	1.5931
Jeffreys	0.5	0.1453	0.3016	0.451	0.5973	0.6934	1.4432
	1	0.1562	0.3185	0.4805	0.6463	0.7545	1.3723
	2	0.1669	0.3295	0.5101	0.692	0.8119	1.3632
	3	0.1669	0.3361	0.5194	0.7046	0.8378	1.3969
Gaussian	0.5	0.1695	0.3319	0.4945	0.6424	0.7515	1.2704
	1	0.1768	0.3506	0.521	0.6805	0.7946	1.2191
	2	0.182	0.3684	0.5463	0.7159	0.8407	1.2096
	3	0.1853	0.3734	0.5525	0.7316	0.8579	1.2026

Table IVa. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=4$

IVa. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=4$ esetén

Concentration of the MFV corrected variances around the true value of the variance for the sample size n=9							
parent distribution type		probability of the event $1-\Delta < \sigma^2_{M_k corr} < 1+\Delta$					$\bar{q}_u - \bar{q}_l$
	k	Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.2266	0.4393	0.6214	0.7372	0.7993	0.9664
	1	0.2326	0.4492	0.6432	0.7726	0.8312	0.9679
	2	0.234	0.4545	0.6652	0.8125	0.8643	0.9811
	3	0.2304	0.4584	0.6758	0.8257	0.8807	0.9967
Jeffreys	0.5	0.2592	0.4966	0.687	0.8067	0.8627	0.8321
	1	0.2631	0.5066	0.7086	0.8316	0.8839	0.8165
	2	0.2652	0.5181	0.7292	0.8564	0.9056	0.8159
	3	0.265	0.522	0.7364	0.8698	0.9158	0.8324
Gaussian	0.5	0.2838	0.5629	0.7722	0.8847	0.9305	0.7263
	1	0.2895	0.5663	0.7724	0.885	0.9328	0.7118
	2	0.2952	0.5745	0.7871	0.9026	0.9445	0.7061
	3	0.3001	0.5761	0.7922	0.9104	0.9499	0.7080

Table IVb. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=9$

IVb. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=9$ esetén

Concentration of the MFV corrected variances around the true value of the variance for the sample size n=16							
parent distribution type		probability of the event $1-\Delta < \sigma^2_{M_k\,corr} < 1+\Delta$					$\overline{q}_u - \overline{q}_l$
	k	Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.3504	0.6379	0.8184	0.8999	0.9346	0.6089
	1	0.3539	0.6436	0.8293	0.9099	0.9413	0.5970
	2	0.3569	0.6533	0.8442	0.9231	0.952	0.6095
	3	0.3572	0.6553	0.8522	0.9292	0.9566	0.6188
Jeffreys	0.5	0.3531	0.6355	0.82	0.8979	0.9353	0.5939
	1	0.357	0.6436	0.8326	0.9088	0.9435	0.6013
	2	0.3578	0.6499	0.8478	0.9224	0.952	0.6059
	3	0.3593	0.6559	0.8551	0.9297	0.9564	0.6118
Gaussian	0.5	0.4067	0.7221	0.8924	0.9559	0.9804	0.5015
	1	0.4101	0.727	0.8984	0.9596	0.9824	0.5035
	2	0.4146	0.7325	0.9062	0.9663	0.9848	0.5040
	3	0.4143	0.7356	0.9106	0.9693	0.9859	0.5041

Table IVc. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=16$

IVc. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=16$ esetén

Concentration of the MFV corrected variances around the true value of the variance for the sample size n=25							
parent distribution type		probability of the event $1-\Delta < \sigma^2_{M_k\,corr} < 1+\Delta$					$\overline{q}_u - \overline{q}_l$
	k	Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.3661	0.6611	0.8258	0.8991	0.9387	0.5777
	1	0.3669	0.6734	0.8468	0.914	0.9496	0.5766
	2	0.3657	0.6827	0.862	0.9286	0.961	0.5848
	3	0.3652	0.6853	0.87	0.9369	0.9681	0.5953
Jeffreys	0.5	0.4329	0.7407	0.8902	0.9492	0.9754	0.4764
	1	0.4354	0.7475	0.9022	0.956	0.9783	0.4763
	2	0.4347	0.7547	0.913	0.9628	0.9819	0.4834
	3	0.4348	0.7594	0.9175	0.9662	0.9841	0.4858
Gaussian	0.5	0.5026	0.8263	0.9507	0.9854	0.996	0.3986
	1	0.5049	0.8315	0.9535	0.9866	0.9963	0.3949
	2	0.5107	0.8362	0.958	0.988	0.9967	0.3929
	3	0.5142	0.8382	0.9597	0.9888	0.997	0.3956

Table IVd. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=25$

IVd. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=25$ esetén

Concentration of the MFV corrected variances around the true value of the variance for the sample size n=49							
parent distribution type		probability of the event $1-\Delta < \sigma^2_{M_k\,corr} < 1+\Delta$					$\bar{q}_u - \bar{q}_l$
	k	Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.4892	0.8101	0.9335	0.9743	0.9902	0.4134
	1	0.4916	0.8172	0.9398	0.9776	0.9915	0.4096
	2	0.4981	0.8269	0.946	0.9813	0.9935	0.4099
	3	0.4979	0.8297	0.9495	0.982	0.994	0.4132
Jeffreys	0.5	0.5752	0.8831	0.9689	0.9915	0.9992	0.3384
	1	0.5776	0.89	0.9727	0.9929	0.9994	0.3387
	2	0.5797	0.8957	0.9763	0.9941	0.9995	0.3416
	3	0.5801	0.8978	0.9775	0.9951	0.9995	0.3434
Gaussian	0.5	0.6676	0.9424	0.9941	0.9996	1	0.2733
	1	0.6687	0.9439	0.9941	0.9996	1	0.2732
	2	0.6706	0.947	0.9943	0.9996	1	0.2716
	3	0.6712	0.9476	0.9947	0.9997	1	0.2722

Table IVe. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=49$

IVe. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=49$ esetén

Concentration of the MFV corrected variances around the true value of the variance for the sample size n=100							
parent distribution type		probability of the event $1-\Delta < \sigma^2_{M_k\,corr} < 1+\Delta$					$\bar{q}_u - \bar{q}_l$
	k	Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.6553	0.9378	0.9867	0.998	0.9997	0.2876
	1	0.6587	0.9412	0.988	0.9983	0.9998	0.2872
	2	0.6644	0.9455	0.9904	0.9985	0.9998	0.2877
	3	0.6629	0.9434	0.9893	0.9987	0.9998	0.2901
Jeffreys	0.5	0.7389	0.9705	0.9964	0.9998	1	0.2382
	1	0.7411	0.9719	0.9969	0.9999	1	0.2382
	2	0.7442	0.9735	0.9974	0.9999	1	0.2396
	3	0.745	0.9742	0.9975	0.9999	1	0.2391
Gaussian	0.5	0.8375	0.9926	0.9999	0.9999	1	0.1952
	1	0.8397	0.9931	0.9999	0.9999	1	0.1944
	2	0.8415	0.9932	0.9999	0.9999	1	0.1950
	3	0.8419	0.9934	0.9999	0.9999	1	0.1949

Table IVf. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=100$

IVf. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=100$ esetén

parent distribution type	k	probability of the event $1-\Delta < \sigma^2_{M_k\,corr} < 1+\Delta$					$\bar{q}_u - \bar{q}_l$
		Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.9447	0.9995	1	1	1	0.1410
	1	0.9446	0.9995	1	1	1	0.1409
	2	0.9447	0.9995	1	1	1	0.1401
	3	0.9442	0.9994	1	1	1	0.1391
Jeffreys	0.5	0.9749	1	1	1	1	0.1217
	1	0.9751	1	1	1	1	0.1210
	2	0.9757	1	1	1	1	0.1210
	3	0.9756	1	1	1	1	0.1208
Gaussian	0.5	0.9953	1	1	1	1	0.0937
	1	0.9953	1	1	1	1	0.0936
	2	0.9953	1	1	1	1	0.0936
	3	0.9954	1	1	1	1	0.0937

Table IVg. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=400$

IVg. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=400$ esetén

parent distribution type	k	probability of the event $1-\Delta < \sigma^2_{M_k\,corr} < 1+\Delta$					$\bar{q}_u - \bar{q}_l$
		Δ=0.2	Δ=0.4	Δ=0.6	Δ=0.8	Δ=1.0	
geostatistical	0.5	0.9396	1	1	1	1	0.0953
	1	0.9391	1	1	1	1	0.0952
	2	0.9389	1	1	1	1	0.0950
	3	0.9386	1	1	1	1	0.0949
Jeffreys	0.5	0.9989	1	1	1	1	0.0795
	1	0.9989	1	1	1	1	0.0794
	2	0.999	1	1	1	1	0.0794
	3	0.999	1	1	1	1	0.0793
Gaussian	0.5	1	1	1	1	1	0.0628
	1	1	1	1	1	1	0.0628
	2	1	1	1	1	1	0.0628
	3	1	1	1	1	1	0.0629

Table IVh. Concentration of the MFV-corrected empirical variances around the true value σ^2 for the sample size $n=900$

IVh. táblázat. Az MFV korrigált varianciáknak a σ^2 körüli koncentrációja $n=900$ esetén

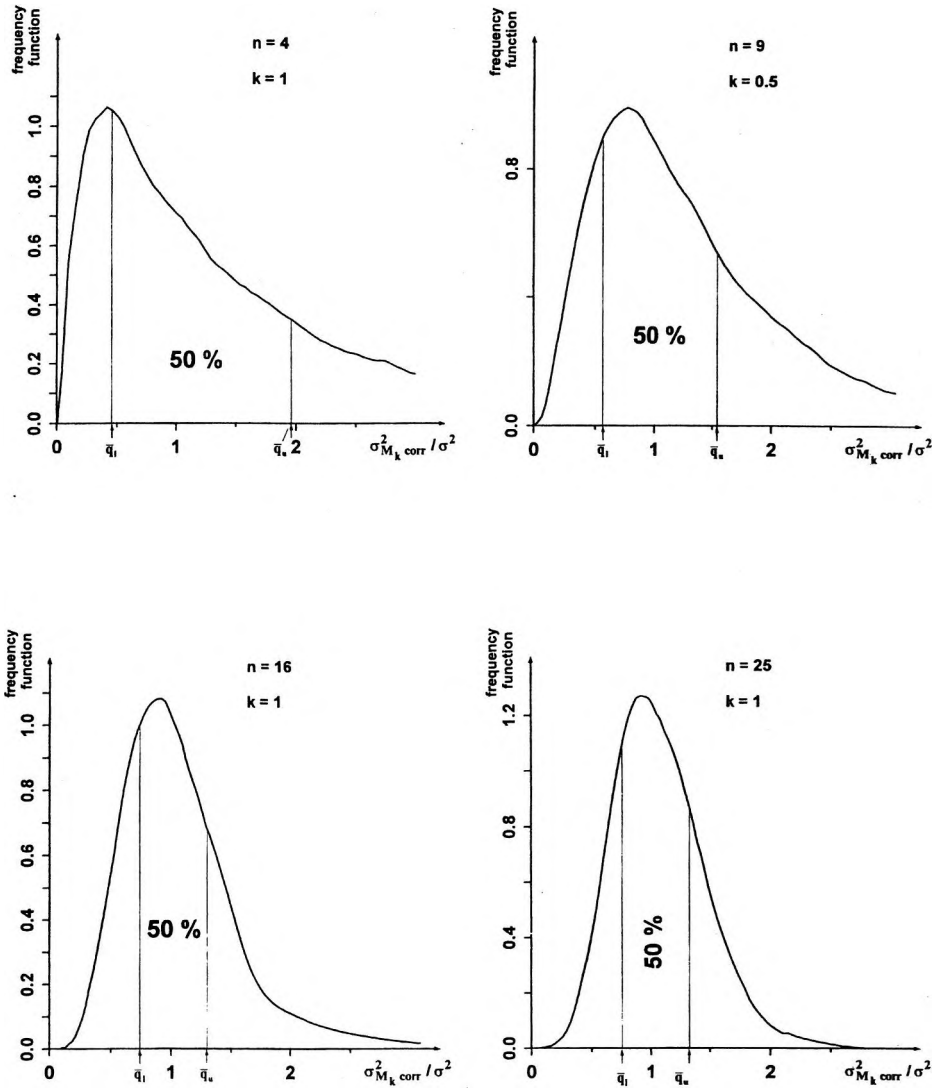


Fig. 6a. (the figure caption is given on the next page)
6a. ábra. (az ábrafelirat a következő oldalon található)

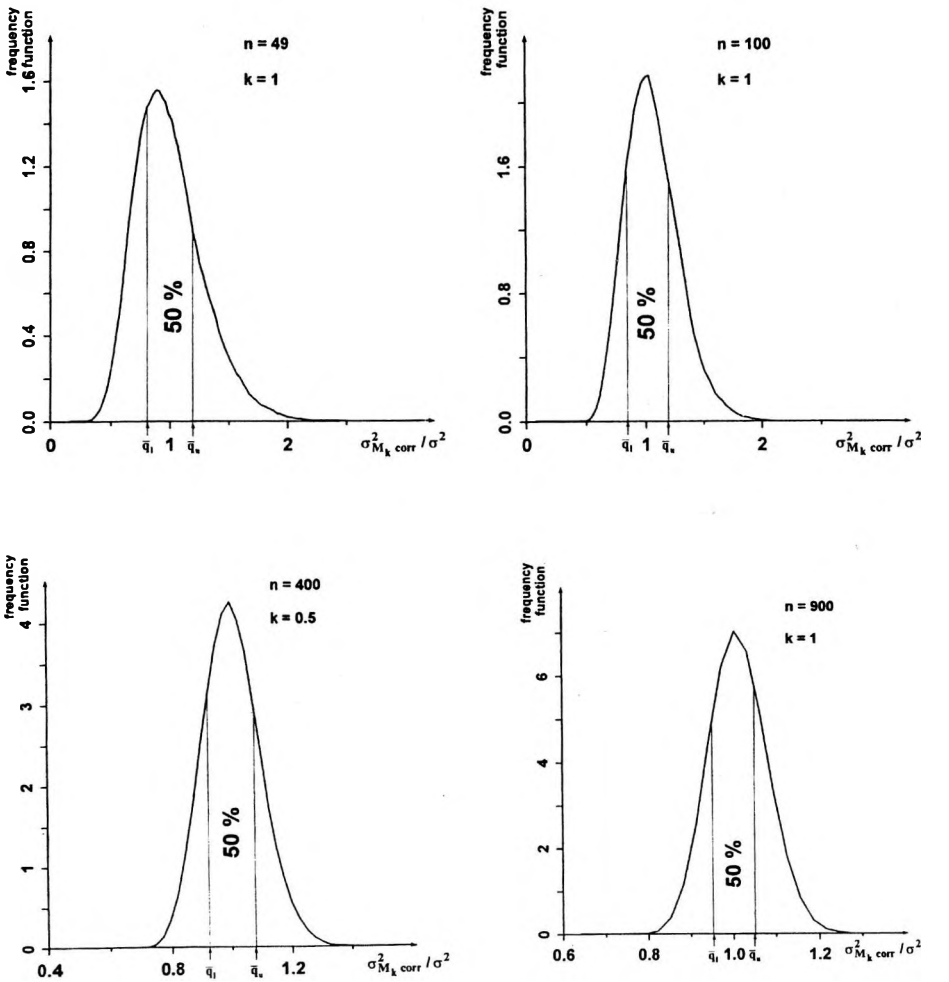


Fig. 6a. Curves of MFV-corrected variances ($\sigma^2_{M_k corr}$) for eight sample sizes (from $n=4$ to $n=900$); the parent distribution is of geostatistical type. The areas between $\bar{q}_l$ and $\bar{q}_u$ are characterized by the probability 0.5; the probability of both events $\bar{q}_l < \sigma^2_{M_k corr} / \sigma^2 < 1$ and $1 < \sigma^2_{M_k corr} / \sigma^2 < \bar{q}_u$ is equally 1/4

6a. ábra. MFV-korrigált varianciák görbéi (azaz $\sigma^2_{M_k corr}$-görbék) nyolcféle mintaméretre ($n=4$-től $n=900$-ig); az anyaeloszlás geostatisztikus típusú. A görbék alatti területeket $\bar{q}_l$ és $\bar{q}_u$ között 1/2 valószínűség jellemzi. Mind a $\bar{q}_l < \sigma^2_{M_k corr} / \sigma^2 < 1$, mind az $1 < \sigma^2_{M_k corr} / \sigma^2 < \bar{q}_u$ valószínűsége 1/4-del egyenlő

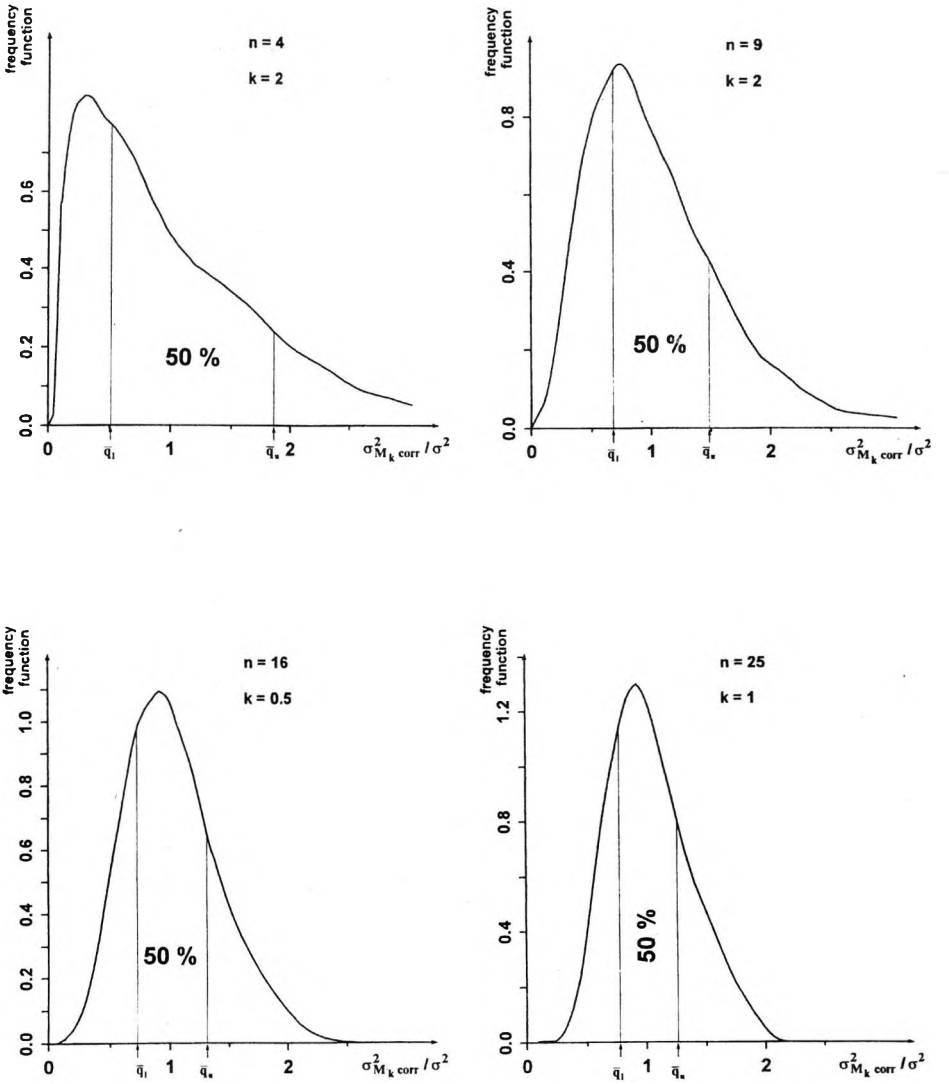


Fig. 6b. (the figure caption is given on the next page)
6b. ábra. (az ábrafelirat a következő oldalon található)

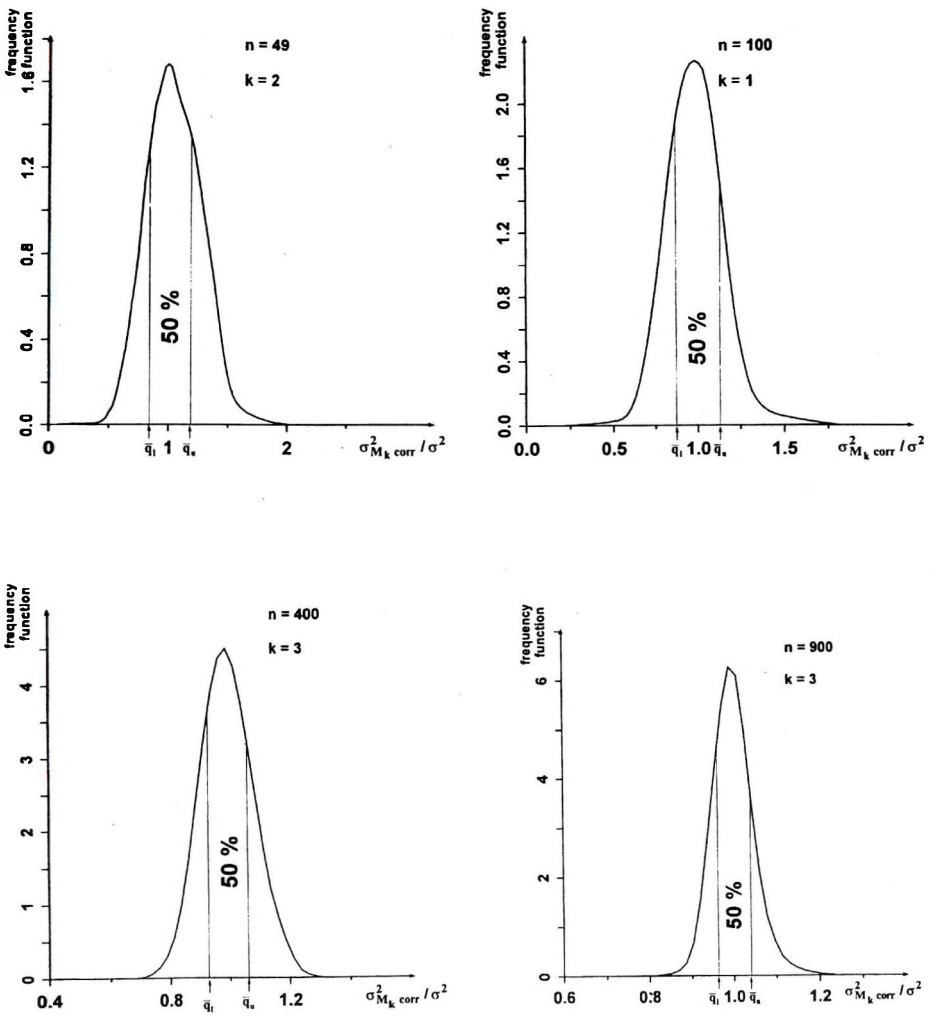


Fig. 6b. Curves of MFV-corrected variances ($\sigma^2_{M_k corr}$) for eight sample sizes (from $n=4$ to $n=900$); the parent distribution is of Jeffreys type. The areas between $\bar{q}_l$ and $\bar{q}_u$ are characterized by the probability 0.5; the probability of both events $\bar{q}_l < \sigma^2_{M_k corr}/\sigma^2 < 1$ and $1 < \sigma^2_{M_k corr}/\sigma^2 < \bar{q}_u$ is equally 1/4

6b. ábra. MFV-korrigált varianciák görbéi (azaz $\sigma^2_{M_k corr}$-görbék) nyolcféle mintaméretre ($n=4$-től $n=900$-ig); az anyaeloszlás Jeffreys típusú. A görbék alatti területeket $\bar{q}_l$ és $\bar{q}_u$ között 1/2 valószínűség jellemzi. Mind a $\bar{q}_l < \sigma^2_{M_k corr}/\sigma^2 < 1$, mind az $1 < \sigma^2_{M_k corr}/\sigma^2 < \bar{q}_u$ valószínűsége 1/4-del egyenlő

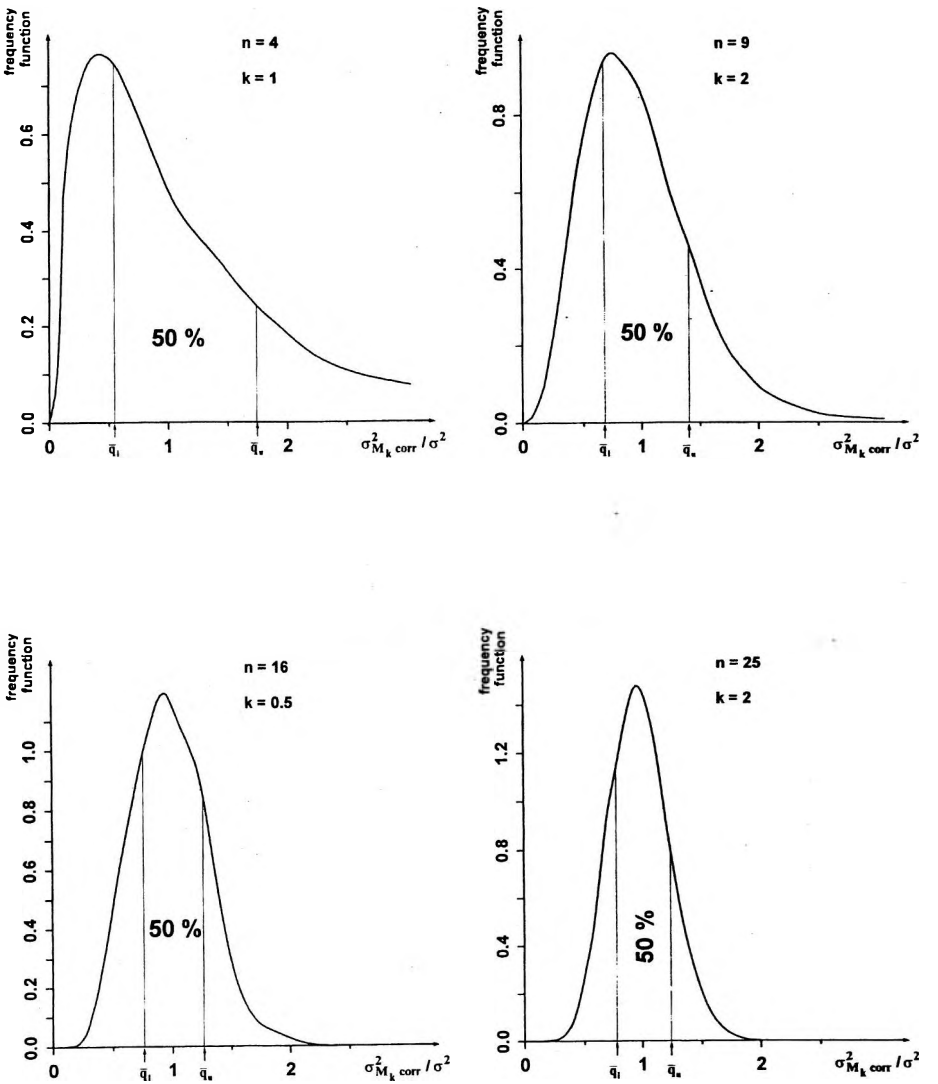


Fig. 6c. (the figure caption is given on the next page)
6c. ábra. (az ábrafelirat a következő oldalon található)

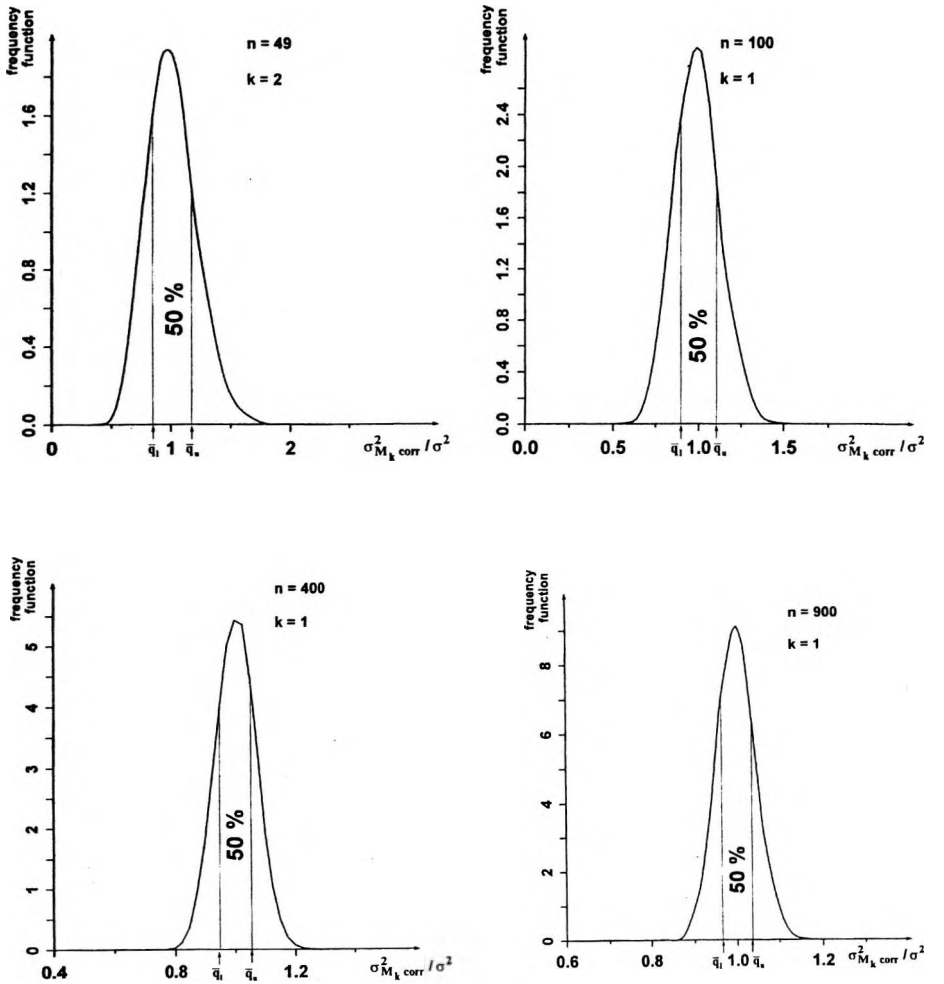


Fig. 6c. Curves of MFV-corrected variances ($\sigma^2_{M_k corr}$) for eight sample sizes (from $n=4$ to $n=900$); the parent distribution is of Gaussian type. The areas between $\bar{q}_l$ and $\bar{q}_u$ are characterized by the probability 0.5; the probability of both events $\bar{q}_l < \sigma^2_{M_k corr}/\sigma^2 < 1$ and $1 < \sigma^2_{M_k corr}/\sigma^2 < \bar{q}_u$ is equally 1/4

6c. ábra. MFV-korrigált varianciák görbéi (azaz $\sigma^2_{M_k corr}$-görbék) nyolcféle mintaméretre (n=4-től n=900-ig); az anyaeloszlás Gauss típusú. A görbék alatti területeket $\bar{q}_l$ és $\bar{q}_u$ között 1/2 valószínűség jellemzi. Mind a $\bar{q}_l < \sigma^2_{M_k corr}/\sigma^2 < 1$, mind az $1 < \sigma^2_{M_k corr}/\sigma^2 < \bar{q}_u$ valószínűsége 1/4-del egyenlő

Appendix

A/ Formulae for calculating M_k and ε_k based on samples: the equations

$$M_k = \frac{\sum_{i=1}^{n} \frac{x_i}{(k\varepsilon_k)^2 + (x_i - M_k)^2}}{\sum_{i=1}^{n} \frac{1}{(k\varepsilon_k)^2 + (x_i - M_k)^2}} \tag{A1}$$

and

$$\varepsilon_k^2 = 3\frac{\sum_{i=1}^{n} \frac{(x_i - M_k)^2}{\left[\varepsilon_k^2 + (x_i - M_k)^2\right]^2}}{\sum_{i=1}^{n} \frac{1}{\left[\varepsilon_k^2 + (x_i - M_k)^2\right]^2}} \tag{A2}$$

must be simultaneously fulfilled. (These formulae were applied in the Monte Carlo investigation of the present paper executed for estimates of variances obtained for samples from geostatistical and Jeffreys parent type distributions.)

B/ Formulae for calculating M_k and ε_k based on given probability density functions $f(x)$: the equations

$$M_k = \frac{\int_{-\infty}^{\infty} \frac{x}{(k\varepsilon_k)^2 + (x - M_k)^2} f(x)\mathrm{d}x}{\int_{-\infty}^{\infty} \frac{1}{(k\varepsilon_k)^2 + (x - M_k)^2} f(x)\mathrm{d}x} \tag{A3}$$

and

$$\varepsilon_k^2 = 3\frac{\int_{-\infty}^{\infty} \frac{(x - M_k)^2}{\left[\varepsilon_k^2 + (x - M_k)^2\right]^2} f(x)\mathrm{d}x}{\int_{-\infty}^{\infty} \frac{1}{\left[\varepsilon_k^2 + (x - M_k)^2\right]^2} f(x)\mathrm{d}x} \tag{A4}$$

must be simultaneously fulfilled.

C/ The density- and distribution functions (for the standard case, i.e., if for the parameter of location (T) and that of scale (S) $T=0$ and $S=1$ hold) for the geostatistical and Jeffreys types of probability distribution and for the Gaussian are the following:

$$f_{st}(x) = \frac{3}{4} \cdot \frac{1}{\left(1+x^2\right)^2 \cdot \sqrt{1+x^2}} \tag{A5}$$

$$F_{st}(x) = \frac{1}{2} + \frac{x}{2\sqrt{1+x^2}} + \frac{x}{4\left(1+x^2\right)\sqrt{1+x^2}} \tag{A6}$$

$$f_J(x) = \frac{35}{32} \cdot \frac{1}{\left(1+x^2\right)^4 \cdot \sqrt{1+x^2}} \tag{A7}$$

$$F_J(x) = F_{st}(x) + \frac{3x}{16\left(1+x^2\right)^2 \sqrt{1+x^2}} + \frac{5x}{32\left(1+x^2\right)^3 \sqrt{1+x^2}} \tag{A8}$$

$$f_G(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{A9}$$

$$F_G(x) = \int_0^x f_G(y)\mathrm{d}y \tag{A10}$$

calculated by computer with the approximation formula given, for example, in STEINER [1990] as Eq.(1–45).

The maximum difference between these approximate values and the correct values of $F_G(x)$ equals 10^{-7}. For generating samples of geostatistical and Jeffreys types, random numbers distributed equally between 0 and 1 are based on the distribution functions $F_{st}(x)$ and $F_J(x)$ — in the sense demonstrated in STEINER [1990] (Fig. 5.2) .

REFERENCES

CRAMÉR H. 1945: Mathematical Methods of Statistics. Almqvist & Wiksells, Uppsala, 575. p.

DUTTER R. 1986/87: Mathematische Methoden in der Montangeologie. Vorlesungsnotizen. Manuscript, Leoben

HAJAGOS B., STEINER F. 1994: Bestimmungsfehler der statistischen Momente. Publ. Univ. Miskolc Series D, **35**, No.4. pp.65–95

KERÉKFY P. 1978: On robust estimates. (in Hungarian) Alkalmazott Matematikai Lapok, **4**, pp. 327-357

KIS M., HURSÁN G. 1995: Current distortions of the corrected empirical variance. (in Hungarian) Poster presentation. V. Conference for Geomathematics 5th Oct. 1995 Szeged

MATHERON G. 1965: Les variables regionalisées et leur estimation. Masson & Cie, Paris, 305. p.

STEINER F. 1990: Principles of geostatistics. (in Hungarian) Tankönyvkiadó, Budapest, 363. p.

STEINER F. ed. 1991: The most frequent value. Akadémiai Kiadó, Budapest, 315. p.

MFV-korrigált varianciák

STEINER Ferenc, HAJAGOS Béla és HURSÁN Gábor

A konvencionálisan korrigált empirikus varianciának ($\sigma^2_{emp,corr}$-nak) még az adatok steril Gauss-típusú eloszlása esetén is vannak ún. „valószínűségi torzulásai". Ezek ellen úgy védekezhetünk, hogy MFV-korrekciót alkalmazunk; az ehhez szükséges korrekciós faktorokat a IIIa. táblázat mutatja be három anyaeloszlás-típusra és különböző n mintaméretekre (görbékkel az 5. ábra szemlélteti ezeket). A IIIb. táblázat együttható-hármasai segítségével e korrekciós faktorok jó közelítéssel számíthatók bármely n mintaméretre (megjegyzendő, hogy ilyen korrekcióra $n < 1000$ esetén a tárgyalt anyaeloszlásoknál már nincs is szükség). A $\sigma^2_{M,corr}$-eket, mint valószínűségi változókat, a gyakorlat szemszögéből a legfontosabb adataikkal a IVa.-IVh. táblázatok mutatják be.

A gyakorlati célú vizsgálatok felvetették azt a kérdést, hogy a b becslések torzítatlanságának általánosan elfogadott elméleti definíciója (nevezetesen hogy ti. b-t torzítatlan becslésnek nevezzük, ha a becslések számértékeinek várható értéke azonos az elméletileg helyes értékkel) kritika nélkül alkalmazható-e a gyakorlatban. A torzítatlanság kritériumaként a gyakorlat inkább azt igényelné, hogy a dolgozat szerint a becslések számértékeinek leggyakoribb értéke legyen egyenlő a valódi értékkel.