Jour. Korean Earth Science Society, Vol. 23, No. 1, p. 15–29, January 2002

# Bayesian Inversion of Gravity and Resistivity Data: Detection of Lava Tunnel

**Byung-Doo Kwon[1,*] and Seok-Hoon Oh[2]**

[1]Seoul National University, San 56-1, Shilim-dong, Kwanak-ku, Seoul, 151-742, Korea
[2]Marine Meteorology and Earthquake Research Laboratory, METRI/KMA, 460-18, Shindaebang-dong, Dongjak-gu, Seoul, 156-720, Korea

**Abstract:** Bayesian inversion for gravity and resistivity data was performed to investigate the cavity structure appearing as a lava tunnel in Cheju Island, Korea. Dipole-dipole DC resistivity data were proposed for *a prior* information of gravity data and we applied the geostatistical techniques such as kriging and simulation algorithms to provide *a prior* model information and covariance matrix in data domain. The inverted resistivity section gave the indicator variogram modeling for each threshold and it provided spatial uncertainty to give *a prior* PDF by sequential indicator simulations. We also presented a more objective way to make data covariance matrix that reflects the state of the achieved field data by geostatistical technique, cross-validation. Then Gaussian approximation was adopted for the inference of characteristics of the marginal distributions of model parameters and Broyden update for simple calculation of sensitivity matrix and SVD was applied. Generally cavity investigation by geophysical exploration is difficult and success is hard to be achieved. However, this exotic multiple interpretations showed remarkable improvement and stability for interpretation when compared to data-fit alone results, and suggested the possibility of diverse application for Bayesian inversion in geophysical inverse problem.

Key words: bayesian inversion, geostatistical simulation, gaussian approximation, cavity detection

## INTRODUCTION

All of the finite approximations for geophysical inversion made the process be a probabilistic inference problem. Of various approaches to the goal, Bayesian frame has provided robust and stable inversion system to incorporate *a prior* information, and the *a posterior* probability density distribution (PDF) is the complete solution of the problem (Tarantola, 1987).

In the past decade, Bayesian analysis (Bayes, 1763) for geophysical inverse problem was tried sometimes as test problems (Duijndam, 1988a, b) or sometimes as alternatives (Mosegaard and Tarantola, 1995; Gouveia, 1996; Moraes, 1996) to the traditional deterministic process (Lines and Treital, 1984). All of them showed the robustness of this method. But most of their approaches were rather arbitrary and subjective, so that it was hard to attack the geophysical inverse problem in a systematic process within Bayesian frame.

The most difficult and significant aspect of Bayesian frame is the preparation of *a prior* information. Duijndam (1988a, b) arbitrarily assumed 1st and 2nd statistical moments of model parameters. Moraes (1996) preliminarily defined the probabilistic structure of the model parameters and then used maximum entropy technique inversely. Gouveia (1996) incorporated various experimental noises into model and data covariance with subjectivity. Scales and Tarantola (1994) thought the smooth components of the P-wave sonic log as representing the geological structure and others erratic random fluctuations. But it is too vague to distinguish between signal and noise. And most of the above algorithms were based on 1-dimensional well logging data.

We present here a practical and objective mechanism to construct two-dimensional *a prior* information from geophysical data by geostatistical simulations and applied it to generate model and data covari-

*E-mail: bdkwon@mantle.snu.ac.kr

ance matrix.

Geostatistics gives tools to analyze and interpret spatially distributed data. It is based on variogram model to estimate and simulate parameters on unsampled points or blocks with *probabilistic* information. The estimation may come from one variable or combined several variables. We used the indicator kriging and simulation technique to generate cumulative probability density functions from various geophysical data for *a prior* information. Such techniques can generate probability density function independent of any statistical moments of data.

We adopted the Gaussian approximation to describe the probabilistic state of our information. Actually Bayesian approach to geophysical inversion is simply the product of two kinds of PDF, *a prior* and likelihood PDF (Tarantola, 1987). But in the case of parameterization of geophysical inversion, high-multidimensional state of *a posterior* PDF makes interpreters dizzy and doesn't give tractable information to infer the substructure. So a marginalization process is required. This process makes the Bayesian approach hard problem to solve, because it needs high dimensional integral as much as the number of parameters. In addition, the kernel of the integral is highly non-linear.

Nonparametric and general treatment of PDF of the data is very powerful and doesn't lose any information in analysing the substructure. But it requires high costs to get marginalized information of each parameter. So we used the local optimization technique to investigate the characteristics of *a posterior* information by Gaussian approximation which makes the mathematical treatment easy. This approach needs two covariance matrices, data and *a prior* model space unlike the process of nonparametric one such as a Monte Carlo technique. Model covariance is easily generated from *a prior* information but the data covariance should be given independently and we applied the cross-validation technique to give objective and practical result. Then we analyse the posterior covariance matrix by geostatistical simulation dependent upon it.

To verify the algorithm, we proposed a model structure that consists of Schlumberger, resistivity well log and DC dipole-dipole data. And finally we applied the above algorithm to in-situ geophysical data. Two-dimensional dipole-dipole resistivity data was used to construct *a prior* information for gravity inversion surveyed for detection of lava tunnels in Cheju Island, Korea. This exotic combination successfully found the cavity area with high probability.

## *A PRIOR* INFORMATION BY GEOSTATISTICAL APPROACH

Geostatistics is largely based upon the random function model, whereby the set of unknown values is regarded as a set of spatially dependent random variables. Such presentation reflects our imperfect knowledge of the unsampled value z(u) and more generally, of the distribution of z within the area (Goovaerts, 1997). Our ultimate goal is to get *a prior* information from various geophysical data and to infer the uncertainty structure of an interpreted region.

After all, this is the estimation of unsampled points or area spatially dependent on random models or variables, so it coincides with that of geostatistical model. Detailed tutorials for geostatistics are found on Isaaks and Srivastava (1989) and Goovaerts (1997), etc.

### Some Terminology of Geostatistics

**Kriging.** Kriging, most frequently quoted terminology in following text, is a linear, unbiased, least-squares spatial interpolation method; a weighted-average, or weighed-mean, estimator whose weights are functions of spatial covariance. And the spatial covariance is derived from simple calculation called *semivariogram* (or traditionally variogram) given by following equation:

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}(Z(x)-Z(x+h))^2 \qquad (1)$$

Variogram (h) describes the variance distribution of the data given by relative location difference h. Z(x) is the random function of being attained at x, and N(h) is the number of sample points conditional to the estimated point. This is the assessment of local uncertainty and we applied it to derivation of data covariance matrix.

Simulation. A geostatistical simulation algorithm aims to draw realizations that reflect the statistics modeled from the data. The question is how well these statistics should be reproduced. The smoothing effect of kriging, and more generally, of any low pass-type interpolation, is due to a missing error components (Deutsch and Journel, 1992). Consider the random function Z(u) as the sum of the estimator and the corresponding error:

$$Z(u) = Z(u)^* + R(u).$$

Kriging, for example, would provide the smoothly varying estimator $Z^*(u)$. To restore the full variance of the random function model, one may think of simulating a realization of the random function error with zero mean and the correct variance and covariance. The simulated z-value would be the sum of the unique estimated value and a simulated error value:

$$z_c^{(l)} = z^*(u) + r^{(l)}(u).$$

In this process, the error component R(u) must be independent or at least orthogonal to the estimatorand the random function R(u) modeling the error must have the same spatial distribution or, at least, the same covariance as the actual error (Deutsch and Journel, 1992).

### Inference of Uncertainties of Blocks

The points where the simulations estimate are denser than the number of the parameterized blocks for inversion. This is due to the fact that we couldn't infer the uncertainty of the blocks parameterized for inversion. So by assessing the values on denser points, we derive values by following indirect method.

We consider the problem of evaluating the block ccdf (conditional cumulative density function) that models the uncertainty about an average z-value over the block V(u) conditional to (n) neighborhood data:

$$F_V(u; z|(n)) = \text{Prob}\{u \le z|(n)\}. \quad (2)$$

Because of the non-linearity of the indicator transform, the block ccdf cannot be derived simply as a linear combination of point ccdfs:

$$[F_V(u; z|(n))]^* \neq \frac{1}{J}\sum_{j=1}^{J}[F(u'_j; z|(n))]^* \quad (3)$$

with the point-ccdf $F_V(u'_j; z|(n))$ being defined at J points $u'_j$ discretizing the block V(u). $[F_V(u; z|(n))]^*$ is the inference of ccdf for the block V by conditional to n data around the point u. In the absence of block data $z_V(u_\alpha)$ and corresponding block statistics and block indicator data, the block ccdf (eq. 2) can be numerically approximated by the cumulative distribution of many simulated block values $z_V^{(l)}(u)$ (Isaaks, 1990; Gomez-Hernandez, 1991; Deutsch and Journel, 1992; Glacken, 1996):

$$[F_V(u; z|(n))]^* = \frac{1}{L}\sum_{l=1}^{L} i_V^{(l)}(u; z) \quad l = 1, \ldots, L \quad (4)$$

with the block indicator value defined $i_V^{(l)}(u; z) = 1$ if $z_V^{(l)}(u) \le z$, and zero otherwise. Each simulated block value $z_V^{(l)}$ is obtained by averaging a set of z-values simulated at J points $u'_j$ discretizing the block V(u):

$$z_V^{(l)}(u) = \frac{1}{J}\sum_{j=1}^{J} z^{(l)}(u'_j). \quad (5)$$

## FAST BAYESIAN INVERSION BY GAUSSIAN APPROXIMATION

Tarantola (1987) showed that the complete solution of geophysical inverse problem is *a posterior* information given by the conjunction of the two states of information, *a prior* and theoretical or likelihood information, and this is accord form to the generalized Bayes' theorem. That is,

$$\sigma(m) = k\rho(m)L(m) \quad (6)$$

where, $\sigma(m)$ is *a posterior* PDF, $\rho(m)$ is *a prior*

PDF, and L(m) is likelihood PDF, is given for the solution of Bayesian inversion. Actually, thus the simple product of *a prior* PDF and likelihood PDF is the solution to the geophysical inverse problem in the viewpoint of probabilistic approach. But high-multidimensional PDF itself gives no choice and when we should determine the most probable model for consulting or inferring the characteristics of the neighborhood of maximum *a posterior* PDF, it requires to solve high dimensional integral or optimization problem called marginalization process. Unfortunately, the kernel of the integral is very highly nonlinear, so that analytic solution is very rare and we should attack it by numerical analysis.

Of the various techniques, the Monte Carlo method to analyze and estimate the characteristics of the *a posterior* PDF is robust and provides full information for the intended substructure to study. But it costs extensive computational power and takes too long to interpret by this technique. These days, operation decisions based on in-situ prompt interpretation of field data have been required, so that fast Bayesian inversion for stable and plausible analysis is presented here.

Thus, as the compensation for the robust and full Monte Carlo analysis, we present Gaussian approximated Bayesian inversion scheme. Gaussian approximation describes, as already known, the distribution by only two moments, mean and varance (or standard deviation). So mathematical implementation and realization is easy and we may apply to local optimization techniques to estimate the state of maximum *a posterior* PDF. But deficient consideration for statistical moments makes the interpreted PDF smooth in terms of maximum entropy method and the assumption for the distribution may be not correct for all kind of the data.

However, as previously mentioned, as the compromise for the post robust interpretation, it provides a swift and simple but stable analysis tools for in-situ operations. We also propose a practical algorithm to implement an objective covariance matrix especially in data domain by geostatistical technique, cross-validation.

### Covariance Matrices

Unlike the Monte Carlo approach that is fundamentally non-parametric, local optimization with Gaussian assumption requires the mean vector and covariance matrix for multidimensional problem. That is, by Tarantola (1987), multidimensional normal probability distribution for observed data and model parameters is given by

$$L(d|m) \propto \exp(G(m)-d_{obs})^T C_d^{-1}(G(m)-d_{obs}) \quad (7)$$

where, L(d|m) is likelihood distribution, G(m) is forward model operator, $d_{obs}$ is observed data and $C_d$ is a covariance matrix for the observed data. To make the maximum of the eq. 7 is called maximum likelihood method (MLM) and of various techniques for MLM, least-square method may be used. Traditionally, the inverse of covariance $C_d^{-1}$ was simply treated like a weighting matrix which have only diagonal components, variances (Constable and Parker. 1987). They allocated low values for points showing bad S/N ratio with subjectivity or simply indiscriminated ones. This was due to the reason that naturally we can't infer the $C_d^{-1}$.

In this study, we adopt a cross-validation technique, one of the geostatistical methods to infer the $C_d^{-1}$ with maximum objectiveness. Originally this was devised to examine the appropriateness of the applied kriging method. As we mentioned above, parameters that define the kriging coefficients are various and depend on spatial configuration of the observed data, so that different configurations make the kriging values different to estimate unsampled points. But areas with small noise and smooth variation will be continuously estimated. Now, the cross-validation scheme is appeared in Fig. 1. As in the figure, we intentionally eliminate a sampled point to estimate kriging error (b), and conduct various krigings with remaining samples to estimate the eliminated point (c), then compare the difference between true and estimated value to figure out the data domain covariance matrix and mean (d).

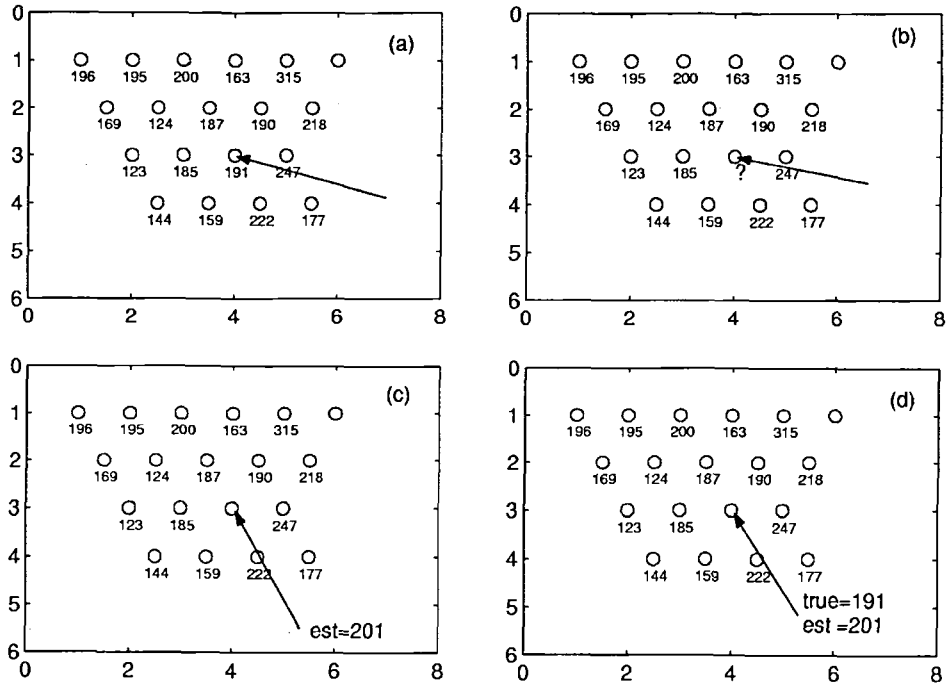


**Fig. 1.** Cross-validation algorithm to present covariance matrix in data domain. (a) an arbitrary sampled point is eliminated with an intention. (b), (c) And estimate the eliminated point with the other sampled points by various Kriging and simulations. (d) the difference provides information of error distribution of observed data.

Goovaerts (1997) and Journel and Huijbregts (1978) published that the mean and variance of cumulative density function by multidimensional normal distribution is identified with those of simple kriging. Thus, we have assumed all of the PDF is Gaussian, this treatment is rational. And in the process of kriging, the remaining samples weigh the inference of unsampled points, we objectively constitute the off-diagonal elements of the covariance matrix. On the other hand, this scheme constrains the data smoothing, due to the property of kriging which minimizes the variance of error between random function and estimation. The covariance matrix for the *a prior* information was composed by simply calculating the mean and covariance from the previously simulated data.

## Approximate Analysis of a Posterior Pdf by Local Optimization

In Gaussian assumption, the Bayesian inversion results to *a posterior* p.d.f.,

$$\sigma(m|d) \propto \exp(G(m) - d_{obs})C_d^{-1}(G(m) - d_{obs}) + (m - m_{prior})C_m^{-1}(m - m_{prior})) \quad (8)$$

where, $C_m^{-1}$, $m_{prior}$ are respectively mean and covariance matrix of the previously presented *a prior* information and $C_d^{-1}$ which is covariance of the data domain was implemented by cross-validation scheme (Tarantola, 1987). Like the case of non-parametric approach previously described, above PDF also makes interpretations difficult. So we used a local optimization technique to extract the information in the neighborhood of maximum *a posterior* PDF. It may lose the generality and robustness but gives promptness for compensation as we intended at first. Now, to get a parameter vector m which makes the eq. 8 maximum, we differentiate S which is the exponent of the $\sigma(m)$ with respect to m,

$$\frac{\partial S}{\partial m} = g = 2A^T C_d^{-1}(G(m) - d_{obs}) + 2C_m^{-1}(m - m_{prior}) \quad (9)$$

that is a gradient of eq. 8, and it should be zero.

Here, A is the sensitivity matrix of the forward model operator G. For ith iteration, we let

$$G(m_{j+1}) = G(m_j) + A\Delta m + O(n)^2 \quad (10)$$

where, $m_{j+1} = m_j + \Delta m$ and eliminate above of second order perturbations, we may write

$$(A^T C_d^{-1} A + C_m^{-1})\Delta m_j = -(A^T C_d^{-1}(G(m_j) - d_{obs}) + C_m^{-1}(m_j - m_{prior})). \quad (11)$$

And again by letting $(A^T C_d^{-1} A + C_m^{-1}) = H$, above equation may be written,

$$H\Delta m_j = -g_j, \quad (12)$$

and to solve above equation, we applied the SVD (Lanczos 1961) and Marquardt-Levenberg constraints (Lines and Treital, 1984). Also in this study, we used Broyden update to calculate the sensitivity matrix A (Broyden, 1965; Jung, 1998). Broyden update is given by,

$$A_{j+1} = A_j + \frac{(y_j - A_j\Delta m_j)\Delta m_j^T}{\|\Delta m_j\|_2^2} \quad (13)$$

where, $y_j = F(m_{j+1}) - F(m_j)$, $\Delta m_j = m_{j+1} - m_j$ which initial sensitivity matrix $A_j$ was taken from reciprocity theorem. Broyden update is somewhat different to the Jacobian matrix given by reciprocity theorem but its behavior is similar to that and it does not need many forward calculation results, so that it's very stable (Jung, 1998).

Finally with the initial model $m_{prior}$, we iterate eq. 12 until discrepancy between observed and calculated data is reduced to standard deviation,

$$(G(m) - d_{obs})^T C_d^{-1}(G(m) - d_{obs}) \le N \quad (14)$$

where N is the total number of observed data (Gouveia, 1996).

### Uncertainty Analysis

By Tarantola (1987), eq. 8 may be re-written

$$\sigma_g(m) \propto [(m - m_{map})^T C'^{-1}_m (m - m_{map})] \quad (15)$$

where, $C'_m$ is covariance of *a posterior* PDF, $m_{map}$ is maximum *a posterior* model vector obtained from above Bayesian inversion. *A posterior* covariance $C'_m$ such a nonlinear system is given,

$$C'_m = [A^T C_d^{-1} A + C_m^{-1}]^{-1}. \quad (16)$$

When we adopt the diagonal elements of the covariance $C'_m$, we may get a model vector that has confidence intervals as a function of the standard deviation. But this analysis may give simple aspects for the inferred model, but more information can be achieved by sampling the various models dependent upon the *a posterior* covariance. So, we applied the LU decomposition simulation (Davis, 1987; Alabert, 1987), one of the geostatistical Gaussian simulation, to sample models dependent on *a posterior* covariance. In geostatistics, people use the variogram model to get spatial covariance information. But we have already have the *a posterior* one, we simply replace the variogram model with covariance. Therefore the algorithm is simple. The *a posterior* covariance matrix from eq. 16 is symmetric and positive definite. The Cholesky decomposition was used to result,

$$C'_m = K = LU,\ L^T = U \quad (17)$$

When a vector w is randomly selected from a normal distribution N(0, 1),

$$y = Lw \quad (18)$$

and expectation of yy' is

$$E(yy') = E(Lww'U) = LE(ww')U. \quad (19)$$

Also

$$E(ww') = I, \quad (20)$$

so that

$$E(yy') = LIU = LU = K \quad (21)$$

is satisfied. And because E(y) is zero, we accomplish the multidimensional Gaussian sampling dependent on $C'_m$ from a numerous simulations of vector y.

## MODEL TEST

### A Resistivity Model Prepared by Geostatistics

Figure 2 shows the resistivity model prepared for

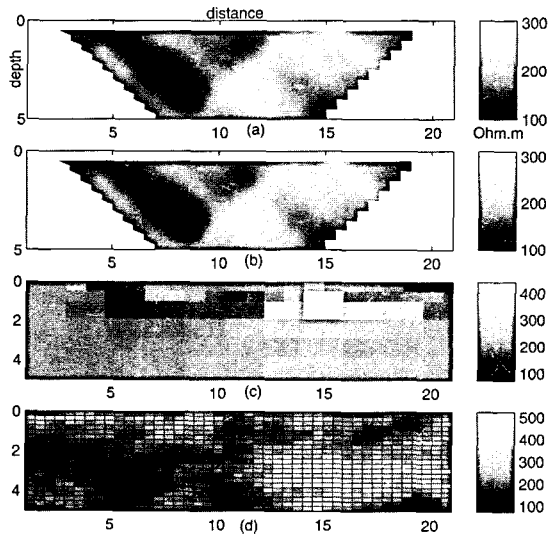


**Fig. 2.** Proposed model for this study by geostatistical simulation. (a) Observed dipole-dipole resistivity data contaminated by 10% Gaussian noise to the forward response of model (d). (b) Calculated response from the inverted model (c) by observed data-fit alone. The depth and distance are arbitrary model unit.

testing our algorithm. Observed apparent resistivity pseudosection and the calculated one from simple inversion of the model were also plotted. The observed data were contaminated by 10% Gaussian noise to the original model values.

We assumed that the real state of the underground is very continuous and has smaller scale variations than we proposed for inversion block as in Fig. 3. So we allocated simple blocks with some high and low resistivities, then applied the geostatsitcal simulation to get a more plausible model to reflect real geology and to have finer scale block variations.

Figure 4 is configuration map where Schlumberger

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | ... |
| 23 | 24 | 25 | 26 | 26 | .. | | | | | | | | | | |
| 45 | | 46 | | 47 | | 48 | | 49 | | ... | | | | | |
| 56 | | 57 | | 58 | | ... | | | | | | | | | |
| 67 | | 68 | | 69 | | ... | | | | | | | | | |
| 78 | | | | 79 | | | | ... | | | | | | | |

**Fig. 3.** Block parameterization prepared for resistivity inversion. Block numbers are displayed for following interpretations.

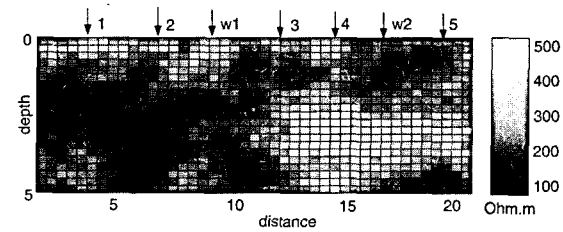


**Fig. 4.** Configuration map where the well logs and Schlumberger sounding simulation is peformed. The prefix w means the position of well logging simulations and the other is Schlumberger sounding positions.

sounding and resistivity well logging data are modeled. Sounding data are inverted to present true resistivities (Fig. 5) by the hybrid of global optimization technique SA and local least-square method (Chunduru *et al.*, 1996). Resistivity well logs (Fig. 6) are synthesized by five layers running average procedure. In addition, 5% Gaussian noise that is exponentially correlated is added to the average.

### A Prior Information and Covariance Matrix

We produced *a prior* information by geostatistical simulation from Schlumberger and well logging data in the way proposed in section 2. The schematic process to describe it is shown in Fig. 7. Indicator covariogram modelings are performed for each threshold with resistivity sounding and well logging data, then on unsampled points where the density of estimation is finer than the parameterized blocks, sequential indicator joint simulation is conducted. Many realizations were provided for assessing the uncertainty of the studied region. As described in section 2.4, the simulated estimations now are used in the block CDF making process. Figure 8 shows the PDFs for the upper blocks for parameterized region for *a prior* information of resistivity inversion. The block numbers in Fig. 3 start left-upper corner and increase to right corner in the same row until last block is reached, then continue on the left of next row. Expected, the blocks far from the sounding and logging data had widespread PDF structure that means large uncertainty and vice versa. So we may conclude that this ap-

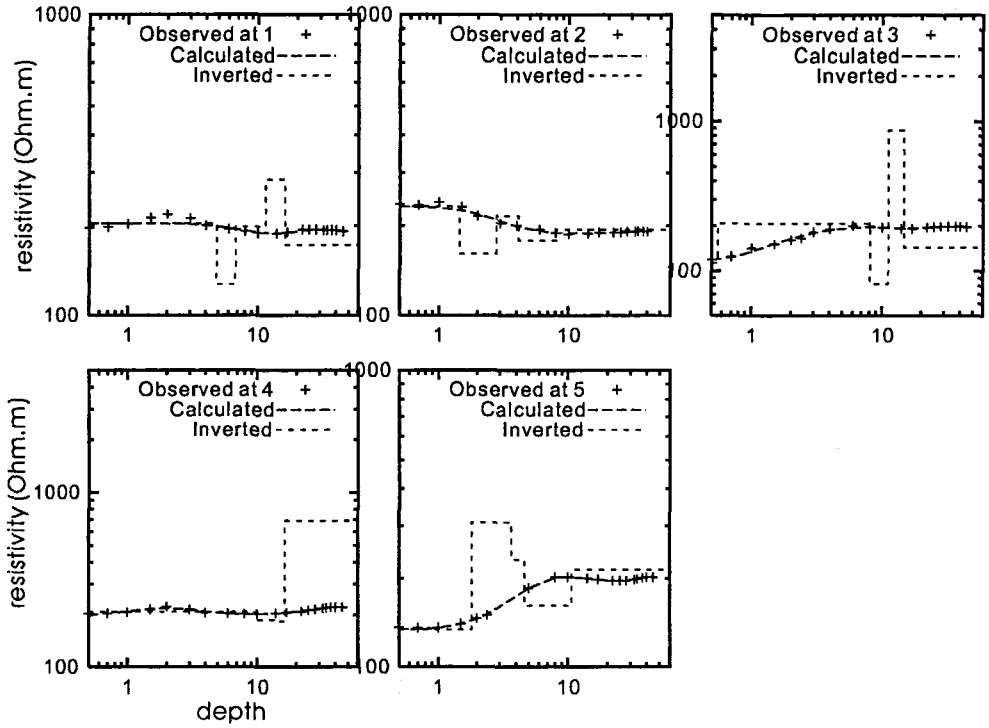


**Fig. 5.** Inversion of Schlumberger sounding data in fig. 4. SA and local least-square hybrid is applied.

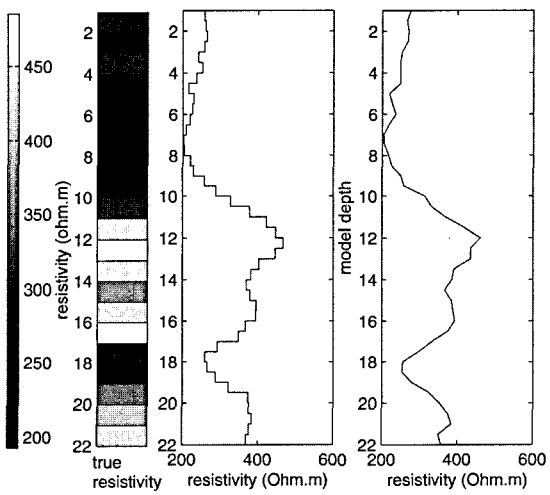


**Fig. 6.** Simulated resistivity well logging data at point w2 in fig. 4.

proach reflects well the information given by partial geologic and geophysical data.

### Maximum a Posterior Model and Marginal PDF

Figure 9 is the maximum *a posterior* model for gravity data with *a prior* information by resistivity inversion data. Compared to data-alone fit (Fig. 2), it enhances the high resistivity zone at lower-right region, and fits the observed data to a almost same degree. In this process, we didn't impose any constraints and it converged with stability. Meanwhile, *a posterior* information is sampled from *a posterior* covariance matrix as the way described in section 3.3. Figure 10 is the comparison between *a prior* and *a posterior* information for each parameter by PDF. It can be easily noticed that *a posterior* ones have less uncertainty that the *a prior* ones, and from the marginal data, we may tackle every kind of uncertainty analyses such as those proposed by Sambridge (1999).

## APPLICATION TO FIELD DATA: DETECTION OF LAVA TUNNEL

Kwon *et al.* (1998) conducted multiple geophysical explorations to investigate the lava tunnel, Man-

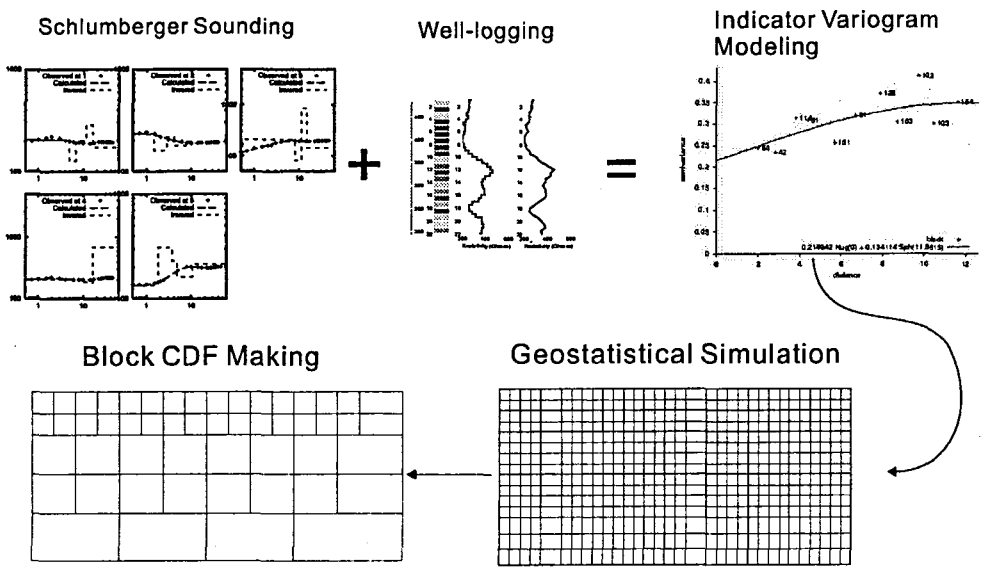


**Fig. 7.** Schematic process to generate *a prior* information. Schlumberger sounding and resistivity well logging data are put in covariogram modeling for a specified threshold for indicator kriging or simulations. Then denser estimations contribute to the block CDF making.

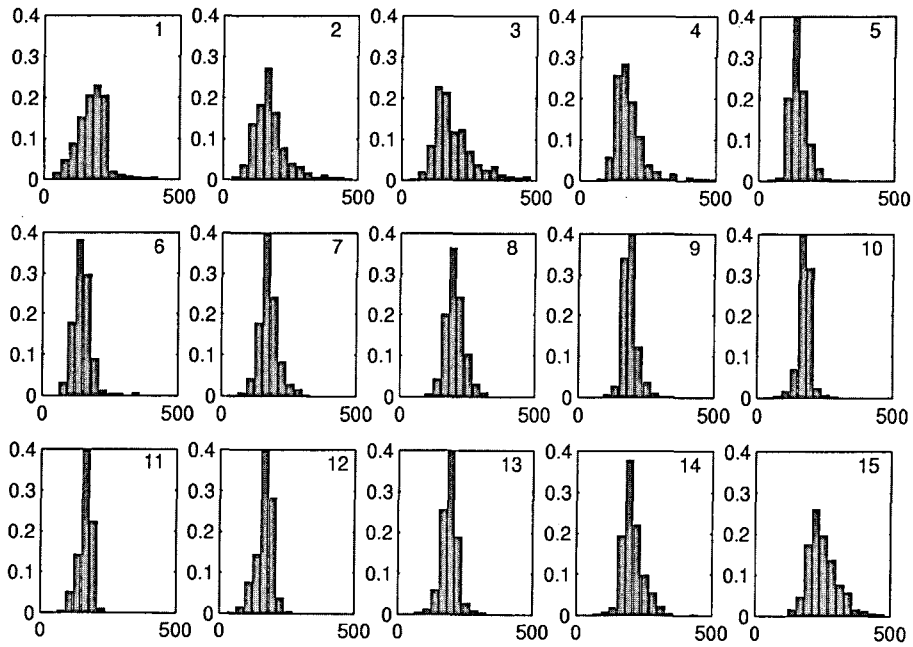


**Fig. 8.** *A prior* information generated by geostatistical simulation from Schlumberger and well logs data. The block numbers are defined in fig. 3. The horizontal axes mean resistivity (Ω-m), and the vertical axes present the probability.

janggul, in Cheju Island, Korea. They applied various geophysical methods such as gravity, magnetic, GPR, multi-frequency EM, AMT and DC dipole-dipole explorations. To test our algorithm, we chose two sets of data, gravity and DC dipole-dipole to map out the location of the lava tunnel. At the right above the cavity, gravity profile should show low density anomaly and DCs do high resistivity zone

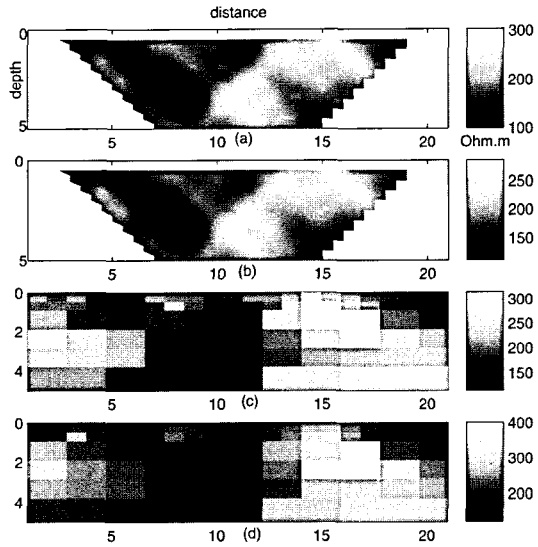


**Fig. 9.** Maximum *a posterior* (MAP) map. (a) Observed pseudosection given by fig. 2(d) model. (b) Calculated pseudosection given by bayesian inversion model (c). (d) *A prior* model map generated from dipole-dipole resistivity data.

based on the cavity modeling algorithm, and this correlation provides Bayesian inversion of the two data sets for high probability to plausible interpretation. In this study, we prepared *a prior* information from DC dipole-dipole inversion result by using

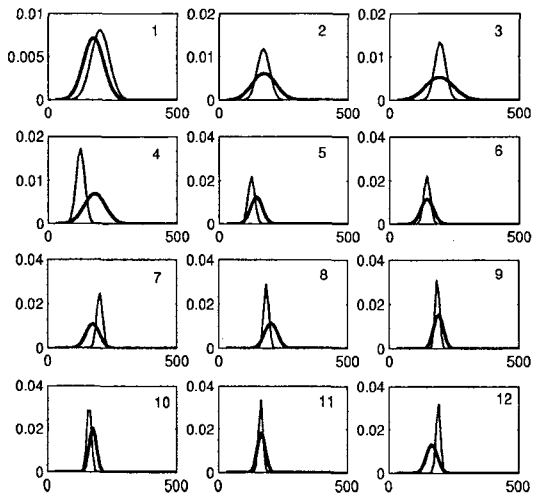

**Fig. 10.** Comparison of the probability density functions between *a prior* (bold line) and a posterior (thin line) PDF for 12 upper blocks. The *a posterior* PDF is concentrated on specific region representing low uncertainties.

the previous geostatistical approach, and gravity data was used as observed ones to reduce the uncertainty.

### Exploration Introduction

Figure 11 shows the map of the area where gravity and DC surveys were carried out. In this study the profile B-B' line was interpreted by Bayesian analysis. The depth from surface to the ceiling of the tunnel is about 20 ~ 25 m, height and width of the tunnel is respectively 5 ~ 10 m, and horizontal position is located on center about ±10 m error in this profile. Actually, we know roughly the geometric configuration of the tunnel but our assumption is we don't know it and try to apply the following method to other suspicious regions. The LaCoste-Romberg gravimeter that has measurement ranges 5000 mgal, sensitivity 0.001 mgal, and degree of accuracy 0.01 mgal was used.

### Gravity Exploration

Figure 12 shows the result of interpretation for gravity profile B-B'. Non-linear inversion section given by data-fit alone result was also plotted. They shows the negative anomalous zone near the x = 20 m, depth = 20 ~ 30 m, but we have no confidence for the result because the gravity inversion is unstable and there are no other supporting data.

### DC Dipole-dipole Exploration

Cavity modeling. Before analyzing the resistivity data from the profile B-B', we examined the behavior of resistivity sections when there exists a cavity. We modified the algorithm presented by Dey and Morrison (1979) to apply to the cavity modeling. Figure 13 shows a cavity located in a substructure which resistivity modeling is being conducted. Then following condition describe the left corner of the cavity block:

$$\text{VLC: } C_L^{ij}\phi_{i-1,j} + C_T^{ij}\phi_{i,j-1} + C_R^{ij}\phi_{i+1,j} + C_B^{ij}\phi_{i,j+1} + C_P^{ij}\phi_{i,j} = \frac{1}{2}\delta(x_s)\delta(z_s)$$

where,

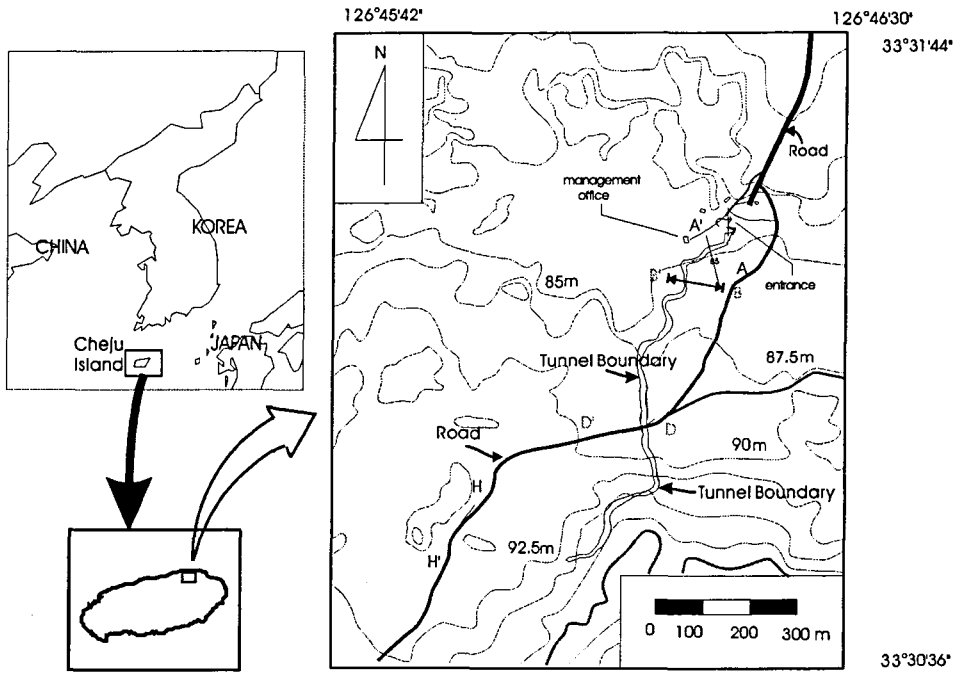


**Fig. 11.** Configuration map of gravity and dipole-dipole explorations. Gravity-DC Bayesian analysis was applied to the profile B-B'. The lava tunnel is thought to continue to the lower-left of the map, but the survey of tunnel interior was ended at the point around the profile H-H'. Contour interval for displaying height is 2.5 m.

$$C_L^{ij} = -\left[\frac{\Delta z_{j-1}\sigma_{i-1,j-1} + \Delta z_j \sigma_{i-1,j}}{2\Delta x_{i-1}}\right]$$

$$C_R^{ij} = -\left[\frac{\Delta z_{j-1}\sigma_{i,j-1}}{2\Delta x_{i-1}}\right]$$

$$C_T^{ij} = -\left[\frac{\Delta x_{i-1}\sigma_{i-1,j-1} + \Delta x_j \sigma_{i,j-1}}{2\Delta z_{j-1}}\right]$$

$$C_B^{ij} = -\left[\frac{\Delta x_{i-1}\sigma_{i-1,j}}{2\Delta z_j}\right]$$

$$A = K_y^2\left[\frac{\sigma_{i-1,j-1}\Delta x_{i-1}\Delta z_{j-1}}{4} + \frac{\sigma_{i,j-1}\Delta x_i \Delta z_{j-1}}{4} + \frac{\sigma_{i-1,j}\Delta x_{i-1}\Delta z_j}{4}\right]$$

$$C_P^{ij} = [C_L^{ij} + C_R^{ij} + C_T^{ij} + C_B^{ij} + A]$$

where, VLC is representative of the Void Left Corner boundary condition, and the above notation is same with those of Dey and Morrison (1979).

As expected, the result showed high resistivity at the zone where the cavity located. Figure 14 shows

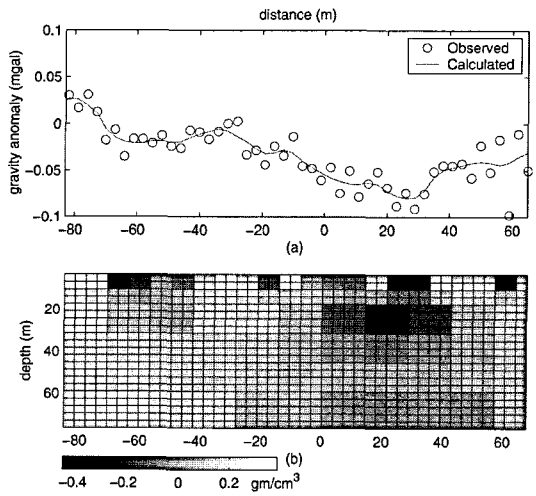


**Fig. 12.** 2-D gravity inversion result for B-B' profile (Fig. 11) showing low density cells around the 20 ~ 30 m depth. The unit of scale-bar at (b) is the gravity anomaly in mgal.

the response of the cavity modeling where the surround resistivity is 3,000 Ω-m and then shows about over 8,000 Ω-m around the cavity in the inversion

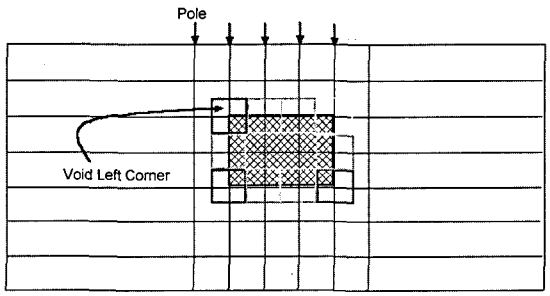


**Fig. 13.** Schematic map for displaying the boundary condition of cavity model.

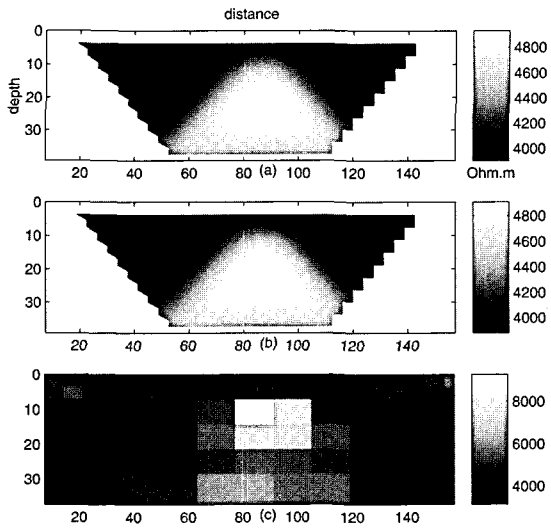


**Fig. 14.** 2-D dipole-dipole resistivity response for cavity model. (a) observed pseudosection (b) Calculated pseudosection from inverted model (c).

section with noise free state. Thus we can conclude that the both methods well delineate the cavity.

DC dipole-dipole data. Figure 15 shows the observed DC data and data-fit alone inversion result of profile B-B'. Seen from the observed pseudosection image with discontinuous pattern, it is apparent that the data are contaminated by various noise sources. The inverted section roughly delineates the high resistivity zones around 10 ~ 20 m depths with some variation. This depth is shallower than we have expected by about 20 ~ 25 m. Considered the circumstances of the lava tunnel generation, we thought there are many small-scale cavities around

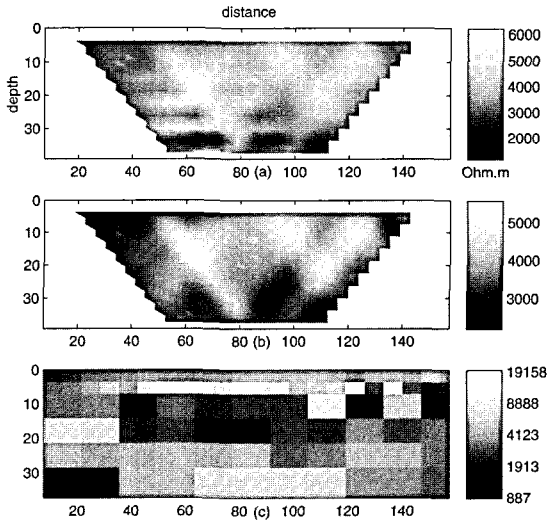


**Fig. 15.** 2-D DC result at B-B' profile (Fig. 11). (a) Observed pseudosection. (b) Calculated section from the data-fit alone inverted resistivity section (c).

and above the main tunnel, so that a high anomalous zone may be located shallow and a heavy noise response appeared. Thus complex and noise dependent situation makes the inversion result very suspicious and hard to interpret. So we apply DC inversion result to gravity Bayesian inversion to reduce the uncertainty and to get more solid evidence.

## Bayesian Inversion of Gravity and DC Data

*A prior* information. As Fig. 15, we postulate the anomalous zone due to cavities for displaying high resistivity areas, so we bring in categorical geostatistical simulation technique to present *a prior* information from the DC inversion result. That is, we conducted variogram modeling with the assumption that the center of the block represents the inverted resistivity value, then performed indicator transform with category. The categorization was made based on whether the resistivity values are higher than 8,000 Ω-m or not. The threshold value of 8,000 Ω-m was chosen by examining the previous resistivity modeling results in the inverted section, on which cavity effect was appeared to be about 8,000 Ω-m contrary to the surround observed val-

ues of about 3,000 Ω-m.

So the points having resistivity value over 8,000 Ω-m were indicated by -3, meaning the cavity, and then the other points was defined by 0. This simulation mechanism was conducted by GSLIB software (Deutsch and Journel, 1992). One hundred times of simulations was accomplished and we got *a prior* information from the DC inversion result for the Bayesian gravity inversion for detection of cavity.

Bayesian inversion and uncertainty analysis. We adopt the Gaussian approximated approach (section 3) to interpret the lava tunnel data. The inversion scheme and iteration stop criterion are same with those of section 4. Figure 16 is the final result for the Bayesian inversion of the gravity and DC data combined for detection of the lava tunnel. As we can see, the negative density anomalous zone is enhanced compared to that of Fig. 16. Also Fig. 17 shows the comparison of the *a prior* and *a posterior* information for some blocks. But unlike Fig. 11, it doesn't indicate significant update of reduction

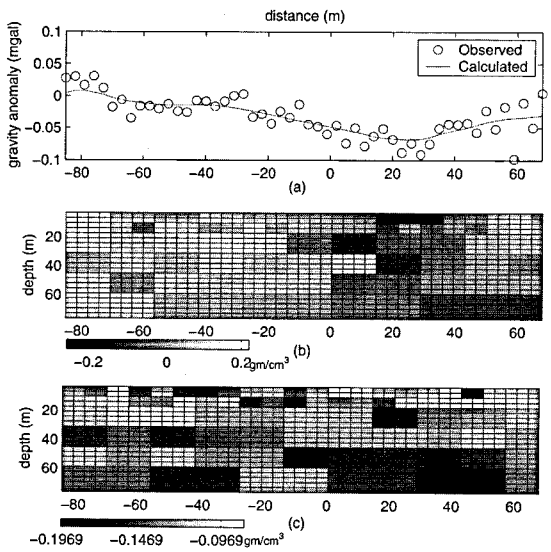


**Fig. 16.** Bayesian inversion result for B-B' profile with *a prior* model from cavity resistivity inversion simulation. It shows (a) Observed and (b) calculated response of gravity and (c) Bayesian inversion result of maximum *a posterior* map with *a prior* gravity model. The unit of scale-bars at (b), (c) is the gravity anomaly in mgal.

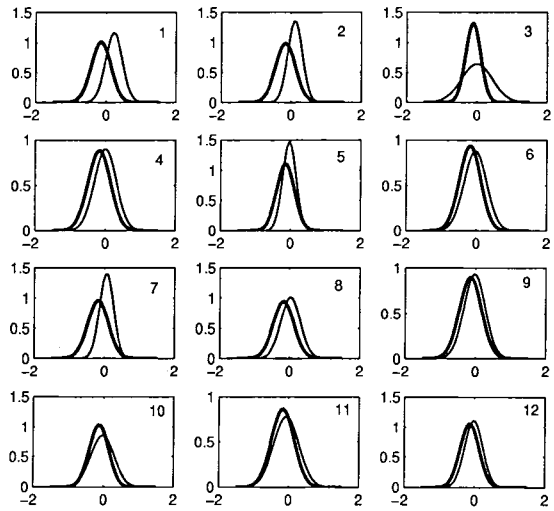

**Fig. 17.** Comparison of the probability density functions between *a prior* (bold line) and a posterior (line) PDF for 12 blocks. The *a posterior* PDF doesn't show drastic differences in uncertainties for some parameter that means low information in gravity observed data at each blocks.

of uncertainty, and we thought it was reasoned from the natural difference of resolution between the two exploration methods. However with this exotic combination, we may have confidence to point out the location of the cavities with geophysical surveys.

## CONCLUSION

At a glance, combined interpretation of gravity and resistivity data is thought to be unfamiliar. But qualitatively they showed the related responses for the underground cavity, and we used Bayesian frame to interpret it quantitatively.

In Bayesian inversion, objective and practical implementation of *a prior* information is very important element in geophysical inversion. So we, at first, present a mechanism to produce *a prior* information by geostatistical simulations. It enables us to achieve various uncertainty analyses and gives full information by a non-parametric approach. Also its great ability to incorporate much information from various sources provides us for manifold potential applications to geophysical data processing.

Next, we adopt Gaussian approximation approaches

to examine *a posterior* PDF. It provides us with a fast and effective mechanism to inspect *a posterior* PDF by local optimization technique. Meanwhile, we readopt geostatistical tools, cross-validation to generate covariance matrix in data domain in a practical and objective way. Then simulations around the maximum *a posterior* PDF were completed and applied to approximated uncertainty analysis.

Our algorithm was successful to interpret gravity and DC dipole-dipole data conducted for the detection of the lava tunnel in Cheju Island, Korea. By Gaussian approximated Bayesian frame, we could successfully delineate the cavity zone and performed uncertainty analysis.

As we have experienced in traditional inversion schemes, data-fit criteria alone produced an implausible model such as sudden fluctuation of parameters although the model validates the observed data, for we are intending to extract more information than is contained in the data. Therefore approaches to the geophysical inverse problem by Bayesian paradigm may provide more robust data interpretation and the methodology to incorporate various information to the inverse problem. It is also stable enough to estimate feasible solutions, and we presented various practical tools to implement this frame.

## ACKNOWLEDGMENT

The work was financially sponsored by Korea Energy Management Coorporation and METRI/KMA. This study was supported by the project "The development of protection techniques against earthquake disaster" of METRI/KMA. We are grateful to anonymous reviewers for their helpful and critical reviews, which allowed us to improve this manuscript significantly.

## REFERENCES

Alabert, F., 1987, The practice of fast conditional simulations through the LU decomposition of the covariance matrix. Mathematical Geology, 19, 369–386.

Bayes, T., 1763, An essay towards solving a problem in the doctrine of chances. reprinted in Biometrika, 45 (1958), 293–315.

Broyden, C.G., 1965, A class of method for solving nonlinear simultaneous equations. Mathematical Computation, 19, 577–593.

Constable, S.C. and Parker, R.L., 1987, Occam's inversion : A practical algorithm for generating smooth models from electromagnetic sounding data. Geophysics, 52, 289–300.

Chunduru, R.K., Sen, M.K., and Stoffa, P.L., 1996, 2-D resistivity inversion using spline parameterization and simulated annealing. Geophysics, 61, 151–161.

Davis, M.W., 1987, Production of conditional simulations via the LU triangular decomposition of the covariance matrix. Methematical Geology, 19, 91–98.

Deutsch, C.V. and Journel, A.G., 1992, GSLIB : Geostatistical software library and user's guide. Oxford Universty Press, New York.

Dey, A. and Morrison, H.F., 1979, Resistivity modelling for arbitrary shaped two-dimensional resistivity structure. Geophysical Prospecting, 27, 106–136.

Duijndam, A.J.W., 1988a, Bayesian estimation in seismic inversion, part i : Principles. Geophysical Prospecting, 36, 878–898.

Duijndam, A.J.W., 1988b, Bayesian estimation in seismic inversion, part ii : Uncertainty analysis. Geophysical Prospecting, 36, 899–918.

Glacken, I., 1996, Change of Support by direct conditional block simulation. Master's thesis, Stanford Univ., Stanford, CA.

Gomez-Hernandez, J., 1991, A stochastic approach to the simulation of block conductivity fields conditional upon data measured at a smaller scale. Doctoral Dissertation, Stanford University, Stanford, CA.

Goovaerts, P., 1997, Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Gouveia, W.P., 1996, Bayesian seismic waveform data inversion : Parameter estimation and uncertainty analysis. Ph. D. thesis, Colo. Sch. of Mines, Golden.

Isaaks E.H. and Srivastava, R.M., 1989, An Introduction to Applied Geostatistics. Oxforf University Press, New York.

Isaaks, E.H., 1990, The application of Monte Carlo Methods to the Analysis of Spatially Correlated Data. Doctoral Dissertation, Stanford University, Stanford, CA.

Jaynes, E.T., 1994, Probability theory : the logic of science.

Journel A.,G. and Huijbregts, C.J., 1978, Mining Geostatistics. Academic Press, New York.

Jung, H.J., 1998, Fast two-dimensional magnetotelluric incersion using approximate update of sensitivity in wavelet domain. Ph. D. Thesis, Seoul Nat'l University, Korea.

Kwon, B.D., Lee, S.J., Oh, S.H., and Lee, C.K., 1998, Application of multiple geophysical methods in investigating the lava tunnel of Manjanggul in Cheju island. Economic and Environmental Geology, 31. 535–545.

Lanczos, C., 1961, Linear differential operator. Von Nostrans Company, Princeton, New Jersey.

Lines, L.R. and Treital, S., 1984, Tutorial : A Review of Least-Squares Inversion and Its Application to Geophysical Problems. Geophysical Prospecting, 32, 159–186.

Loredo, T.J., 1990, From Laplace to Supernova SN 1987A : Bayesian Inference in Astrophysics. Reprinted from: P.F. Fougere (ed.), Maximum Entropy and Bayesian Methods 81–142, Kluwer Academic Publishers, The Netherland.

Moraes, F.S., 1996, The application of marginalization and local distributions to multidimensional Bayesian inverse problems, Ph. D. Thesis, Colorado School of Mines, Golden, Co.

Mosegaard, K. and Tarantola, A., 1995, Monte Carlo sampling of solutions to inverse problem. Journal of Geophysical Research, 100, 12,431–12,447.

Sambridge, M., 1999, Geophysical inversion with a neighbourhood algorithm II. Appraising the ensemble. Geophysical Journal International, in press.

Scales, J.A. and Tarantola, A., 1994, An example of geologic prior information in a Bayesian seismic inverse calculation. Center for Wave Phenomena 159, Golden, Co.

Tarantola, A., 1987, Inverse problem theory : Methods for data fitting and model parameter estimation. Elsevier.

Manuscript received October 8, 2001
Revised manuscript received November 1, 2001
Manuscript accepted November 20, 2001