

지구물리 자료의 고속 베이지안 역산

오석훈* · 권병두** · 남재철* · 이덕기*

*기상연구소 해양기상지진연구실 · **서울대학교 지구과학교육과

요 약

베이지안 역산(Bayesian inversion)은 불충분한 자료를 가지고 지하구조를 추정해야 하는 지구물리 자료의 해석에 있어서 안정적이고 신뢰를 줄 수 있는 방법 중의 하나이다. 관측 자료가 측정 과정부터 불확실성을 함유하고 있으며, 역산에 이용되는 이론 자료 또한 모델의 매개변수화에 따른 각종 불확실성을 포함하고 있다. 따라서 지구물리 자료의 역산은 확률적으로 접근하는 것이 가장 바람직하며 베이지안 역산은 이에 대한 처리뿐만 아니라, 추정에 대한 신뢰도와 불확실성에 대한 이론적 근거를 제공한다. 그러나 대부분의 베이지안 역산이 고차원의 적분을 필요로 하므로 몬테 카를로 방법과 같은 대규모의 계산이 요구되는 방법에 의해 사후 확률분포가 구해지는 경우가 많다. 이는 특히 지구물리 자료와 같이 고도의 비선형 자료에 대하여 매우 적합한 접근 방법이기 는 하지만, 점차 현장화, 고속화 되어가는 자료의 해석 경향에 맞추어 간략하게 사후 확률분포를 근사할 수 있는 기법의 연구 또한 필요하다. 따라서 이 연구에서는 관측자료와 사전 확률분포가 정규분포에 의해 근사 될 수 있는 지구물리 자료에 대한 베이지안 역산에 대해 논의 하고자 한다. 사전 확률분포의 작성을 위해 지구통계학적 기법이 이용되었으며, 관측자료의 통계적 불확실성을 추정하기 위해 교차 검사(cross-validation) 방법을 이용하여 공분산(covariance)을 유도하고 그것에 의한 우도 함수(likelihood function)를 작성하였다. 베이지안 해석을 위해 두 확률분포를 곱하여 근사적인 사후 확률분포를 얻을 수 있었으며, 이에 대해 최적화(optimization) 기법을 이용하여 최대 사후 확률(Maximum a Posterior)을 따르는 지하 구조를 얻을 수 있었다. 또한 사후 확률 분포의 공분산 항을 이용하여 지하 비저항 구조를 시뮬레이션 하여 불확실성 분석을 수행하였다.

주요어: 베이지안 역산, 지구통계학, 사후 확률분포

Seok Hoon Oh, Byung Doo Kwon, Jae Cheol Nam and Duk Kee Lee, 2000, Fast Bayesian Inversion of Geophysical Data, Journal of the Korean Geophysical Society, v. 3, n. 3, p. 161-174

ABSTRACT: Bayesian inversion is a stable approach to infer the subsurface structure with the limited data from geophysical explorations. In geophysical inverse process, due to the finite and discrete characteristics of field data and modeling process, some uncertainties are inherent and therefore probabilistic approach to the geophysical inversion is required. Bayesian framework provides theoretical base for the confidency and uncertainty analysis for the inference. However, most of the Bayesian inversion require the integration process of high dimension, so massive calculations like a Monte Carlo integration is demanded to solve it. This method, though, seemed suitable to apply to the geophysical problems which have the characteristics of highly non-linearity, we are faced to meet the promptness and convenience in field process. In this study, by the Gaussian approximation for the observed data and a priori information, fast Bayesian inversion scheme is developed and applied to the model problem with electric well logging and dipole-dipole resistivity data. Each covariance matrices are induced by geostatistical method and optimization technique resulted in maximum a posteriori information. Especially a priori information is evaluated by the cross-validation technique. And the uncertainty analysis was performed to interpret the resistivity structure by simulation of a posteriori covariance matrix.

Key words: Bayesian inversion, geostatistics, a posteriori information

(Seok Hoon Oh, Jae Cheol Nam and Duk Kee Lee, Marine Meteorology and Earthquake Research Laboratory, Meteorological Research Institute, Dongjak-Gu, Seoul 156-720, Korea; Byung Doo Kwon, Department of Earth Science Education, Seoul National University, Seoul 151-742, Korea. email: gimul@metri.re.kr)

1. 서 론

지구물리 탐사를 이용하여 해석하고자 하는 지하구조를 계산하기 쉽게 임의로 구획하는 매개변수화(parameterization)는 순산모델(forward modeling)에 적용할 수 있는 이론식 자체가 한계가 있기 때문에 생겨나는 피할 수 없는 과정이다. 또한 우리가 얻을 수 있는 관측자료의 공간도 유한하기 때문에 실제로 무한한 요소로 이루어진 지하구조를 유한 공간에서 관측하고, 제한된 매개변수 공간으로 다시 투영하므로, 지구물리 역산은 처음부터 해의 비유일성과 불안정성이 내재되어있다. 이로 인하여, 단일 자료만에 의한 역산 해석은 완전성을 가지지 못하며, 계산 과정에서 무의미한 지하구조를 만들어 내기도 한다. 이것을 해결할 수 있는 방안으로 다양한 정규화(regularization) 기법이 연구되었다(Constable *et al.*, 1987; Parker, 1994). 정규화 과정은 지구물리 역산 해의 불안정성을 해소하기 위해, 매개변수에 대해 공간적인 제한조건을 가하는 Occam의 가정(Constable *et al.*, 1987)이 많이 이용되었다. 그러나 이 방법은 실제 존재하지 않는 제한조건을 가함으로써, 해의 안정성을 얻는 대신 정확한 구조를 얻지 못할 수도 있다는 단점을 지녔다.

이 연구에서는 단일 관측자료는 물론이고, 한 지역에 대한 지질학적, 지구물리학적 여러가지의 정보를 가지고 있을 때, 이들을 역산과정에 도입하여 객관적으로 의미있는 지하구조를 찾아내고자 하는 방법에 대해 논의하고자 한다.

임의의 지역에서 지구물리 자료를 이용하여 지하구조를 해석하는데 있어 최근의 추세는 각종 탐사방법들을 동원한 복합적인 탐사가 많이 이루어지고 있는 실정이다(권병두 외, 1999). 따라서 복합 탐사 자료에 적합한 해석 기법들이 요구되고 있으나 이에 대한 통합적이고 체계적인 연구는 의외로 많지 않다. 다시 말하면, 각기 독립적으로 해석된 자료에 의해 전체를 추정하는 방법 이외에는, 정량적으로 존재하는 자료를 통합하여 해석할 수 있는 방법에 대한 연구는 미흡한 실정이다. 이에 대한 시도로써 Vozoff · Jupp(1975)가 발표한 결합 역산(joint inversion)방법을 들 수

있다. 이것은 모델 매개변수를 공유하여 서로 다른 방식에 의한 관측자료의 민감도 행렬을 구하여 역산 문제에 적용하는 방법이다. 그러나 이 방법을 실제로 적용할 때, 반응특성이 서로 다른 자료를 단지 동일한 매개변수에 적용함으로써, 매우 불안정한 민감도 행렬이 발생하고, 측정치와 계산치의 오차도 잘 수렴하지 않으며, 바람직하지 않은 역산 결과가 많이 발생한다. 또한 다수의 탐사가 수행되었을 때, 이들을 이용한 효율적인 계산에도 어려움이 많다.

이 연구에서 추구하는 궁극적인 지구물리 역산의 방향은 관심 대상이 되는 지역에 대해 다수의 정보가 있을 때 이들 모두를 반영하여 최적의 정보를 얻어내는 것으로써 베이저안 역산(Bayesian inversion) 방법을 적용하고자 한다. 이 방법의 개념에 의하면, 현재 존재하고 있는 모든 자료를 이용하여 사전 확률분포(a priori PDF)라는 정보를 얻는다. 독립적으로 관측 자료가 얻어지면, 사전 확률 분포와 관측자료를 최적화하는 사후 확률 분포(a posteriori PDF)를 얻는다. 이후 또 다른 독립적인 자료의 최적화를 위해 직전의 사후 확률분포는 다시 사전 확률분포가 된다. 이 방법은 불완전한 자료에 대한 확률적 접근의 틀을 제공하고, 다수의 정보가 존재하는 경우 가장 효율적으로 이용할 수 있는 방안을 제시한다.

사전 확률분포가 결정론적(deterministic)으로 주어지게 되면, 관측자료의 분산분포에 의해서만 사후 확률 분포를 추정하게 되며(Stark, 1992), 반대로 사전 확률분포가 매우 복잡한 형태로 나타나면, 매개변수에 의한 적분으로 주어지는 사후 확률 분포를 계산하기 위해 몬테 카를로 적분기법과 같은 방법을 이용하여야 한다(오석훈, 2000). 또한 이 두 가지의 중간 정도로서 사전 확률분포가 정규분포를 따를 경우 비교적 용이하게 사후 확률분포를 추정할 수 있다. 이 연구에서는 베이저안 고속 역산을 위해 모든 확률 분포를 정규분포로 가정하였다.

이 연구에서는 다양한 자료를 사전 정보(a priori information)로 통합하기 위하여 지구통계학적 접근(geostatistical approach)을 시도하였다. 각기 다른 지구물리 탐사자료를 통합하기 위해서는 기존

의 모수적 확률 분포(parametric PDF)에 의한 접근은 안정적이지 못하다. 따라서 비모수적 접근(non-parametric approach)을 위해 지시자 변환(indicator transform)을 수행하여 다양한 자료가 통합되는 과정에서 불안정 요소를 제거하였다.

이상에서 개발된 베이지안 고속 역산 기법의 적용을 위해 복잡하게 나타나는 지하 전기비저항 구조의 해석에 이용하여 보았다. 슬림버저 수직 전기탐사와 전기 비저항 검층 자료를 시뮬레이션 하여 사전 확률분포를 작성하고, 쌍극자-쌍극자 전기 탐사 자료를 관측자료로 하여 역산을 수행하였다.

2. 이론 및 방법

2.1. 쌍극자-쌍극자 전기탐사 자료

Fig. 1은 이 연구에서 해석대상으로 정한 지하 구조 (d)와 이에 대한 2 차원 쌍극자-쌍극자 전기

탐사 반응(a)을 나타낸다. 전기탐사 반응에는 표준편차를 1로 주고 10%의 잡음을 추가하였다. (d)에서 볼 수 있듯이 지하구조는 역산에 적용된 블록(Fig. 2)보다 세밀하게 변화도록 구성되었다. 이는 기존의 역산 모델 연구가 미리 정해놓은 역산 블록에 맞추어서 모델 변수 값을 변화시켰던 것을 극복하고, 보다 현실에 가까운 현상을 표현하기 위한 의도이다. 차후의 검토를 위하여 기존의 Marquardt-Levenberg 방법(Lines · Tretal, 1984)에 의한 역산결과 (c)와 그에 의한 계산반응 (b)도 함께 도시하였다. 단순히 관측자료에 의한 역산결과는 (d)에 나타난 실제 모델을 잘 반영하고 있지 못함을 알 수 있다.

2.2. 관측자료의 공분산(Covariance) 모델링

관측자료와 모델 매개변수에 의한 계산 자료의 다변수 정규 확률분포는 식 (1)과 같이 주어질 수

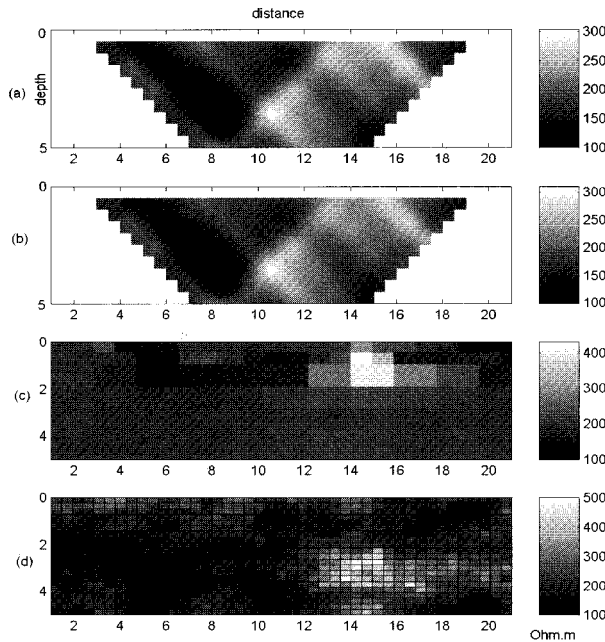


Fig. 1. Proposed model for this study by geostatistical simulation. (a) Observed dipole-dipole resistivity data contaminated by 10% Gaussian noise for the response of input model (d). (b) Calculated response for the data-fit alone inverted model (c). The depth and distance arbitrary model unit.

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

Fig. 2. Blocky parameterization for resistivity inversion. Block numbers are referenced for the further explanation.

있다(Tarantola, 1987).

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

여기서 L은 우도(likelihood)로서 모델 파라미터 m에 대한 관측자료 d의 적합 정도를 나타내며, G는 비선형 순산모델링 함수를 가리킨다. C_d^{-1} 은 관측자료에 대한 공분산 항이다.

위의 식 (1)을 최대로 만드는 것을 최대 우도 기법(maximum likelihood method) 이라 하며, 이를 위해 지수함수의 내부를 최소로 하는 방법의 일종으로 최소자승법(least-square method) 이 사용된다(Gouveia, 1996).

지금까지 대부분의 연구에서 식 (1)의 공분산의 역수 항 C_d^{-1} 은 단순히 중량(weighting) 벡터가 대각행렬을 이루는 중량 행렬(weighting matrix) W로 간단히 취급하였으며(Constable et al., 1987), 주관적으로 자료의 질이 떨어지는 지역에 낮은 값을 적용하는 방식이나, 일률적으로 똑같은 값을 주는 방식으로 적용되었다. 이것은 C_d^{-1} 를 구하는 것이 원칙적으로 불가능한 것에 그 원인이 있다.

이 연구에서는 최대한 객관적으로 C_d^{-1} 를 추정하기 위해 지구통계학적 기법 중의 하나인 교차 검사(cross-validation) 기법(Isaaks · Srivastava, 1989)을 이용하였다. 이것은 지구통계학 기법 중의 하나인 크리깅(kriging: Isaaks · Srivastava, 1989)을

수행하기 전에 주어진 자료에 대해 가장 오차를 작게하는 크리깅 변수를 찾기 위한 방법이다. 이 방법은 Fig. 3에서와 같이 크리깅 오차를 추정하고자 하는 지점의 관측 자료를 제거하고(b), 주변 관측자료에 의해 자료가 삭제된 지점의 크리깅 추정값을 구해(c) 원래 관측 값과의 차이를 통해 오차를 추정하는 방법이다(d).

이 연구에서는 관측 지점에 대한 보다 정확한 조건부 적확률분포(conditional cumulative density function)의 추정을 위해 지시자 크리깅(indicator kriging)에 의한 교차 검사를 수행하였고, 이를 통하여 평균과 분산을 구하였다.

Goovaerts(1997)나 Journel · Huijbregts(1978) 등에 의하면 다변수 정규 분포에 의한 적확률분포의 평균과 분산은 단순 크리깅(simple kriging)의 그것과 동일하다고 발표한 바 있다. 이 연구에서는 고속 베이지안 역산을 위해 모델 변수와 관측 자료의 확률분포가 정규분포를 따른다고 가정하였기 때문에 이와 같은 교차 검사 기법에 의한 공분산의 추정은 타당하다. 또한 크리깅 과정에서 주변의 관측자료들이 중량 항(weighting value)을 가지고 관심 대상의 관측지점자료의 추정에 기여하므로, 공분산 행렬의 대각 행렬 이외의 요소를 객관적으로 구성할 수 있다.

한편 이 방법은 자료 공간에 대한 평활화 제한이라 생각할 수 있다. 이는 크리깅 기법이 추정값의 분산을 최소로 하는(Goovaerts, 1997) 기법이므로

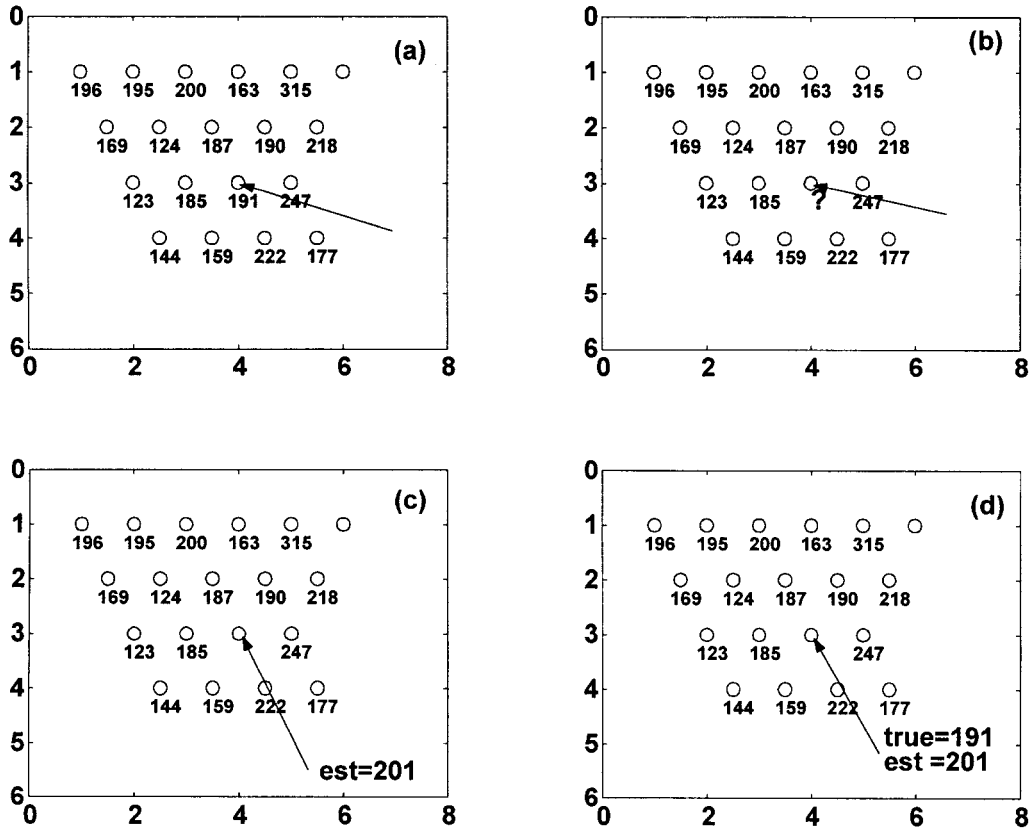


Fig. 3. Cross-validation algorithm to present covariance matrix in part of dipole-dipole resistivity data domain. The scale of each axis follows the data appeared in Fig. 1. An arbitrarily sampled point is selected(a) and eliminated(b). (c) Estimation of the eliminated point with the other sampled points by various Kriging and simulations. (d) The difference between the true value and estimation provides information of error distribution of observed data.

로, 부드러운 곡선으로 임의의 지점을 추정하기 때문이다. 이것은 모델 공간에 대해 평활화 제한을 가하는 Occam의 역산(Constable *et al.*, 1987)과 비교될 수 있다. Occam의 역산이 해의 안정적 수렴을 위한 모델에 대한 인위적 정규화 과정이라면, 이 연구에서 제안된 자료 공간에 대한 평활화 제한은 지구통계학에 기반하여 관측자료를 정규화 할 수 있는 기술로서 보다 타당하다 여겨진다.

2.3. 사전 정보(A Prior information) 자료의 공분산(covariance) 모델링

사전 정보로 주어지는 지하의 전기비저항 m_{prior}

의 다변수 정규 확률분포는 다음과 같이 주어진다(Tarantola, 1987).

$$\rho(m) \propto \exp(-m - m_{prior})^T C_m^{-1} (m - m_{prior}) \quad (2)$$

여기서 $\rho(m)$ 은 모델 파라미터 m 에 대한 사전 확률정보를 가리키며, m_{prior} 는 사전 모델 파라미터이다. C_m^{-1} 은 모델 파라미터에 대한 공분산 항이다.

$\rho(m)$ 의 유도는 지구통계학적 방법(오석훈, 2000)에 의해 수행되었다. 먼저 사전 정보를 표현하기 위해 Fig. 1과 같은 지하구조에 대해 슬럼

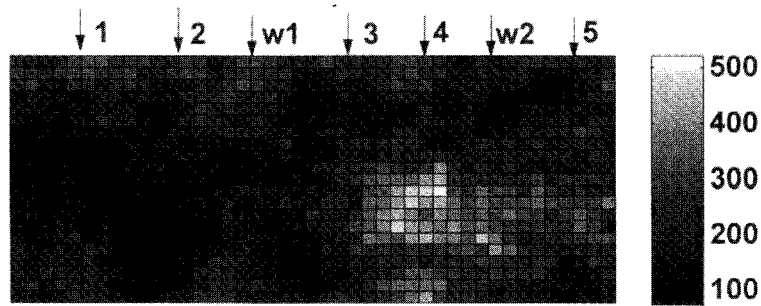


Fig. 4. Location of the sample points where the well logging and Schlumberger sounding simulation are performed, as shown by the arrows. The prefix w indicates the position of well logging simulations and the cited numbers represent Schlumberger sounding positions. The unit and scale are same with those of Fig. 1(d).

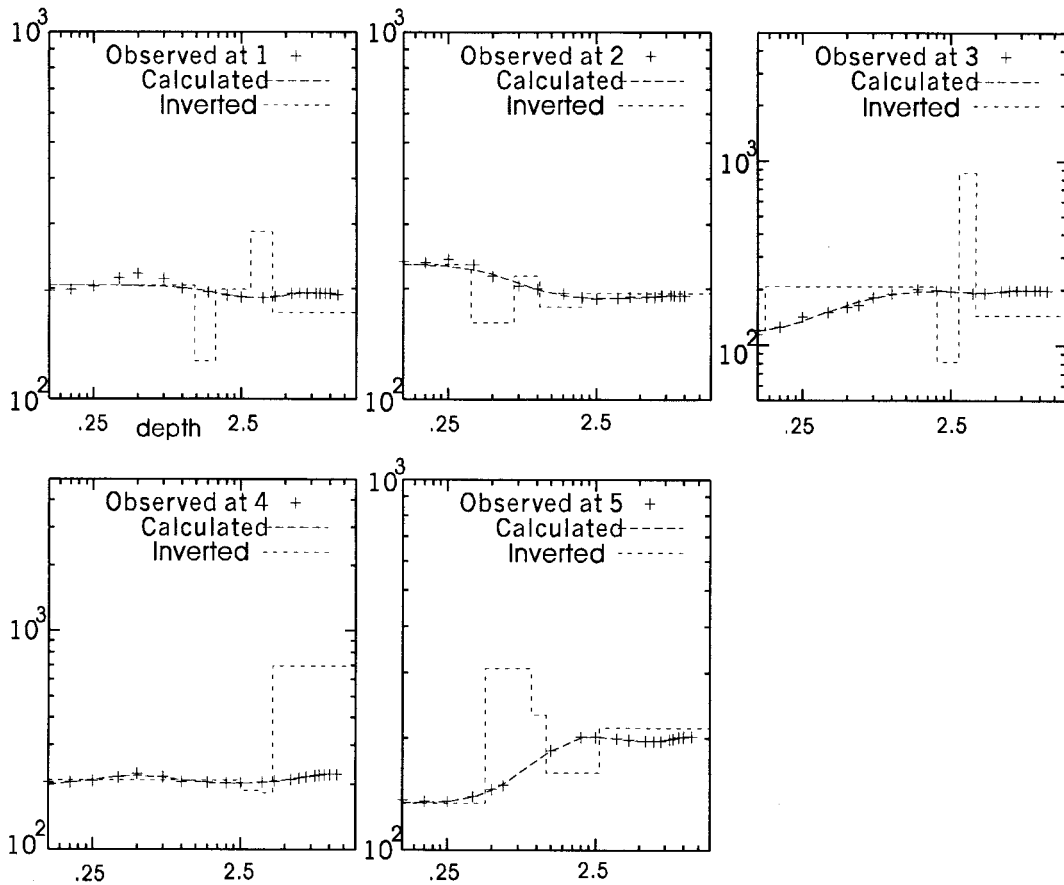


Fig. 5. Schlumberger sounding inversion result for the model of Fig. 4. SA(Simulated Annealing) and local least-square hybrid are jointly applied.

버저 수직탐사와 전기 비저항 검층 자료를 시물레이션 하였다. 슬림버저 탐사는 총 5 개 지점에서 시물레이션 되었으며, Fig. 4는 그 위치를 나타내고, Fig. 5는 반응값 및 역산 결과를 나타낸다. 슬림버저 수직 탐사자료의 역산은 Simulated Annealing에 의한 광역적 최적화 기법(오석훈과 권병두, 1997)과 기존의 최소자승법을 결합한 방법으로 수행하였다.

또한 Fig. 6은 Fig. 4에 도시한 지점에서 수행된 전기비저항 검층 자료의 시물레이션 결과를 나타낸다. 전기비저항 검층을 시물레이션 하기 위하여, Fig. 1의 전기비저항 구조에 대해 5개의 블록에 대한 윈도우 평균을 취하였고, 지수함수적인 상관관계에 의한 5% 정규분포 잡음을 추가하였다.

이 두 자료를 이용하여 사전 확률분포를 작성하기 위하여 지구통계학적 기법의 하나인 동시 크

리깅(co-kriging) 기법을 이용하였으며, 최종적으로 지시자 동시 크리깅(indicator co-kriging: 오석훈, 2000)을 이용하여 지하구조에 대해 시물레이션을 수행하고 사전 확률 분포에 의한 공분산 행렬을 작성하였다.

2.4. 베이지안 역산을 위한 사후 확률 분포

(A Posterior PDF)

사전 확률분포의 개선(update)을 위해 관측자료의 해석을 첨가하는 베이지안 역산은 식 (1)과 (2)를 곱하는 것에 의해 수행된다(Loredo, 1990; Ceman · Ceman, 1984). 즉, 사후확률분포 $\sigma(m|d)$ 는

$$\sigma(m|d) \propto \exp((G(m) - d_{obs}) C_d^{-1} (G(m) - d_{obs}) + (G(m) - d_{obs}) C_d^{-1} (G(m) - d_{obs})) \quad (3)$$

에 의해 주어진다(Tarantola, 1987).

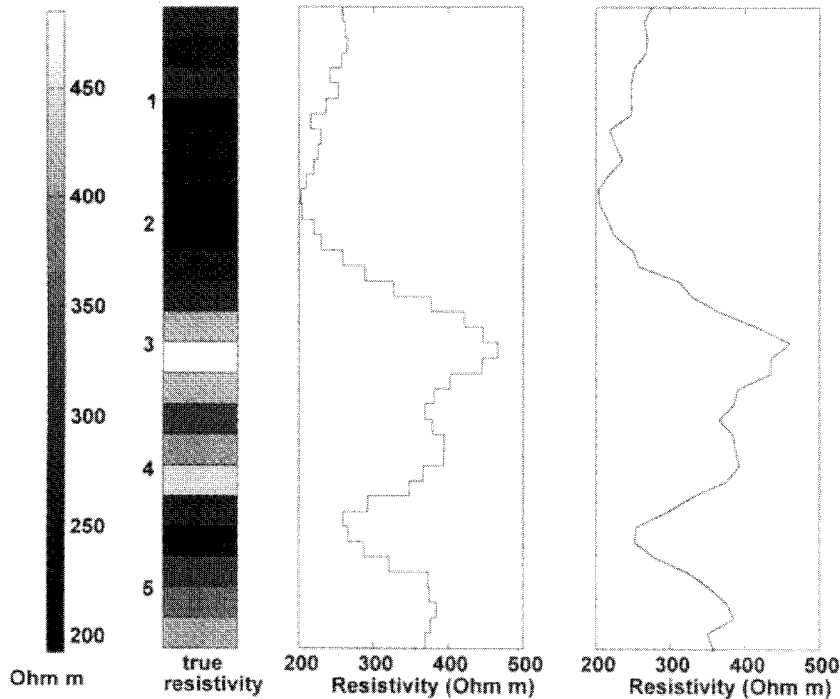


Fig. 6. Simulated resistivity well logging data at the position of w2 in Fig. 4. The left stripe shows the true resistivity structure and the log in the middle is generated by five-point averaging with the resistivity values read at every 0.125 unit. The well log on the right is simulated by adding correlated Gaussian noise.

기존의 베이지안 해석의 어려움은 식 (3)으로 주어지는 사후 확률분포의 추정에서 시작된다. 실제 다차원 문제에서 매개변수 벡터 m 의 성질을 규명하기 위해서는 식 (3)을 관심 대상이 되는 성분 m_j 를 제외한 나머지 성분들로 모두 적분하여야 m_j 에 대한 주변 확률분포(Marginal probability distribution)를 구할 수 있다(Sambridge, 1999). 이를 위해 Mosegaard · Tarantola (1995)나 오석훈(2000)은 몬테카를로 적분을 이용하였으나, 계산량이 방대하므로 고속으로 자료를 처리하는데는 어려움이 있다. 이에 대조적으로 Mora(1987)와 Crase 등(1990)은 공분산 행렬 C_d, C_m 을 추정하지 않고 단순히 식 (3)을 최대로 만드는 모델 벡터를 구하였다. 이는 베이지안 역산의 장점인 불확실성 해석을 이용할 수 없으므로 무의미하다. 이 연구에서는 이미 앞 절에서 공분산 행렬 C_d, C_m 를 추정하는 방법에 대해 논의하였고, 이를 이용하여 불확실성 해석하였다.

식 (3)의 사후 확률 분포를 최대로 만드는 벡터 m 을 얻기 위하여, m 에 대한 미분을 수행하면,

$$\frac{\partial S}{\partial m} = g = 2A^T C_d^{-1}(G(m) - d_{obs}) + 2C_m^{-1}(m - m_{prior}) = 0 \quad (4)$$

을 만족하여야 한다. 여기서 A 는 순산 모델링 행렬 G 의 민감도 행렬(sensitivity matrix)이며, 이제 j 번째 반복에 의하여,

$$G(m_{j+1}) = G(m_j) + A\Delta m + O(n)^2 \quad (5)$$

$$\text{where, } m_{j+1} = m_j + \Delta m$$

이라 두고, 2 차 이상의 미분항을 제거하면,

$$(A^T C_d^{-1} A + C_m^{-1})\Delta m_j = -(A^T C_d^{-1}(G(m_j) - d_{obs}) + C_m^{-1}(m_j - m_{prior})) \quad (6)$$

로 쓸 수 있다. 이제 $(A^T C_d^{-1} A + C_m^{-1}) = H$ 라 두면 식 (6)은

$$H\Delta m_j = -g_j \quad (7)$$

라 둘 수 있고, Δm_j 를 계산하기 위해 특이치 분해(SVD; singular value decomposition; Lanczos, 1961)와 Marquardt-Levenberg 제한(Lines · Treital, 1984)을 가하여 최적화를 수행할 수 있다.

또한 이 연구에서는 관측자료의 민감도 행렬의 계산에 Broyden 개선(Broyden, 1965; 정호준, 1998)을 이용하였다. 즉, 최초의 민감도 행렬 A 는 상반성 원리(reciprocity theorem; 김정호, 1987)를 이용하여 구하고, 이후의 갱신은 Broyden 개선에 의해 신속하게 구할 수 있었다.

관측치의 평균으로 초기 모델 벡터 m 을 주고 식 (7)에 의해 반복을 수행하여, 관측자료의 모델 반응에 대한 오차가 표준편차의 크기와 동일한 수준으로 낮아질 때까지 반복을 계속한다. 즉,

$$(G(m) - d_{obs})^T C_d^{-1} (G(m) - d_{obs}) \leq N \quad (8)$$

수준까지 식 (7)을 반복한다. 여기서 N 은 총 관측 자료의 개수이다.

3. 결과 및 분석

3.1. 베이지안 역산 결과

Fig. 7은 이상의 알고리즘을 적용하여 2 차원 전기비저항 베이지안 역산을 수행하여 사후 확률분포를 최대화하는 전기비저항 단면과 사전 정보에 의해 주어진 단면을 함께 도시한 것이다. Fig. 1의 단순 역산 자료와 비교할 때 깊이 2 이하의 지역에서 상대적으로 고비저항대가 다수 위치하고 있으며, 이는 사전 비저항 정보가 많이 반영된 것으로 볼 수 있다.

이제 불확실성 해석을 위해 식 (3)을 다시 정리하면,

$$\sigma_g(m) \propto [(m - m_{map})^T C_m^{-1} (m - m_{map})] \quad (11)$$

로 쓸 수 있다(1987, Tarantola). 여기서 C_m^{-1} 은 사후 확률분포의 공분산 행렬, m_{map} 은 최대 사후 확률분포 값으로 위의 베이지안 역산에서 구한 모델 벡터를 이용할 수 있다. Tarantola(1987)에 의하면, 식 (3)과 같은 비선형 모델자료에 대한

식 (11)의 정규 확률분포의 공분산 C_m 는

$$C_m = [A^T C_d^{-1} A + C_m^{-1}]^{-1} \quad (12)$$

와 같이 구할 수 있다.

이를 이용하면 Fig. 7의 최대 사후 확률분포 모델에 대해 비교적 간단히 신뢰구간을 추정할 수 있다. Fig. 8은 Fig. 7의 역산 결과에 대해 식 (12)의 사후 공분산에 의해 표준편차 만큼의 오차를 표현한 것이다. 이는 식 (12)의 행렬에서 대각 성분(분산)만을 이용하여 구할 수 있다.

그러나 이는 식 (12)로 주어지는 사후 공분산 행렬의 일부 정보만을 이용하는 것이므로 보다 일반적인 해석이 필요하다. 그래서 이 연구에서는 평균 m_{map} 공분산 C_m 로 주어지는 정규확률분

포에 의한 샘플링을 수행함으로써 완전한 불확실성 해석을 수행하였다.

3.2. 사후 확률분포에 의한 불확실성(uncertainty) 해석
식 (3)으로 주어지는 사후 확률 분포에 의해 각각의 모델 벡터를 성분 별로 분석하기 위해서는

$$m_j = \int \sigma(m|d) dm$$

을 계산하여야 하며, 이를 주변 확률분포(marginal probability distribution)라 한다(Gilks *et al.*, 1996). 여기서 m 벡터는 m_j 를 제외한 나머지 성분들이다. 이를 위해 몬테 카를로 방법 등이 적용되었지만(Mosegaard · Tarantola, 1995; 오석훈, 2000), 이 연구에서는 모든 확률분포가 정규분포를 따른다고

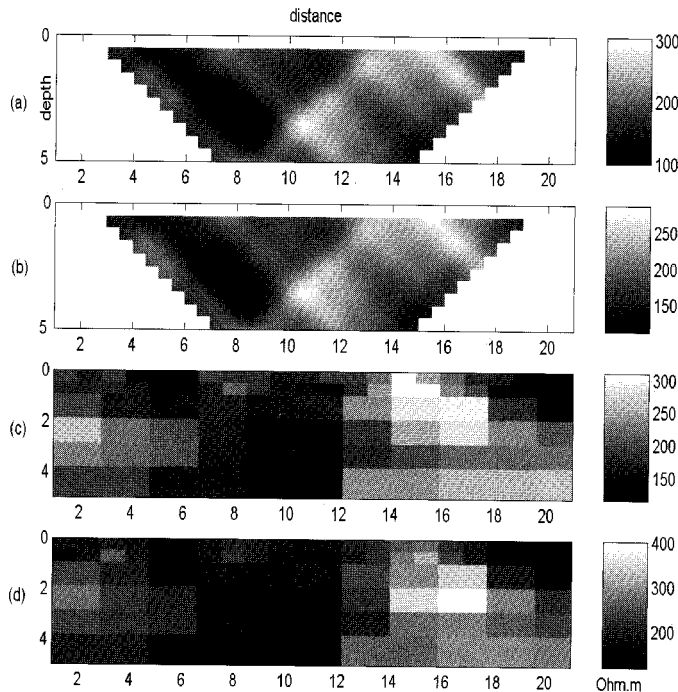


Fig. 7. Maximum *a posteriori*(MAP) map. (a) Observed pseudosection for the model of Fig. 1(d). (b) Pseudosection calculated by Bayesian inversion model (d). (c) *A priori* model map generated from dipole-dipole DC resistivity data.

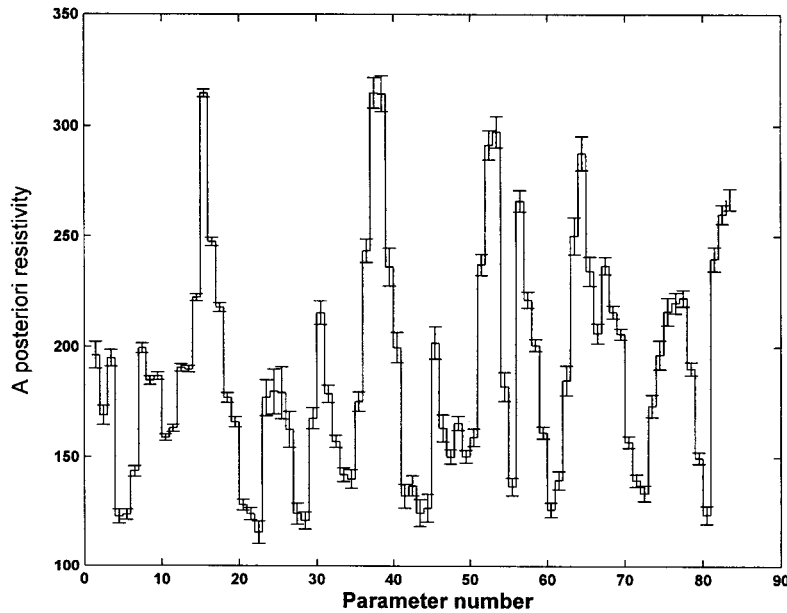


Fig. 8. Confidence interval given by standard deviation length. The block numbers are defined in Fig. 2.

가정하고, 몇 가지 선형대수 기법을 이용하여, 사후 확률분포 역시 정규분포에 의해 정의되며, 그 평균은 m_{map} 이고 공분산은 $m_{map} (A^T C_d^{-1} A + C_m^{-1})$ 이라 규정하였다. 정규분포의 가정 하에서 완전한 불확실성 분석을 위해 사후 확률분포의 공분산에 의한 샘플링을 수행하며, 이를 위해 지구통계학적 기법을 다시 이용하고자 한다.

3 차원 이상의 문제에서 평균 m_{map} 공분산 ($A^T C_d^{-1} A + C_m^{-1}$)을 따르는 정규분포에 의해 확률변수를 실현하는 것은 정규분포에 의한 지구통계학적 시뮬레이션을 수행하는 것과 동일하다. 지구통계학에서는 공간적 공분산을 구하기 위해 배리오그램(variogram)을 작성한다. 이는 일정 거리 만큼 떨어져 있는 자료들의 상관성 혹은 비상관성을 나타내는 것으로 공간 자료의 자기 상관(auto covariance) 혹은 일반 상관(covariance) 관계를 표현한다. 이를 이용하여 관측점의 값을 기준으로 임의의 미관측지점에 대한 값을 추정 혹은 시뮬레이션 한다. 특히 관측점 부근의 공분산

이 정규분포를 따를 경우, 임의의 지점의 추정을 정규분포를 이용해 구한다(Deutsch · Journel, 1992). 정규분포를 이용한 지구통계학적 시뮬레이션 방법으로는 순차적 가우스 시뮬레이션(sequential Gauss simulation), LU 분해 시뮬레이션(LU decomposition simulation), 밴드 변환 시뮬레이션(turning band algorithm) 등이 있다.

이 연구에서는 사후 확률분포에 의한 공분산을 LU 분해 시뮬레이션에 적용하여 샘플링을 수행하였다. 이 방법은 공분산을 LU 분해하여, 시뮬레이션 결과의 분산이 공분산을 따르고, 평균은 0 인 점에 착안한다(Davis, 1987; Alabert, 1987; Parker, 1994).

이제 식 (12)로 주어지는 사후 확률분포의 공분산은 대칭 행렬(symmetric) 이며, 그 값은 유한(positive definite) 하다. 따라서 이 식을 LU 분해하여 $L^T = U$ 인 관계를 만들기 위해 Cholesky 분해를 수행한다. 즉,

$$C_{m'} = K = LU, L^T = U \quad (13)$$

여기서, C_m 는 $n \times n$ 공분산 행렬이다. 벡터 w 가 평균은 0 이고 분산은 1 인 정규분포를 따른다고 하면,

$$y = Lw \quad (14)$$

이고, yy' 의 기댓값은

$$\begin{aligned} E(yy') &= E(Lww'U) \\ &= LE(ww')U \end{aligned} \quad (15)$$

이다. 여기서 w 는 평균 0, 분산 1 의 벡터이므로,

$$E(ww') = I \quad (16)$$

이 성립하여,

$$\begin{aligned} E(yy') &= LIU \\ &= LU \\ &= K \end{aligned} \quad (17)$$

를 만족하며, $E(y)$ 또한 0 이므로, y 값을 시뮬레이션 하여, m_{map} 값에 더하면 식 (12)에 의한 정규분포 샘플링을 수행할 수 있다.

Fig. 9 는 상부 역산 블록에 대한 사전확률분포

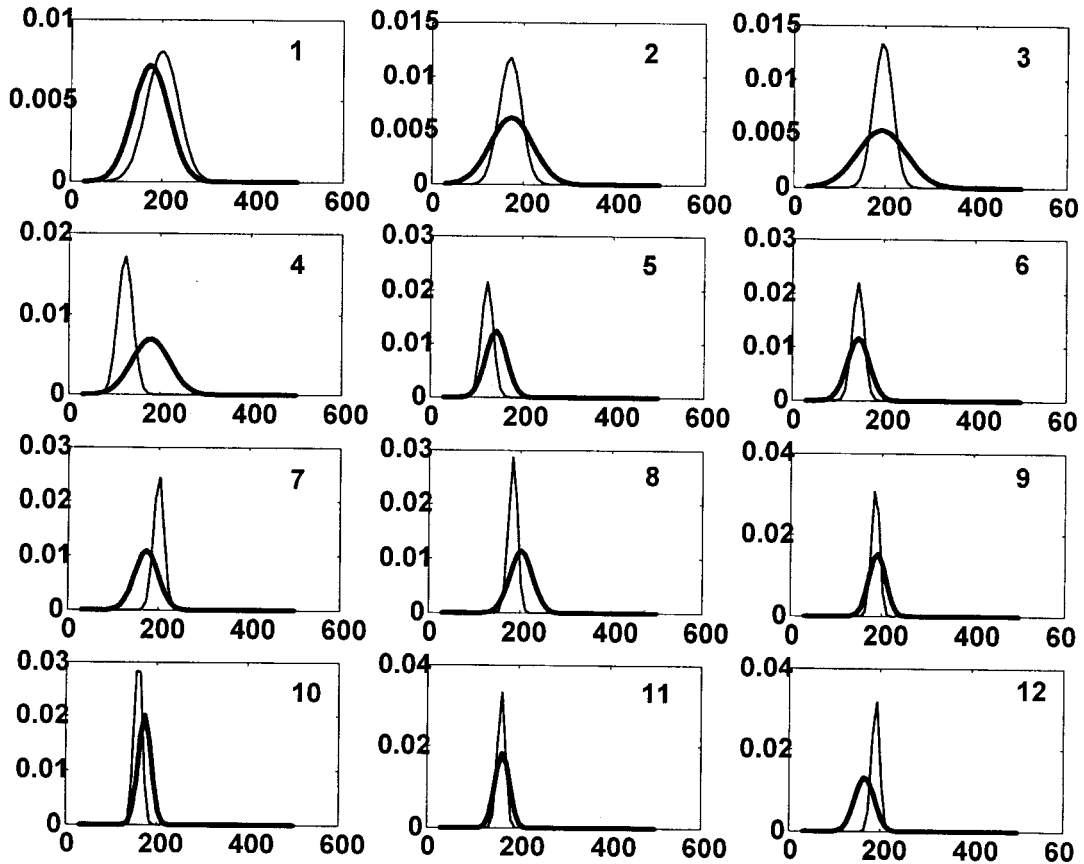


Fig. 9. Comparison of the probability density functions between *a priori*(bold line) and *a posteriori*(thin line) PDF for 12 upper blocks. The *a posteriori* PDFs are concentrated on specific region representing low uncertainties. Block numbers specified in Fig. 2 are displayed for the interpretation. The horizontal axes mean resistivities($\Omega \cdot m$), and the vertical axes present the probability.

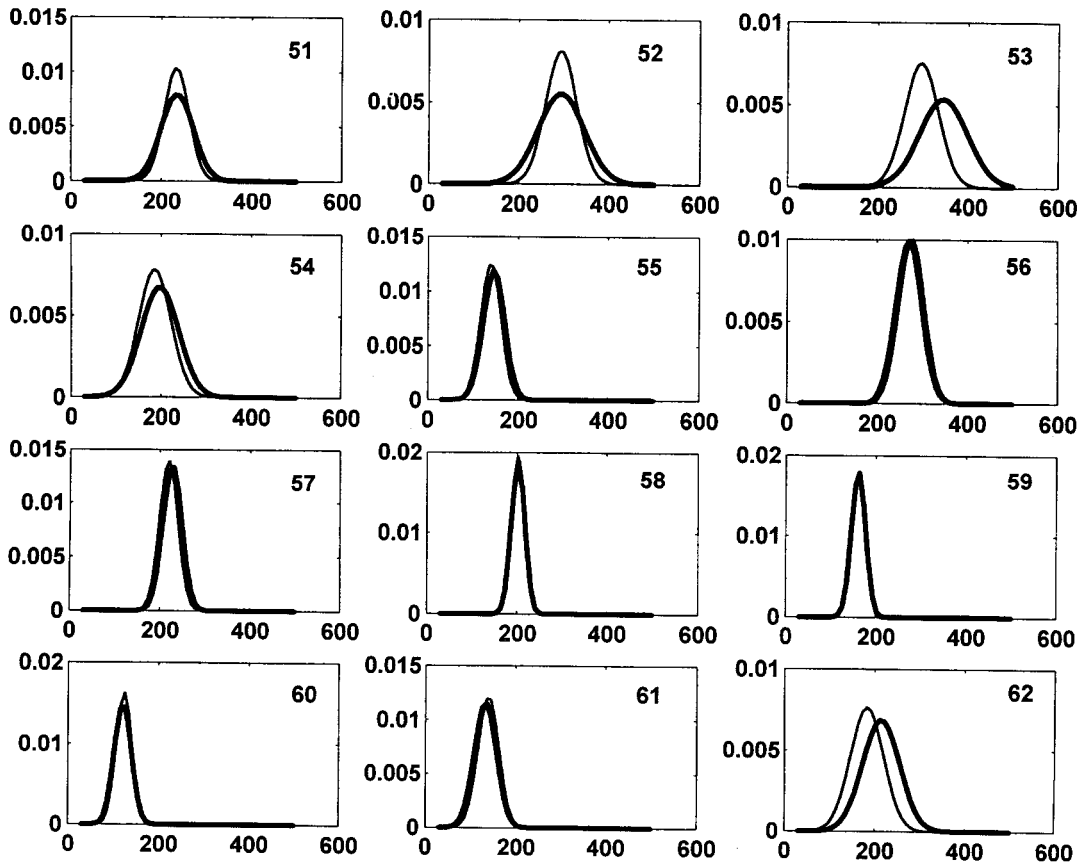


Fig. 10. Comparison of the probability density functions between a *a priori*(bold line) and a *a posteriori*(thin line) PDF for 12 lower blocks. Some a *a posteriori* PDFs do not appear to show prominent improvement in uncertainty. Block numbers specified in Fig. 2 are displayed for interpretations. The horizontal axes mean resistivity ($\Omega \cdot m$), and the vertical axes present the probability.

와 사후 확률분포에 대한 비교이다. 각 블록의 번호는 Fig. 1에 나타나는 전기비저항 역산 블록을 좌상단부터 우하단 방향으로 정한 것이며, 블록의 총 갯수는 83 개이다. 상부에 위치하는 이 블록들은 사전 정보를 크게 개선하고 있으며, 불확실성도 많이 감소하고 있는 것을 알 수 있다. 반면에 Fig. 10에 나타나 있는 하부 블록의 비교는 평균값이나, 불확실성이 사전 정보를 크게 벗어나지 못하고 있다. 이는 관측자료가 유의미한 자료를 제공하지 못하고 있다는 것을 의미하며, 사전 확률정보에 많이 의존하고 있다.

4. 결론

베이지안 역산은 다양한 지구물리 자료를 확률적으로 통합함으로써 새로운 방식의 복합역산 가능성을 제시한다. 이는 계속적으로 추가되는 지질학적, 지구물리학적 정보를 역산에 포함시킴으로써 보다 의미있는 해석 결과를 제공할 수 있을 것으로 생각된다. 또한 인위적 정규화 과정이 필요하지 않으며 해에 안정적으로 수렴하는 양상을 보인다. 어려운 점은 객관적인 공분산 자료의 획득과 사후 확률분포에 대한 추정이다. 이 연구에

서는 지구통계학적 기법을 최대한 이용하여 다양한 접근에 의해 공분산 자료를 획득하였으며, 신속한 해석을 위해 사전 확률분포와 우도 함수가 정규분포라는 가정에 의해 사후 확률 분포를 도출하였다. 정확한 베이지안 해석을 위해서는 비모수적 확률분포의 적용과 몬테 카를로 방법에 의한 일반적 접근이 필요하지만 신속하게 불확실성에 대한 고찰이 필요한 경우를 위하여 정규분포의 가정을 이용할 수 있었다. 이는 최대 엔트로피 방법(maximum entropy method)의 관점에서 볼 때, 모멘트가 3 이상인 첨도(kurtosis) 등과 같은 정보는 제외하므로 비교적 평활화 된 결과로 볼 수 있다. 또한 사후 확률 분포를 분석하기 위해 사후 공분산 행렬에 의존하는 샘플링을 수행하여 불확실성 분석을 수행하였다. 그 결과 모델 변수 각각에 대한 확률적 정보를 알 수 있었고, 필요한 모델 변수의 정확한 해석을 위해 다른 탐사가 필요한 지 여부를 결정할 수 있었다. 이는 최근 지구물리탐사가 대규모 복합 탐사로 가는 시점에 있어 적절한 복합 자료 해석 기술과 효율적 프로젝트 수립을 위한 객관적 정보를 제공할 수 있을 것으로 생각된다.

사 사

이 연구는 기상연구소 주요사업 “수퍼컴을 활용한 예보능력 향상 연구”의 일부 지원으로 수행되었습니다.

참 고 문 헌

- Alabert, F., 1987, The practice of Fast Conditional Simulations Through the LU decomposition of the covariance matrix, *Mathematical Geology*, 19, 369-386.
- Broyden, C. G., 1965, A class of method for solving nonlinear simultaneous equations, *Maths. Comp.*, 19, 577-593.
- Constable, S. C., Parker, R. L. and Constable, C. G., 1987, Occam's inversion : A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, 52, 289-300.
- Cruse, E., Pica, A., Noble, M., McDonald, J. and Tarantola, A., 1990, Robust elastic nonlinear waveform inversion : Application to real data, *Geophysics*, 55, 527-538.
- Davis, M. W., 1987, Production of conditional simulations via the LU triangular decomposition of the covariance matrix, *Mathematical Geology*, 19, 91-98.
- Deutsch, C. V., and Journel, A. G., 1992, *GSLIB : Geostatistical software library and user's guide*. Oxford University Press, New York.
- Geman, S. and Geman D., 1984, Stochastic Relaxation, Gibbs distributions, and the Bayesian Restoration of Images, *IEEE Transactions on pattern analysis and machine intelligence*, PAMI-6, 721-741.
- 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*. Oxford University Press, New York.
- Journel A. G. and Huijbregts, C. J., 1978, *Mining Geostatistics*. Academic Press, New York.
- Vozoff, K. and Jupp, D. L. B., 1975, Joint inversion of geophysical data, *Geophys. J. R. astr. Soc.* 42, 977-991.
- 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*, p.81-142, Kluwer Academic Publishers, The Netherland.
- Mosegaard K. and Tarantola A., 1995, Monte Carlo sampling of solutions to inverse problem, 100, 12431-12447.
- Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multi-offset seismic data, *Geophysics*, 52, 1211-1228.
- Parker, R. L., 1994, *Geophysical Inverse Theory*. Princeton University Press, New Jersey.
- Sambridge, M., 1999, Geophysical inversion with a neighbourhood algorithm-II. Appraising the ensemble, *Geophys. J. Int.*, 138, 727-746.
- Stark, P. B., 1992, Minimax confidence interval in geomagnetism, *Geophys. Jour. International*, 108, 329-338.
- Tarantola, A., 1987, *Inverse problem theory : Methods for data fitting and model parameter estimation*. Elsevier.
- Gilks, W. R., S. Richardson, D.J., Spiegelhalter, 1996, *Markov Chain Monte Carlo in practice*. Chapman

& Hall

권병두, 이희순, 오석훈, 이춘기, 1999, 제주도 만장굴에서의 복합지구물리 탐사, 대한자원환경 지질학회지, 31, 535-545.

김정호, 1987, 상반성 정리를 이용한 2차원 전기탐사 자료의 역산. 서울대학교 대학원 박사 학위 논문.

오석훈, 2000, 지구통계학적 방법에 의한 지구물리 자료의 베이지안 역산. 서울대학교 대학원 박사학위 논문.

오석훈, 권병두, 1997, Simulated Annealing을 이용한

전기탐사자료의 역산, 한국자원공학회지, 34, 285-293.
정호준, 1998, 웨이블릿 영역에서의 근사적 민감도 갱신을 이용한 2차원 MT 신속역산. 서울대학교 대학원 박사학위 논문.

2000년 5월 29일 원고접수

2000년 9월 21일 원고채택