

베이지안 방식에 의한 지구물리 역산 문제의 접근

오석훈¹⁾ · 정승환²⁾ · 권병두³⁾ · 이희순⁴⁾ · 정호준⁵⁾ · 이덕기¹⁾

A Bayesian Approach to Geophysical Inverse Problems

Seokhoon Oh¹⁾, Seung-Hwan Chung²⁾, Byung-Doo Kwon³⁾, Heuisoon Lee⁴⁾, Ho Jun Jung⁵⁾ and Duk Kee Lee¹⁾

요 약 : 본 연구에서는 지구물리 자료의 베이지안 역산을 효과적으로 수행하는 방법에 관해 논의하였다. 베이지안 처리에서 가장 문제가 되는 사전확률분포를 구하기 위해 지구통계학적 방법을 적용하였으며, 사후확률분포의 추정을 위해 MCMC(Markov Chain Monte Carlo) 방법을 적용하였다. 쌍극자배열 전기비저항 탐사 자료의 2차원 역산을 위해 슬럼버저배열 전기비저항탐사 자료와 시추공 자료를 사전 정보로 이용하였으며, 이들 사전정보에 대해 지구통계학적 방법을 적용하여 사전확률분포를 작성하였다. 쌍극자배열 전기비저항 탐사 자료를 최대 우도함수로 하는 사후확률분포는 차원이 매우 높은 적분을 요구하므로, 이를 추정하기 위해 MCMC기술을 적용하였으며, 보다 효율적인 접근을 위해 Gibbs 샘플링 방법을 이용하였다. 그 결과 비모수적 방식으로 사후확률분포를 분석함으로써 보다 신뢰성 있는 해를 구할 수 있었으며, 주변화(marginalization)된 사후확률분포를 이용하여 다양한 분석을 적용할 수 있었다.

Abstract : This study presents a practical procedure for the Bayesian inversion of geophysical data. We have applied geostatistical techniques for the acquisition of prior model information, then the Markov Chain Monte Carlo (MCMC) method was adopted to infer the characteristics of the marginal distributions of model parameters. For the Bayesian inversion of dipole-dipole array resistivity data, we have used the indicator kriging and simulation techniques to generate cumulative density functions from Schlumberger array resistivity data and well logging data, and obtained prior information by cokriging and simulations from covariogram models. The indicator approach makes it possible to incorporate non-parametric information into the probabilistic density function. We have also adopted the MCMC approach, based on Gibbs sampling, to examine the characteristics of a posteriori probability density function and the marginal distribution of each parameter.

Keywords : Bayesian inversion, MCMC, Gibbs sampling, posterior PDF

서 론

80년대 이후 지구물리 역산문제를 확률론적으로 인식하고, 베이지안(Bayesian) 방식으로 접근하는 연구들이 진행되어 왔다(Duijndam, 1988a, b; Gouveia, 1996; Moraes, 1996). 역산문제를 변수 추정(parameter estimation) 문제로 다루고자 하는 베이지안 방법은 많은 연구를 통해 그 장점이 부각되었지만, 측정 자료 및 추정하고자 하는 물성의 사전 정보(prior information)에 대한 확률화 문제 등을 객관적으로 마련하는 것이 쉽지 않았다. 실제로 베이지안 역산에서 가장 난해한 부분은 사전확률분포(prior probability distribution)의 구현이다. 일반적으로 기존의 확률적 역산 방법과 베이지안 역산의 차이는

사전확률분포의 이용과 추정하고자 하는 대상의 비선형성에서 나타나게 된다(Loredo, 1990). 선행 연구자들은 사전확률분포의 작성을 위해 임의로 정규분포를 가정한 후, 평균과 표준편차를 정의하거나(Duijndam, 1998b), 최대 엔트로피 방법(maximum entropy method)의 적용(Moraes, 1996), 다양한 실험 오차들의 도입(Gouveia, 1996) 등을 시도하였다. 현재까지 사전정보의 확률분포를 객관적으로 구현할 수 있는 기술로는 최대 엔트로피 방법이 많이 이용되고 있지만 지구물리 역산 문제는 차원이 높아서 풀어야 할 비선형 연립방정식이 매우 복잡해져서 이 방법을 적용하기 어렵다(Moraes, 1996). 본 연구에서는 지구통계학적 방법을 도입하여 사전확률분포를 객관적으로 구현하는 문제를 해결하고 지구과학 자료의 공간적 특성을 반영할

*2002년 10월 14일 접수

- 1) 기상연구소 해양기상지진연구실(Marine Meteorology & Earthquake Res. Lab/METRI)
- 2) 한국지질자원연구원 탐사개발연구부(Geophysical Exploration and Mining Division, Korea Institute of Geoscience and Mineral Resource)
- 3) 서울대학교 지구과학교육과(Dept. Earth Sciences Education, Seoul National University)
- 4) 인천교육대학교 과학교육과(Dept. Science Education, Incheon Nat'l. Univ. of Education)
- 5) (주)희송지오테크(Heesong Geotek, Co. Ltd.)

수 있는 확률분포를 작성하였다.

또한 사전정보와 관측자료를 이용하여 사후확률분포(*posterior probability distribution*)를 작성하더라도 지구물리 역산과 같이 변수의 차원이 매우 높은 문제의 경우, 각각의 매개변수의 특성을 파악하기 위해서는 변수 차원 만큼의 다중적분을 수행해야 한다(Loredo, 1990). 이는 수치적분으로도 어려운 문제이며 간접적으로 확률분포를 파악할 수밖에 없다. 이를 위해 베이지안 연구에서 많이 이용하는 방법 중의 하나가 마르코프 사슬에 의한 몬테카를로(Markov Chain Monte Carlo; MCMC) 적분 기술이다. 이는 독립적으로 임의의 확률분포로부터 변수를 추출하여 추정하고자 하는 확률분포의 특성을 구해내는 기술이다(Geman and Geman, 1984). 본 연구에서는 사전확률분포로부터 변수를 추출하여 MCMC 적분기술을 적용하였다.

본 연구에서는 지구통계학적 방법을 이용하여 슬림버저배열 전기비저항 탐사자료와 시추공 자료의 사전정보를 확률화하고, 쌍극자배열 전기비저항 탐사자료를 관측자료로 이용하여 MCMC 적분기술을 통해 베이지안 방식으로 사후확률분포를 분석한 결과에 대하여 논하고자 한다.

사전정보에 대한 지구통계학적 접근

지구통계학(Geostatistics)은 광체의 위치를 추정하기 위해 연구된 분야이지만 현재는 크리깅(kriging)이나 모사(simulation)방법을 이용하여 공간자료의 불확실성 추정 등에 널리 이용되고 있다(Goovaerts, 1997; Isaaks and Srivastava, 1989). 본 연구에서는 지구통계학적 방법을 이용하여, 슬림버저배열 전기비저항 역산자료와 시추공 자료에 대해 배리오그램 모델링을 수행하여 역산 블록의 확률분포를 추정하는데 적용하였다. 여기서는 지구통계학에 대한 상세한 설명은 피하며 이에 대해서는 Goovaerts(1997)를 참조할 수 있다.

역산 블록의 불확실성 추정

일반적으로 지구통계학의 모사 기술을 이용하면 역산에서 매개변수화하는 블록보다 훨씬 많은 지점에서 샘플 값을 얻을 수 있다. 이를 이용하면 간접적인 방법으로 역산 블록에 대한 사전확률분포를 추정할 수 있다.

사전정보를 이용하여 역산블록의 확률분포를 추정하는 것은 블록의 조건부 누적확률밀도함수(CCDF; Conditional Cumulative Density Function)를 구하는 문제가 되는데, 이 CCDF는 $F_V(\mathbf{u};z|n)$ 으로 주어지며, 역산 블록 $V(\mathbf{u})$ 에 대해 샘플값의 불확실성을 모델링 할 수 있는 것이어야 한다. 이 F_V 는 n 개의 주변 데이터에 의존하여 \mathbf{u} 의 위치에서 샘플된 것에 의해 구현된 것을 의미한다. 본 연구에서는 사전정보의 배리오그램 모델링을 위해 지시자 변환을 이용하였는데(Goovaerts, 1997), 이 변환의 비선형성 때문에 우리가 사전정보(본 연구에서는 슬림버저배열 전기비저항 역산모델과 시추공 자료)를 통해 구해놓

은 point CCDF를 단순히 선형 결합하는 것만으로 블록의 CCDF를 구할 수는 없다. 즉,

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

이다. 여기서 point CCDF $F_V(\mathbf{u};z|n)$ 는 블록 $V(\mathbf{u})$ 를 분할하는 J 개의 포인트, \mathbf{u}_j 로 정의되며, $[F_V(\mathbf{u};z|n)]^*$ 는 샘플로부터 추정되는 위치 \mathbf{u} 점 주변의 n 개의 데이터에 의존하는 블록 V 에 대한 CCDF의 추정을 의미한다. 풀어서 말하면, 사전자료의 샘플링을 통해 임의의 역산 블록에 대한 사전확률분포를 구현할 때, 단순히 그 블록 안에 있는 샘플 값들의 평균을 이용해서 확률분포를 작성하면 안 된다는 의미이다. 블록 자체에 대한 자료나 이에 대한 통계값이 주어질 경우에는 쉽게 구할 수 있지만, 지구물리 역산문제의 블록은 임의적으로 결정되므로 각 블록의 통계정보는 존재하지 않는다.

이를 해결하기 위해서는, 즉 $[F_V(\mathbf{u};z|n)]^*$ 를 제대로 추정하기 위해서는, 블록내의 임의의 지점에 대한 다량의 모사를 통해 누적확률분포를 수치적으로 근사하는 방법을 이용해야 한다(Isaaks, 1990; Gomez-Hernandez, 1991; Deutsch and Journel, 1992; Glacken, 1996). 즉,

$$[F_V(\mathbf{u};z|n)]^* = \frac{1}{L} \sum_{l=1}^L i_V^{(l)}(\mathbf{u};z) \quad (2)$$

을 수행해야 하는데, 여기서 블록 지시자 값 $i_V^{(l)}(\mathbf{u};z)$ 는 $z_V^{(l)}(\mathbf{u}) \leq z$ 일 경우는 1, 아닐 경우는 0으로 처리하여, 임의의 V 번째 블록에 대한 CCDF를 값의 변화에 따라 이산적으로 추정한다. 또한 각각의 모사블록 $z_V^{(l)}(\mathbf{u})$ 는 다음과 같이 주어지는데

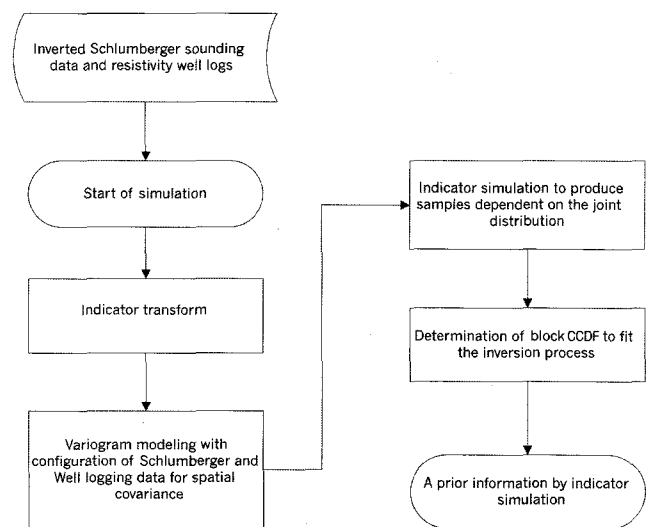


Fig. 1. Flowchart for preparation of prior information through a geostatistical simulation using Schlumberger array resistivity and well logging data.

$$z_V^{(l)}(\mathbf{u}) = \frac{1}{J} \sum_{j=1}^J z^{(l)}(\mathbf{u}_j), \quad (3)$$

이는 임의의 블록 $V(\mathbf{u})$ 를 분할하는 J 개의 포인트 중, \mathbf{u}_j 의 위치에서 모사되는 값의 평균으로부터 구해진다. 즉, 배리오그램 모델에 기반하여 모사를 수행함으로써 간접적으로 블록의 CCDF를 추정할 수 있다. Fig. 1은 이에 대한 알고리즘을 순서도로 나타낸 것이다.

Markov Chain Monte Carlo 방법

Tarantola(1987)는 지구물리 역산 문제의 완전한 해는 사전 정보와 관측자료로부터 구해지는 사후확률분포로부터 얻어질 수 있다는 것을 정보이론을 이용하여 규명하였으며 이는 베이 지안 접근과 동일하다. 즉,

$$\sigma(\mathbf{m}) = k\rho(\mathbf{m})L(\mathbf{m}) \quad (4)$$

으로 주어질 수 있으며, $\sigma(\mathbf{m})$ 은 사후확률밀도 함수(posterior PDF; Probability Density Function), $\rho(\mathbf{m})$ 은 사전확률밀도 함수(prior PDF), $L(\mathbf{m})$ 은 관측자료에 의한 우도 함수(likelihood PDF)로 정의하며, \mathbf{m} 은 물성을 나타내는 변수 벡터이다. k 는 확률밀도함수를 만족하기 위해 주어진 비례상수이다. 이와 같이 단순히 사전확률밀도 함수와 우도 함수를 곱함으로써 완전한 역산 문제의 해를 사후확률밀도 함수로 구할 수 있다. 그러나 일반적으로 \mathbf{m} 벡터가 다차원이므로, 각 성분의 확률분포를 추정하기 위해서는 주변화(marginalization) 과정을 거쳐야 한다(Moraes, 1996). 주변화란 확률밀도함수가 결합확률분포(joint probability distribution)의 형태를 가질 때, 원하는 변수 이외의 나머지 부분(nuisance parameter)을 모두 적분하여 제거하는 과정을 말한다. 이 과정은 (매개변수-1)차원의 다중 적분을 수행해야 하는데, 이는 해석적으로는 불가능하며 수치적으로도 계산량이 너무 방대하다. 따라서 사후확률분포를 정량적으로 완전히 계산하는 방법보다는, 확률분포의 통계적 특성을 간접적으로 추정하는 방법이 쓰이고 있는데, 그 대표적인 것이 마르코프 사슬에 의한 몬테카를로(MCMC) 적분 기술이다(Gilks et al., 1996; Oh and Kwon, 2001). 상세한 설명은 Gilks et al.(1996)의 저서를 참고할 수 있다.

MCMC 적분방법의 핵심을 요약하면, 임의의 확률밀도함수로부터 일련의 무작위 샘플링을 수행하여 사후확률분포의 특성에 맞는 것들을 골라내는 것이다. 대표적인 MCMC 방법으로 Metropolis 알고리즘(Hastings, 1970)이 있다. 이는 균일한 랜덤 함수에서 변이량 $\delta\mathbf{x}$ 를 샘플하여, 변수 벡터 $\mathbf{x} + \delta\mathbf{x}$ 가 주어진 규칙에 따라 허가 혹은 거절되는 과정을 반복함으로써 일정한 구간이 지나면 추정하고자 하는 확률분포의 특성을 반영할 수 있도록 하는 것이다(Fig. 2). 이 방법은 계산이 간단하므로 쉽게 구현할 수 있으나, 벡터 성분 전체에 대해 한번에 변이가 수행되므로 허가되는 비율이 매우 낮아지게 되며, 따라

Metropolis algorithm

- Select initial parameter vector \mathbf{x}_0
- Iterate as follows: at iteration number k
 - (1) create new trial position $\mathbf{x}^* = \mathbf{x}_k + \Delta\mathbf{x}$, where $\Delta\mathbf{x}$ is randomly chosen from $\mathcal{L}(\Delta\mathbf{x})$
 - (2) calculate ratio $r = q(\mathbf{x}^*)/q(\mathbf{x}_k)$
 - (3) accept trial position, i.e. set $\mathbf{x}_{k+1} = \mathbf{x}^*$ if $r \geq 1$ or with probability r , if $r < 1$ otherwise stay put, $\mathbf{x}_{k+1} = \mathbf{x}_k$

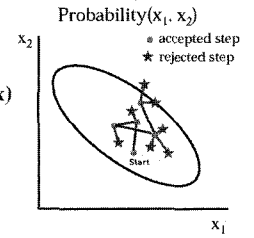


Fig. 2. Metropolis algorithm for MCMC method.

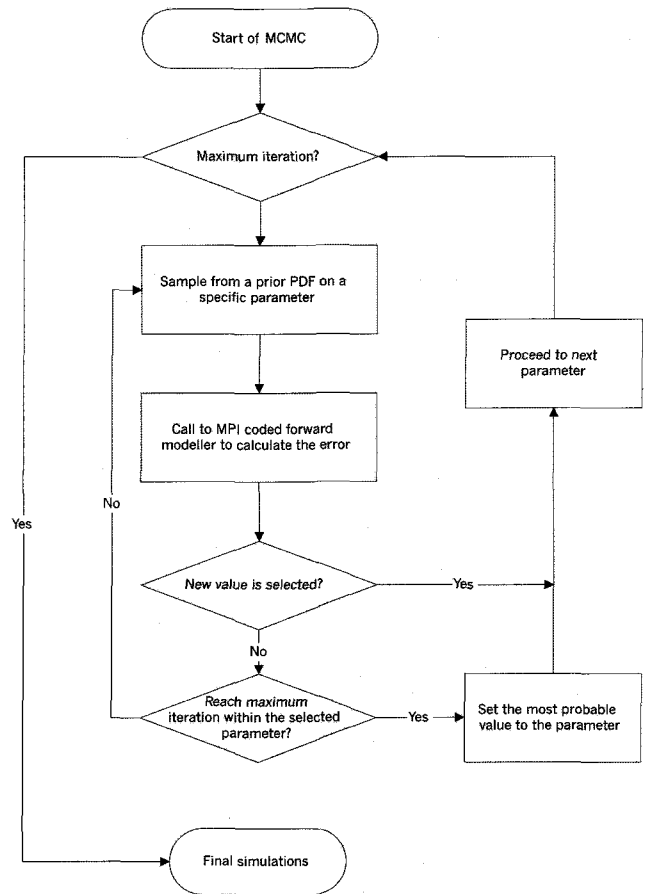


Fig. 3. Flowchart for the MCMC method by Gibbs sampling applied to the analysis of the posterior information.

서 확률분포의 추정에 매우 많은 시간이 소요된다.

이에 본 연구에서는 확률분포 추정의 효율성을 높이기 위해 Gibbs 샘플링을 이용하였다(Gilks et al., 1996). Gibbs 샘플링은 변수 벡터의 한 성분씩만을 변화시키게 되는데, 그 변화값은 변화 대상의 성분을 제외한 나머지 성분에 의존하여 가장 높은 확률 값을 갖는 값이 샘플된다. 즉, 변수 벡터 \mathbf{x} 의 x_j 성분을 구현할 때, 조건부확률분포(전 조건 확률분포: full conditional distribution) $q(x_j|x_1x_2\dots x_{j-1}x_{j+1}\dots)$ 를 이산적으로 구하여 가장 확률이 높은 값을 선택하게 된다. 특히 본 연구에

서는 임의의 확률분포 q 를 앞절에서 구현한 사전확률분포로 대체하여 더욱 효율성을 높였다. Fig. 3은 본 연구에서 적용한 MCMC 샘플링에 대한 알고리즘을 순서도로 표현한 것이다.

모델에 대한 적용

모델의 특성

Fig. 4(b)는 본 연구에서 제시한 알고리즘을 적용하기 위해 준비한 전기비저항 모델이다. 기존의 모델링 연구들이 역산 블록과 모델 블록을 동일하게 설정하여 현실성이 떨어지는 점을 극복하기 위해, 본 연구에서는 모델 블록을 더욱 조밀하게 구성하여 실제 자연적인 모델이 구현될 수 있도록 하였다. Fig. 4(b)의 모델은 임의의 지점(모델 블록보다 훨씬 산발적으로 존재)에 대표값으로써 전기비저항값을 지정하고 배리오그램 모델을 통해 지구통계학적 모사를 수행하여 작성되었다.

이 모델은 Fig. 4의 (b)에 나타나 있는 바와 같이, 역산 블록(Fig. 5)보다 훨씬 조밀하며, 이 모델로부터 계산된 순산모델(Fig. 4(a))은 관측값(V/m)에 대해 10%의 Gaussian 잡음을 포함하도록 구성되었다. Fig. 4(d)는 (a)의 관측값에 대해 기존에 사용되고 있는 쌍극자탐사 자료의 2차원 M-L(Marquardt-Levenberg) 역산 방법(Lines and Treital, 1984)을 이용하여 얻

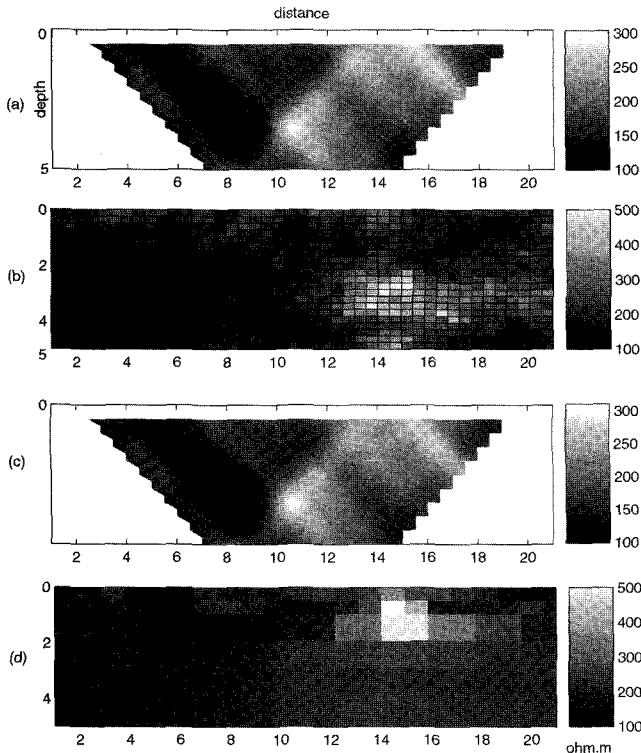


Fig. 4. Proposed model for this study generated by geostatistical simulation. (a) Synthetic dipole-dipole resistivity data contaminated by 10% Gaussian noise for the forward response (V/m) of model (b). (c) Calculated responses from the inverted resistivity section (d) by observed data-fit alone. The depth and horizontal distance are in arbitrary model unit.

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. 5. Block parameterization prepared for resistivity inversion. Block numbers are displayed for following interpretations.

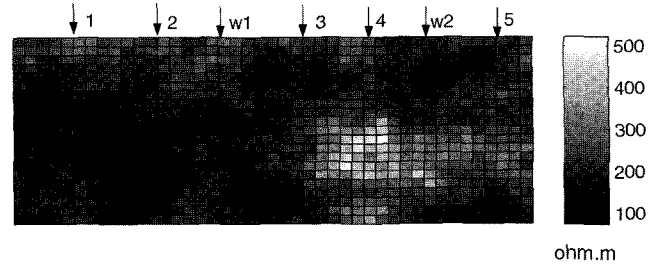


Fig. 6. Configuration map showing the location of the well logs and Schlumberger array resistivity soundings in synthetic way. The prefix w means the position of well logging simulations and the others are Schlumberger array sounding positions.

은 결과이다. 쉽게 비교할 수 있듯이, 실제 구조와 많은 차이를 보이고 있으나 관측값과 이론 모델링 결과들((a)와 (c))은 별 차이가 없음을 알 수 있다. 본 연구에서는 M-L 역산과정에서 평활화 제한을 가하지는 않았다. 그러나 평활화 제한과 관계없이 하부에 위치하고 있는 고비저항대의 탐지는 실패하고 있다.

사전확률밀도분포

Fig. 6은 사전정보로 이용되어진 슬럼버저배열 전기비저항 탐사 및 시추공 자료가 획득된 지점을 나타낸다. Fig. 7과 Fig. 8은 각각 Fig. 4의 모델 상에서 얻어진 슬럼버저배열 전기비저항 자료의 역산 모델 및 전기비저항 시추공 자료의 모델을 보여준다. 슬럼버저배열 전기비저항 역산은 오석훈과 권병두(1997)가 이용한 M-L과 SA 복합 역산을 통해 획득되었다. 각기 5%의 Gaussian 잡음이 첨가되었다. Fig. 9는 두 가지의 사전정보로부터 사전확률분포를 구하기 위한 과정을 도식적으로 나타낸다. 슬럼버저배열 전기비저항 역산 자료와 전기비저항 시추 자료에 대하여 지시자 코베리오그램 모델링(indicator covariogram modeling)을 수행하여, 앞절에서 언급한 바와 같이 역산 블록 확률분포 추정을 위한 모사를 적용하였다. 슬럼버저배열 전기비저항 역산 자료와 시추공 자료의 해상도가 다르기 때문에, 배리오그램 모델링 과정이 까다로웠지만 안정적으로 수행되었다. Fig. 10은 이상의 방법을 이용하여 구현한 사전확률분포이다. 우측 상단의 번호는 Fig. 5에 있는 블록 번호를 의미한다. 예상할 수 있는 바와 같이, 슬럼버저배열 전

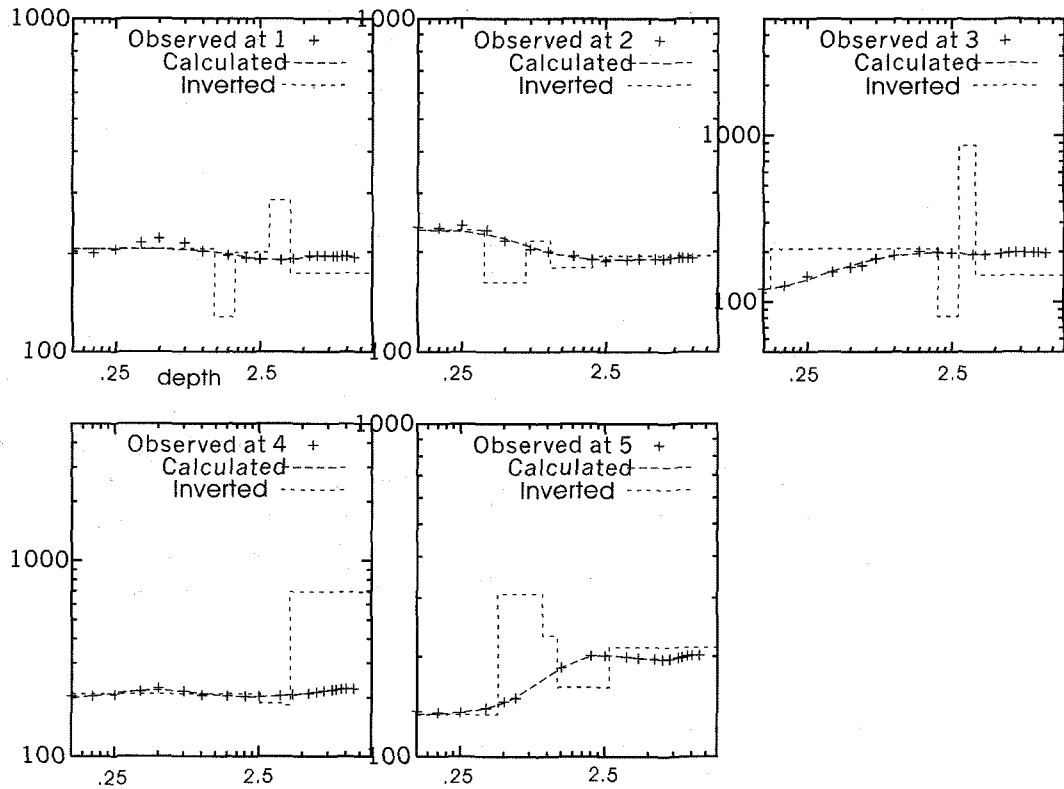


Fig. 7. Schlumberger array sounding inversion results from the model in Fig. 6.

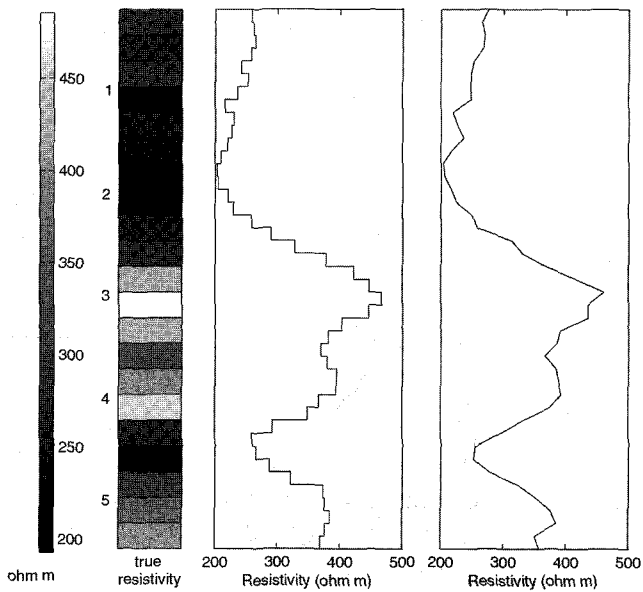


Fig. 8. Simulated resistivity well logging data at position w2 in Fig. 6. The left stripe shows the true resistivity structure and the log in the middle is generated by five-point averaging with the resistivity values read at every 0.125 unit. The well log in the right is simulated by adding correlated Gaussian noise.

기비저항 탐사자료와 시추공 자료로부터 멀리 떨어져 있는 블록(1,2,3,14,15 등)의 경우 불확실성이 큰 양상을 보이고 있다.

Gibbs 샘플링과 Burn-in 계수

사후확률분포를 분석하기 위해서 앞절에서 언급한 바와 같이 Gibbs 샘플링에 의한 MCMC 적분을 수행하였다. 본 연구에서는 특별히 지구통계학적 방법에 의해 얻어진 사전확률분포를 전 조건 확률분포로 가정하여 Gibbs 샘플링을 수행하였다. 사전확률분포를 샘플링하기 위해, Fig. 10의 확률분포를 누적확률분포 형태로 변환하여 주어진 각각의 변수에 대해 0에서 1사이의 무작위 샘플링을 수행하고 이를 역함수로 대응시켰다.

한편 MCMC를 이용하여 확률분포를 분석하기 위해서는 일정기간의 샘플이 진행 된 이후의 결과를 이용하여야 한다. 이 기간을 burn-in 계수라 하며(Gilks *et al.*, 1996), 이것이 필요한 이유는 초기의 안정화 과정을 거쳐야 확률분포를 반영하는 샘플을 추출할 수 있기 때문이다. 현재 이 계수를 결정하는 정량적 기술은 없으며 Fig. 11과 같이 마르코프 사슬에 의한 샘플링 과정에서 그 변이가 수렴하는 수준에서 결정하게 된다. 본 연구에서는 MPI 병렬 라이브러리(Pacheco, 1997; MPI Forum, 1995)를 이용하여 모델링을 수행하였으며, burn-in 계수는 2500으로 결정하였다.

사후확률분포 분석

Fig. 12의 (a)와 (b)는 Gibbs 샘플링에 의해 얻어진 상부 층 15개의 블록변수의 사후확률분포에 대한 dots분포표와 누적

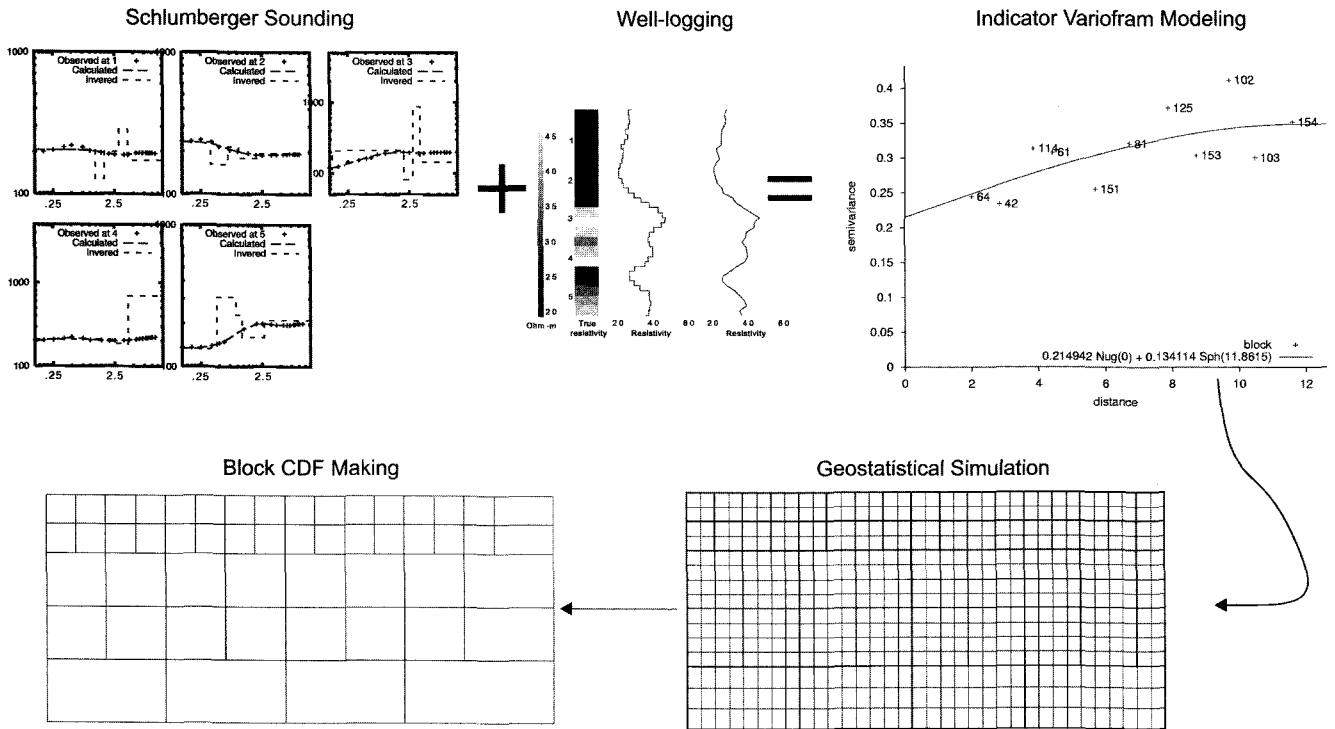


Fig. 9. Schematic process to generate prior information. Schlumberger array sounding and resistivity well logging data are put in covariogram modeling for a specified threshold for indicator kriging or simulation. Variogram model shows the data are rough from the high nugget value (Deutsch and Journel, 1992) but they have correlation and sustain a tendency. Then denser estimations contribute to the block CDF making, because it is impossible to estimate the resistivity of each block directly. So indirect estimation for blocks are made by the method proposed in previous section.

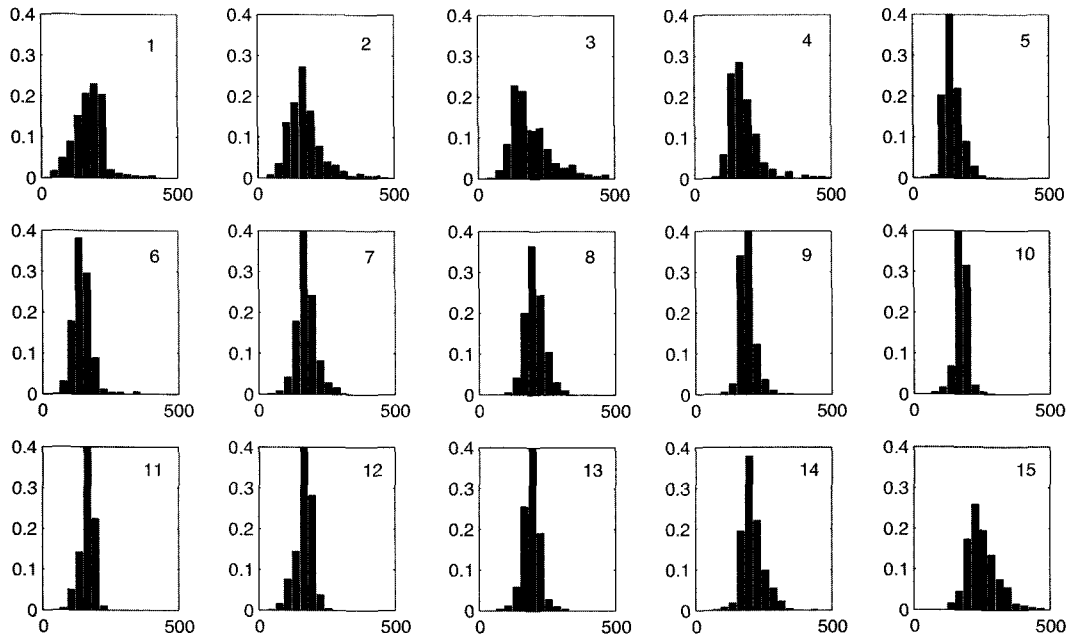


Fig. 10. Prior information generated by geostatistical simulation from Schlumberger array and well logging resistivity data. The block numbers are defined in Fig. 5. The horizontal axes mean resistivity in ohm-m, and the vertical axes present the probability.

확률분포를 보여준다. 특히 Fig. 12(b)에서는 Fig. 10의 사전누적확률분포를 같이 도시하였다. 베이지안 추정 의 핵심 중 하나

는 불확실성의 감소이다. 그림에서 보듯이 사후 누적확률분포의 변이가 특정부분에서 급격히 변하는 것이 쉽게 비교될 수

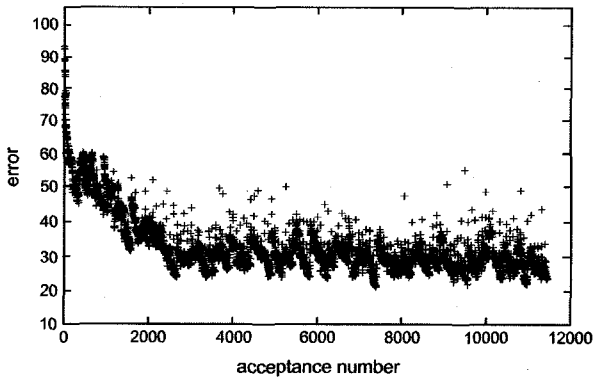


Fig. 11. Evolution of data-fit error as the function of the number of acceptance for the proposed model in Fig. 4.

있으며 이를 통해 사후확률분포의 불확실성이 매우 줄어들어 있음을 알 수 있다. 그러나 Fig. 13에 나타나 있는 바와 같이 심부의 블록들에 대한 불확실성은 그다지 많이 개선되지 못하고 있는 것을 알 수 있다. 이는 우도 함수로 이용한 쌍극자배열 탐사자료가 사전정보를 많이 개선시키지 못하고 있다는 것을 의미하며 이를 극복하기 위해서는 심부 정보를 포함하는 탐사자료가 필요하다는 것을 알 수 있다.

Fig. 14는 베이지안 역산에 의해 추정된 각 블록의 확률분포를 보다 쉽게 이해하기 위해, 최대 확률값(a) 및 평균값(b)으로 환산하여 역산 결과로 도시한 것이다. Fig. 4의 일반적 M-L 역산 방식과 비교하였을 때, 훨씬 참 값에 접근하고 있는 양상을 보인다.

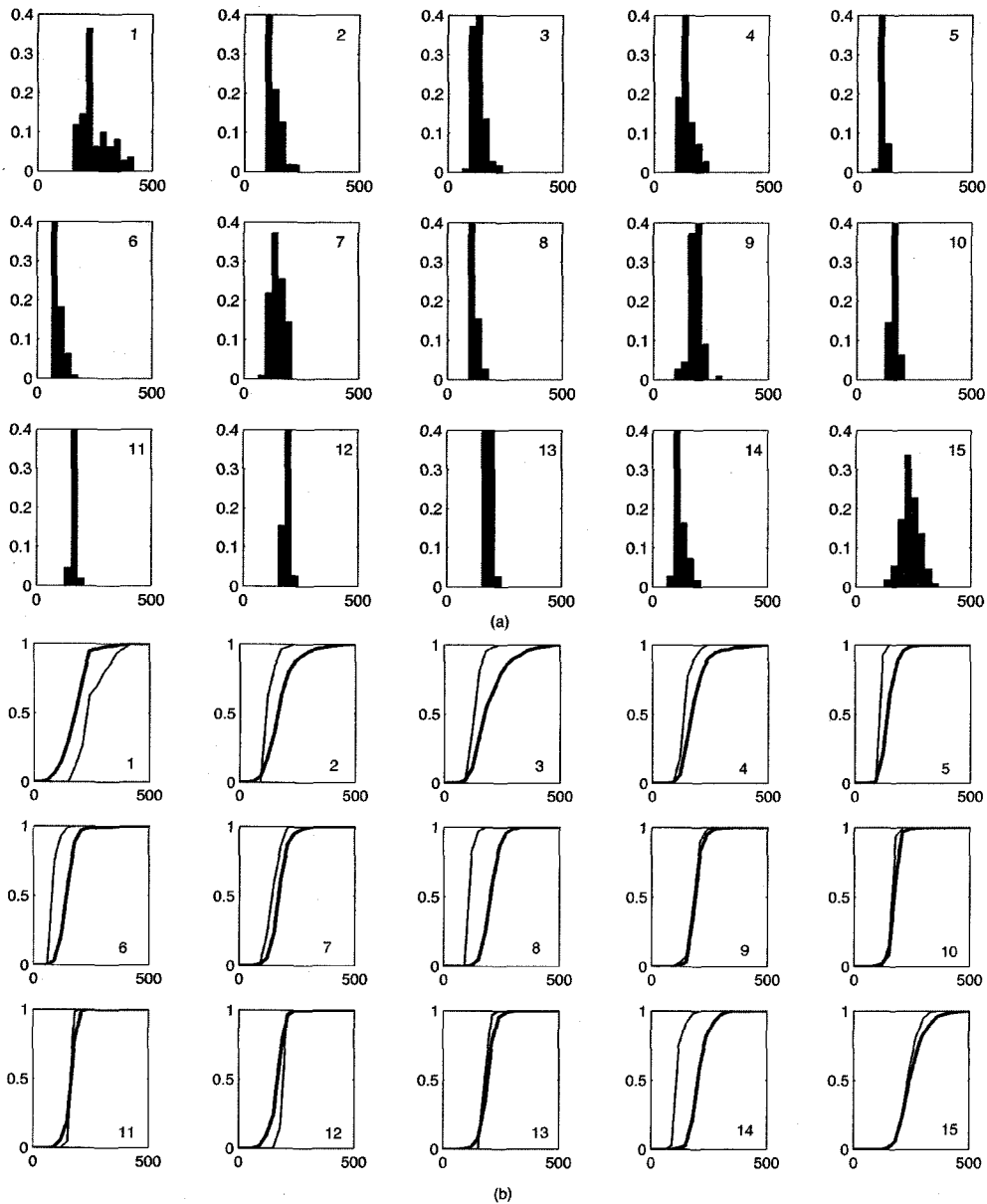


Fig. 12. (a) Posterior information given by probability density functions. The horizontal axes mean resistivity in ohm-m, and the vertical axes present the probability. (b) Cumulative Density functions for prior information (bold line) and posterior information (thin line) for upper blocks. The posterior CDF is concentrated on specific region representing low uncertainties. The block numbers are defined in Fig. 5. The vertical axes present CDF.

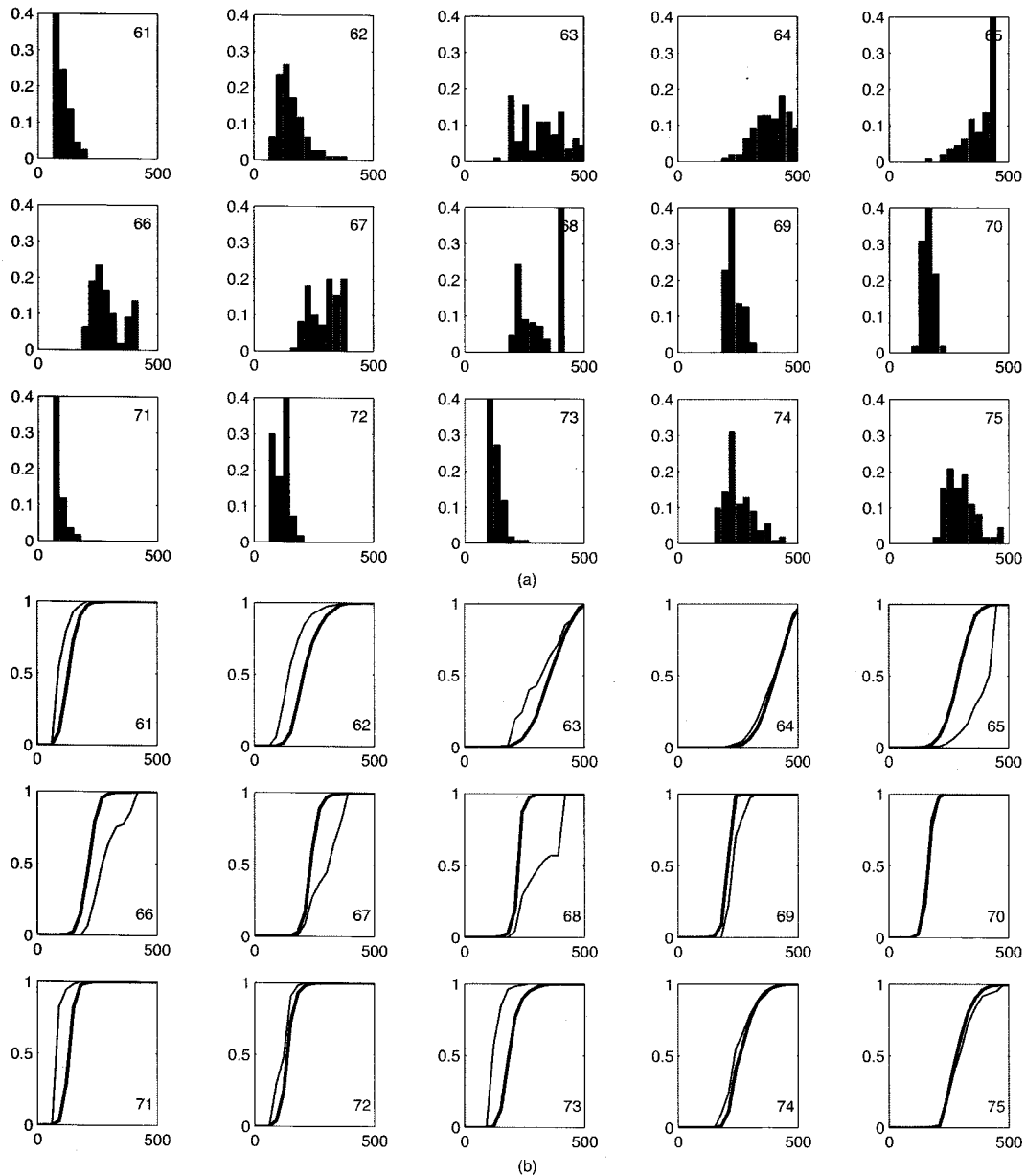


Fig. 13. (a) Posterior information given by probability density funtions. The horizontal axes mean resistivity in ohm-m, and the vertical axes present the probability. (b) Cumulative Density funtions for prior information (bold line) and posterior information (thin line) for lower blocks. The posterior CDF does not show drastic difference in uncertainties that means low information in observed dipole-dipole array resistivity data at lower blocks. The block numbers are defined in Fig. 5. The vertical axes present CDF.

결론

지구물리 역산 문제를 확률적으로 접근하는 것은 매우 당연하고 강력한 방법이지만 현재까지 구현의 어려움과 개념의 불일치 등으로 인해 많은 문제점을 가지고 있었다. 본 연구에서는 이를 극복하고 다양한 사전 정보들을 역산과정에 도입할 수 있는 베이지안 역산 방식을 제안하고 이를 이용하여 복잡한 구조에 대하여 적용하였다.

베이지안 역산을 적용하는 것은 크게 세 단계의 절차를 가지게 되는데, 그 첫째는 사전확률분포의 구현이다. 관측자료에

대해 독립적으로 획득할 수 있는 사전정보를 이용하는 것은 베이지안 방법에 있어 매우 중요하며, 효율적이고 정확한 변수 추정의 밑바탕이 된다. 변수가 극히 적은 인문학이나 단순한 문제에 대해서는 최대 엔트로피 방법이 주로 이용되어 왔으나 지구물리 역산은 변수의 차원이 매우 높아 이를 극복할 수 있는 방법이 필요하였다. 본 연구에서는 지구통계학을 이용하여 공간자료의 특성을 잘 반영하며 고차원의 변수를 확률분포로 변환할 수 있는 방법을 제안하였다.

두 번째 단계로는 사후확률분포의 적분문제이다. 사전확률 밀도함수와 우도 함수의 곱으로 주어지는 사후확률밀도함수는

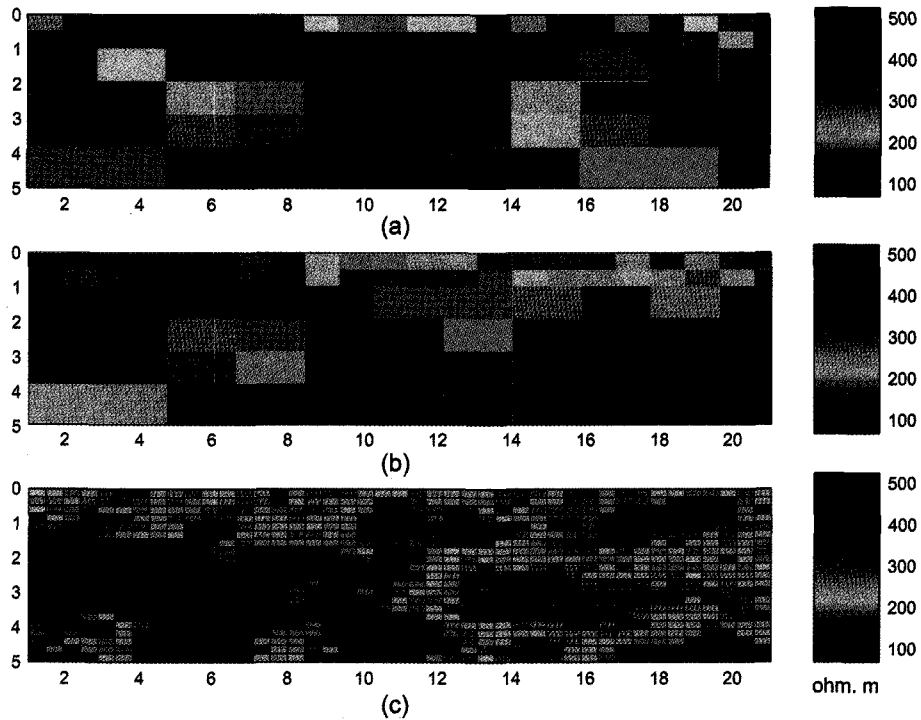


Fig. 14. Some representations from marginal PDF of posterior PDF. (a) Maximum PDF resistivity section. (b) Mean resistivity section of marginal posterior PDF. (c) True resistivity section plotted to be compared.

차원이 매우 높은 다중적분을 수행하여야 각각의 변수에 대한 추정을 할 수 있다. 이를 그대로 적분하는 것은 불가능하므로 본 연구에서는 MCMC 방법을 적용하였다. 특히 사전확률분포를 샘플링하여 Gibbs 방법에 적용하므로써 효율성을 더욱 높였다.

마지막으로 MCMC를 통해 주변화 과정을 거친 후, 얻어지는 각각의 변수에 대한 불확실성 분석을 수행하게 된다. 베이저안 역산으로부터 얻어지는 변수의 확률분포는 매우 많은 정보를 포함하고 있으며 이를 언제든지 원하는 형식으로 표현할 수 있다. 본 연구에서는 각 블록의 평균이나 최대값 등으로 표현하여 참 값과 비교하였다.

베이저안 역산은 다양한 정보를 통합(integration)할 수 있는 효율적인 방법이며, 각 단계를 통해 변수들이 어떻게 새로운 정보에 의해 변형되고 있는가를 체크할 수 있는 강력한 도구이다. 최근 들어 연구대상 지역에 대한 집중적인 지구물리학적 조사가 이루어지는 경향이 많은데, 본 연구가 이런 경향을 잘 소화할 수 있는 방안을 제시할 수 있을 것으로 생각된다.

사 사

본 연구는 기상청 기상연구소 특정사업 “지진예지 및 지진해일 예측기술 개발”의 연구비 지원으로 이루어졌으며, 이에 사의를 표합니다.

참고문헌

오석훈, 권병두, 1997, Simulate Annealing을 이용한 전기비저항 자료의 역산: 한국자원공학회지, 34, 285-293.
 Deutsch, C. V., and Journel, A. G., 1992, GSLIB: Geostatistical Software Library and User's Guide: Oxford University Press, New York.
 Duijndam, A. J. W., 1988a, Bayesian estimation in seismic inversion, part i: Principles: Geophys. Prosp., 36, 878-898.
 Duijndam, A. J. W., 1988b, Bayesian estimation in seismic inversion, part ii: Uncertainty analysis: Geophys. Prosp., 36, 899-918.
 Gilks, W. R., Richardson, S., and Spiegelhalter, D. J., 1996, Markov Chain Monte Carlo in Practice: Chapman & Hall/CRC.
 Glacken, I., 1996, Change of support by direct conditional block simulation: Master's thesis, Stanford University., Stanford, CA.
 Geman, S., and Geman, D., 1984, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images: IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-6, 721-741.
 Gomez-Hernandez, J. A., 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, *Colorado School of Mines*, Golden, CO.
- Hastings, W. K., 1970, Monte Carlo sampling methods using Markov chains and their applications: *Biometrika*, **57**, 97-109.
- Isaaks, E. H., and Srivastava, R. M., 1989, An Introduction to Applied Geostatistics: *Oxford 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.
- Lines, L. R., and Treital, S., 1984, Tutorial: A review of least-squares inversion and its application to geophysical problems: *Geophys. Prosp.*, **32**, 159-186.
- Loredo, T. J., 1990, From Laplace to Supernova SN 1987A: Bayesian Inference in Astrophysics, Reprinted from Fougere, P. F. ed., *Maximum Entropy and Bayesian Methods*, *Kluwer Academic Publishers*, The Netherlands, 81-142.
- 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.
- MPI Forum, 1995, MPI: A Message-Passing Interface Standard: *University of Tennessee*, Knoxville, TN.
- Oh, S.-H., and Kwon, B.-D., 2001, Geostatistical approach to bayesian inversion of geophysical data: Markov chain Monte Carlo method: *Earth Planets Space*, **53**, 777-791.
- Pacheco, P. S., 1997, *Parallel Programming with MPI*: *Morgan Kaufmann Publisher, Inc.*, San Francisco.
- Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: *Elsevier*, The Netherlands.