

# Inversion of Schlumberger resistivity sounding data from the critically dynamic Koyna region using the Hybrid Monte Carlo-based neural network approach

S. Maiti<sup>1</sup>, G. Gupta<sup>1</sup>, V. C. Erram<sup>1</sup>, and R. K. Tiwari<sup>2</sup>

<sup>1</sup>Indian Institute of Geomagnetism (DST), Navi Mumbai-410218, India

<sup>2</sup>National Geophysical Research Institute (CSIR), Hyderabad-500007, India

Received: 26 July 2010 – Revised: 14 January 2011 – Accepted: 17 February 2011 – Published: 9 March 2011

**Abstract.** Koyna region is well-known for its triggered seismic activities since the hazardous earthquake of  $M = 6.3$  occurred around the Koyna reservoir on 10 December 1967. Understanding the shallow distribution of resistivity pattern in such a seismically critical area is vital for mapping faults, fractures and lineaments. However, deducing true resistivity distribution from the apparent resistivity data lacks precise information due to intrinsic non-linearity in the data structures. Here we present a new technique based on the Bayesian neural network (BNN) theory using the concept of Hybrid Monte Carlo (HMC)/Markov Chain Monte Carlo (MCMC) simulation scheme. The new method is applied to invert one and two-dimensional Direct Current (DC) vertical electrical sounding (VES) data acquired around the Koyna region in India. Prior to apply the method on actual resistivity data, the new method was tested for simulating synthetic signal. In this approach the objective/cost function is optimized following the Hybrid Monte Carlo (HMC)/Markov Chain Monte Carlo (MCMC) sampling based algorithm and each trajectory was updated by approximating the Hamiltonian differential equations through a leapfrog discretization scheme. The stability of the new inversion technique was tested in presence of correlated red noise and uncertainty of the result was estimated using the BNN code. The estimated true resistivity distribution was compared with the results of singular value decomposition (SVD)-based conventional resistivity inversion results. Comparative results based on the HMC-based Bayesian Neural Network are in good agreement with the existing model results, however in some cases, it also provides more detail and precise results, which appears to be justified with local geological and structural de-

tails. The new BNN approach based on HMC is faster and proved to be a promising inversion scheme to interpret complex and non-linear resistivity problems. The HMC-based BNN results are quite useful for the interpretation of fractures and lineaments in seismically active region.

## 1 Introduction

The dynamically active Koyna region has been one of the focuses of colossal research activities in India ever since the killer earthquake of  $M = 6.3$  occurred around the Koyna reservoir on 10 December 1967 (Gupta et al., 1969, 1972, 2002; Talwani, 1997). The hazardous seismic activities in and around the region have been continuing since then, though sporadically. The layered structures in and around the region have impacted on the near surface distribution pattern of electrical properties. The modeling and interpretation of Direct Current (DC) resistivity sounding in this region, therefore, assume a special significance to understand the inhomogeneous infiltrations of fluids through porous media and geologically weak zones, such as faults and fractured zones, near the sub-surface area. The DC resistivity sounding method is one of the most popular methods that has been invariably applied for solving hydrological, geothermal, environmental and engineering problems (Ekinici and Demirci, 2008; Singh et al., 2010). In this method, current is induced directly into the ground through a pair of current electrodes and resulting voltage difference is measured between a pair of potential electrodes. The method provides the apparent resistivity distribution against depth. The depth of penetration of electrical signal is generally found to be approximately one-third of the distance between the electrode separations.



Correspondence to: S. Maiti  
(saumen\_maiti2002@yahoo.co.in)

The estimation of true resistivity distribution against depth from the apparent resistivity data essentially lead to solving the inverse problem. Further, relation between the observed “apparent resistivity” and the model parameters (“true resistivity” and “layer thickness”) is non-linear. Forward mathematical models are generally used to relate the measured/observed data (here apparent resistivity) to desired model parameters (true resistivity and layer thickness). Essentially, the forward modeling is a process of predicting results of measurements on the basis of some general principle or model and specific conditions relevant to the problem. Inverse modeling, in contrast starts with the data and a general principle or a model, in which the model parameters are estimated by minimizing the error/misfit function set up between the data and model parameters (Menke, 1984). During the past decades researchers have made several attempts to solve the resistivity inverse problems (Ghosh, 1971; Zohdy, 1989; Macias et al., 2000; Qady and Ushijima, 2001; Singh et al., 2005; Ekinici and Demirci, 2008). However, these approaches have some common drawbacks. These algorithms critically depend on the initial parameter chosen for it. The general class “Monte Carlo” e.g., genetic algorithm and simulated annealing techniques have proven to be useful whilst a good starting model is not available (Rubinstein, 1981; Kirkpatrick et al., 1983; Horne and MacBeth, 1994) for offering global solution. These methods are also computationally expensive and sometimes found unfeasible for processing voluminous amount of data. Thus it is imperative to search for a more robust cost effective approach for solving non-linear resistivity inverse problems.

In the recent past, Artificial Neural Network (ANN)-based techniques have been widely applied to solve non-linear problems in various branches of geophysics (Van der Baan and Jutten, 2000; Poulton, 2001). For example: (1) for seismic event classification (Dystart and Pulli, 1990), (2) well log analysis (Aristodemou et al., 2005; Maiti et al., 2007; Maiti and Tiwari, 2007, 2009, 2010b), (3) first arrival picking (Murat and Rudman, 1992), (4) earthquake prediction (Feng et al., 1997), (5) inversion (Raiche, 1991; Devilee et al., 1999), (6) parameter estimation in geophysics (Macias et al., 2000), (7) prediction of aquifer water level (Coppola Jr. et al., 2005), (8) magneto-telluric data inversion (Spichak and Popova, 2000), (9) magnetic interpretations (Bescoby et al., 2006), (10) signal discrimination (Maiti and Tiwari, 2010a), (11) modeling (Sri Lakshmi and Tiwari, 2009), (12) DC resistivity inversion (Qady and Ushijima, 2001; Lampinen and Vehtari, 2001; Singh et al., 2005, 2006, 2010).

There are, however, several limitations in conventional neural network approaches also (Bishop, 1995; Maiti and Tiwari, 2009). One of the major limitations is that the network is trained by maximizing a likelihood function of the parameters or equivalently minimizing an error function in order to obtain the best set of parameters starting with an initial random set of parameters. Sometimes a regularization term with an error function is used to prevent over-fitting. In that case,

a complex model can well be used to fit the training data but it does not necessarily guarantee the smaller errors in the unseen data (Bishop, 1995). This is due to the fact that ensuing procedure does not take into account of uncertainty in the estimation of parameters (Bishop, 1995; Nabney, 2004).

In the present work, we developed a Bayesian neural network approach via a powerful Hybrid Monte Carlo approach (Bishop, 1995; Nabney, 2004; Maiti and Tiwari, 2009, 2010a) for estimating true resistivity and layer thickness precisely from the apparent resistivity data acquired from Chiplun to Patan (26 number of Schlumberger sounding profile data) of Koyna region (Latitude 17.25° N–17.52° N, Longitude 73.5° E–74.00° E; profile length is about 50 km roughly in E-W direction), Maharashtra (Fig. 1a–b). We compared the new HMC-based BNN results with conventional singular value decomposition (SVD) based damped least-squares solution. Our results suggest that the HMC-based BNN technique is comparatively superior to the conventional SVD-based technique in a sense that the present technique takes care of the problem of uncertainty analysis in a natural way. Thus, besides introducing a novel HMC-based BNN approach for solving non-linear inverse problem, the analyses have brought out precise distribution of true resistivity which appears to be quite significant, especially for constraining the model in seismically active region Koyna.

## 2 Geological setting of the study area

The study area is located on the western side of Peninsular India (Fig. 1a). It is covered by Deccan Traps, a Continental Flood Basalt (CFB). The CFB provinces are believed to be intimately related to passage of continental masses over static mantle plumes (hot spot), and to the subsequent continental rifting triggered by plume activities (Morgan, 1972). It is believed that Deccan Traps were erupted during the separation of the Seychelles micro-continent from Indian, 50–60 Ma (end of Cretaceous-early Tertiary) (Duncan and Pyle, 1988) during a rifting event induced by the northward movement of the Indian subcontinent over the Late Cretaceous manifestation of Reunion plume. A remarkable feature of Deccan volcanism is the horizontal flow throughout the region. There are many lava flows with varying thickness ranging from 10 to 160 m. Variations in thickness of the trap flows are attributed mainly to pre-Trip topography. These massive compact lava sequences have low permeability, which implies that there is significant migration of water through faults, fractures, columnar jointing and vesicles. A general regional strike varies from NS to NNW. Geological observations indicated the existence of a NNE trending strike slip fault, passing through Koyna (Gupta and Rastogi, 1974; Gupta et al., 1969, 2002; Talwani, 1997). Several lineament features have been identified in the Deccan trap region of western India. Structural shear zones are speculated to exist along the lineament, based on study of imageries and some

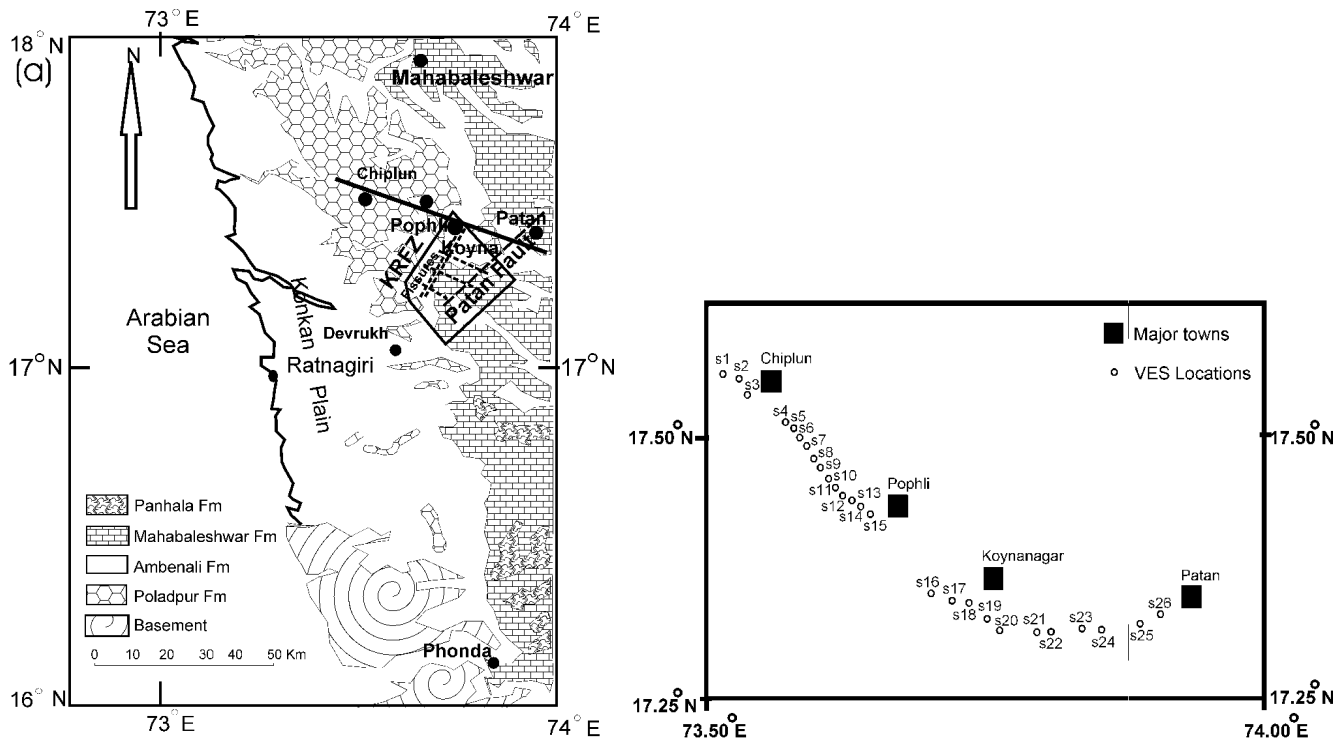


Fig. 1. (a) Geological map of the study area, (b) location map of the Schlumberger resistivity profile.

field observations (Athavale and Mohan, 1976). The seismic activities in the western region are bounded by the Koyna River Fault Zone (KRFZ) that dips steeply towards the West, and to the East by the NE-SW trending Patan Fault (Fig. 1a). The area between KRFZ and Patan Fault is intersected by a number of NW-SE fractures zones.

The present resistivity profile covers the major parts of Koyna region, starting from Chiplun (Latitude 17.51° N, Longitude 73.52° E) to Patan (Latitude 17.36° N Longitude 73.90° E) with maximum current electrode separation (AB) of 200 m. The measurements were taken using the SSR-MP-AT instrument (manufactured by IGIS, Hyderabad) during March 2009. In order to improve the quality of data, the resistivity meter performed a minimum of three stacks for each data point. The locations of the sounding points have been marked in Fig. 1b. The existing geophysical and geological studies show that the shallow distribution of the resistivity is important for understanding the fault and lineament pattern of the area. The exposed basement is fractured at many places of the area which could be speculated as the shear zones. In view of the established utility of resistivity method in locating and demarcating the fractured and weathered zones, especially in the seismically active Koyna region, it was thought worth to conduct electrical resistivity survey over the area to get a precise true resistivity value against depth.

### 3 Methods

#### 3.1 Forward modeling

Schlumberger sounding is the most popular sounding method in DC resistivity survey. For Schlumberger sounding, the relationship between the apparent resistivity ( $\rho_a$ ) and the layer parameters (e.g. layer thickness, layer true resistivity) can be expressed by an integral equation considering an earth model consisting of homogeneous and isotropic layers. Following (Koefoed, 1970) we write the equation as follows

$$\rho_a(s) = s^2 \int_0^\infty T(\lambda) J_1(\lambda s) \lambda d\lambda \tag{1}$$

Where,  $s$  is half of the current electrode spacing ( $AB/2$ ) in Schlumberger electrode configuration,  $J_1$  denotes the first-order Bessel function of the first kind and  $\lambda$  denotes the integral variable. Following Koefoed (1970) we write recurrence relationship of the resistivity transform function,  $T(\lambda)$  as,

$$T_i(\lambda) = \frac{T_{i+1}(\lambda) + \rho_i \tanh(\lambda h_i)}{[1 + T_{i+1}(\lambda) \tanh(\lambda h_i) / \rho_i]}, \quad i = n - 1, \dots, 1 \tag{2}$$

Where,  $n$  denotes the number of layers,  $\rho_i$  and  $h_i$  are the true resistivity and thickness of the  $i$ -th layer, respectively.

### 3.2 Inversion scheme

The standard linearized inversion approach to solving the non-linear inverse problems in geophysics are largely based on iterative processes. The forward model is developed based on the specific relation between physical models to the observed data. “Inversion” processes update the model parameter at each step to best fit the observed data. However, the inversion of the resistivity sounding data is an ill-posed problem. This is due to the fact that the contradictory information on model parameter cannot be assessed due to lack of information. Therefore, small changes in the data may lead to large changes in the model. Successful optimization depends heavily on choosing the correct initial model. The problem may be reduced by introducing damping into the system of equations (Roy, 1999). This resulted in a solution of damped least-squares which we can write following (Menke, 1984)

$$\Delta m = (\mathbf{G}^T \mathbf{G} + \beta^2 \mathbf{I})^{-1} \mathbf{G}^T \Delta d \quad (3)$$

Where  $\Delta m$  is the parameter correction vector;  $\Delta d$  is the data difference vector;  $\mathbf{G}$  is the Jacobian matrix containing partials derivative of data with respect to the initial model parameter.  $\mathbf{I}$  is the identity matrix, and the term  $\beta$  is called damping factor which is a scalar quantity, actually controls both speed of convergence and solution. This solution is also known as Tikhonov regularization (Levenberg, 1944; Marquardt, 1963; Menke, 1984).

### 3.3 Singular value decomposition (SVD)

Singular value decomposition (SVD) technique is popular in many areas of geophysical inversion. It is a very useful technique for small scale geophysical inverse problem especially, for unstable, rank-deficiency problem. SVD provides numerically stable results in addition to the information related to model parameter resolution and covariance analysis (Meju, 1994).

We factorize an  $n \times n$  or  $n \times m$  Jacobian matrix  $\mathbf{G}$  in the above Eq. (3) as follows

$$\mathbf{G} = \mathbf{U} \mathbf{Q} \mathbf{L}^T \quad (4)$$

Where for  $n$  data and  $m$  parameters,  $\mathbf{U}(n \times m)$  and  $\mathbf{L}(m \times m)$  are two orthonormal matrix, containing respectively the data space and the parameter space eigenvectors and  $\mathbf{Q}$  is a  $(m \times m)$  diagonal matrix containing at most  $r$  non-zero eigenvalues of  $\mathbf{G}$ , with a condition  $r \leq m$ . These diagonal entities in matrix  $\mathbf{Q}(\alpha_1, \alpha_2, \dots, \alpha_p)$  are called singular values of  $\mathbf{G}$ . We write SVD-based damped least squares solution,

$$\Delta m = (\mathbf{L} \mathbf{Q}^2 \mathbf{L}^T + \beta^2 \mathbf{I})^{-1} \mathbf{L} \mathbf{Q} \mathbf{U}^T \Delta d \quad (5)$$

We get the form by adding the damping factor to the diagonal elements

$$\begin{aligned} & (\mathbf{L} \mathbf{Q}^2 \mathbf{L}^T + \beta^2 \mathbf{I}) \\ & = (\mathbf{L} \text{diag} \{ \alpha_j^2 \} \mathbf{L}^T + \beta^2 \mathbf{I}) = \mathbf{L} \text{diag} (\alpha_j^2 + \beta^2) \mathbf{L}^T \end{aligned} \quad (6)$$

We write the inverse of the Eq. (6) as follows,

$$\left( \mathbf{L} \text{diag} \{ \alpha_j^2 + \beta^2 \} \mathbf{L}^T \right)^{-1} = \mathbf{L} \text{diag} \left\{ \frac{1}{\alpha_j^2 + \beta^2} \right\} \mathbf{L}^T \quad (7)$$

Substituting the Eq. (7) in Eq. (5) we obtain

$$\Delta m = \mathbf{L} \text{diag} \left\{ \frac{1}{\alpha_j^2 + \beta^2} \right\} \mathbf{L}^T \mathbf{L} \mathbf{Q} \mathbf{U}^T \Delta d \quad (8)$$

We obtain the parameter correction vector as:

$$\Delta m = \mathbf{L} \text{diag} \left\{ \frac{\alpha_j}{\alpha_j^2 + \beta^2} \right\} \mathbf{U}^T \Delta d \quad (9)$$

The Eq. (9) provides damped least-squares solution via the SVD. Usually, in initial iteration, the damping factor is set to be a large positive value while making the full use of steepest descent method. Subsequently, at each iteration the damping factor is multiplied by a factor less than unity so that the least-squares method dominates near the solution (Meju, 1994). Following Arnason and Hersir (1988) we determine the damping factor as follows:

$$\beta = \alpha_W \Delta c^{\frac{1}{W}} \quad (10)$$

Where  $W$  is the test number for the damping factor at any iteration,  $\alpha$  is the parameter eigenvalue and the term  $\Delta c$  is given by

$$\Delta c_r = \frac{(c_{r-1} - c_r)}{c_{r-1}} \quad (11)$$

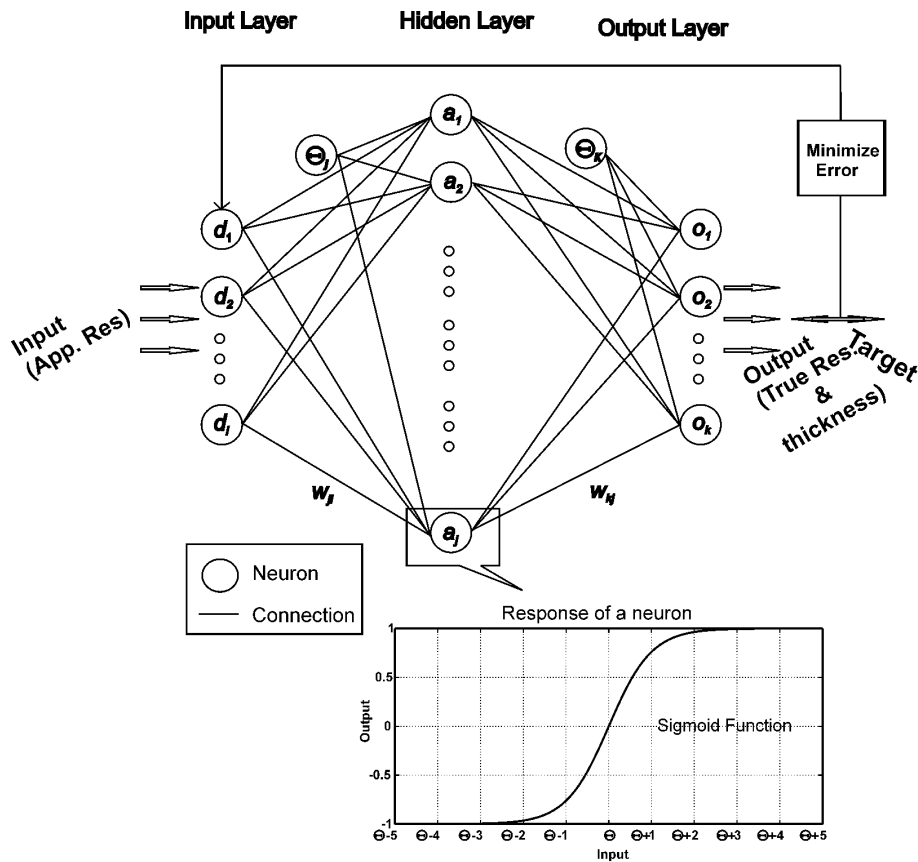
Where,  $c_{r-1}$  is the misfit value obtained at previous iteration and  $c_r$  is the misfit found at the current iteration. In this study, Eqs. (10) and (11) were used to set the damping factor in each iteration.

## 4 Bayesian neural network approach

In Bayesian neural network approach, we establish the forward relationship between models predicted by forward-modeling with the data (apparent resistivity) through a function as

$$d = f(x) + \varepsilon \quad (12)$$

Where  $f$  is the function relating between model space and data space,  $\varepsilon$  is error,  $d$  is data, and  $x$  is model. It is worthwhile to mention that  $f$  is non linear (discussed above) to establish data-model relationship, so direct estimation of model is not possible. A common way of inverting for model  $x$  in Eq. (12) is via iterative least squares method. However this does not provide uncertainty measures. Instead, to solve the Eq. (12) in Bayesian sense, we considered sufficient number of representative realizations samples (model /data pair)



**Fig. 2.** Layout of MLP with a three-layer neural network with “ $d$ ” representing input, subscript “ $i$ ” representing the number of nodes in the input layer  $w_{ji}$  representing the connection weight between the  $i$ -th node in the input layer and the  $j$ -th node in the hidden layer and  $w_{jk}$ , representing the connection weight between the  $j$ -th node in the hidden layer and the  $k$ -th node in the output layer.  $\Theta_j$  and  $\Theta_k$  are bias vectors for the hidden and the output layer. When the sum of the argument of a neuron is comparable to the threshold value  $\Theta_j$ , the sigmoid function squashes linearly, otherwise it saturates with value  $+1$ ;  $-1$  gives non-linearity for non-linear mapping between an input and an output space. Sigmoid function plays key role for mapping non-trivial geo-electrical problem.

from a finite data set  $s = \{d_k, x_k\}_{k=1}^N$  conditioned on the explicit limit of resistivity data. So we write the equation now as

$$x = f_{NN}(d; w) \tag{13}$$

Where,  $x$  is model desired,  $f_{NN}$  is the output predicted by the network and  $w$  is network weight parameter. In standard/conventional approach of neural network for solving Eq. (12), often, regularization is included to modify total misfit function

$$E(w) = \mu E_S + \lambda E_R \tag{14}$$

Where, for example,  $E_R = \frac{1}{2} \sum_{i=1}^R w_i^2$  is known as weight-decay regularization which favors small values for the network weight,  $R$  is total number of weights and biases in the network,  $\lambda$  and  $\mu$ , which control other parameters (synaptic weight and biases), are known as hyper parameters (Maiti and Tiwari, 2010b).  $E_S$ , least square error measure function,

can be written in the form of  $E_S = \frac{1}{2} \sum_k^N \{x_k - o_k(d_k; w_k)\}^2$ ,

where  $x_k$  and  $o_k$  are respectively the target/desired and the actual output at each node in the output layer. It is important to note that forward functions used in node are nonlinear tan sigmoid in nature eases to solve the non trivial problem (Fig. 2). In the traditional approach, the training of a network starts with an initial set of weights and biases and ends up with the single best set of weights and biases given the objective function is optimized.

In the Bayesian neural network approach, a suitable prior distribution, say  $P(w)$  of weights is considered before observing the data instead of single set of weights. Using Bayes’ rule, a posteriori probability distribution for the weights, say  $P(w|s)$  can be given as (Bishop, 1995; Khan and Coulibaly, 2006),

$$P(w|s) = \frac{P(s|w)P(w)}{P(s)} \tag{15}$$

Where,  $P(s|w)$  is the data set likelihood function, and the denominator,  $P(s)$  is a normalization factor. The denominator  $P(s)$  is intractable so direct estimation of posterior  $P(w|s)$  is not possible. Using the rules of conditional probability, the distribution of outputs for a given input vector,  $x$  can be written in the form (Bishop, 1995; Khan and Coulibaly, 2006),

$$P(x|d, s) = \int P(x|d, w) P(w|s) dw. \quad (16)$$

The major problem in Bayesian computation is evaluating integrals for posterior pdf of weights (Eq. 15) and network output (Eq. 16). Fortunately numerical methods Monte Carlo combine with Markov Chain (MCMC) sampling based method plays an important role for evaluating posterior integrals. The Eq. (16) can be approximated as

$$P(x|d, s) = \frac{1}{N} \sum_{n=1}^N P(x|d, w_n). \quad (17)$$

Where  $\{w_m\}$  represents a MCMC sample of weight vectors obtained from the distribution  $P(w|s)$  and  $N$  is the number of points  $w$  sampled from  $P(w|s)$ .

#### 4.1 Markov Chain Monte Carlo (MCMC)/Hybrid Monte Carlo (HMC)

It is well known that simple gradient descent algorithm can get trapped at shallow local minima. One way of overcoming this problem is to define the error surface in terms of a Hamiltonian statistical mechanism that accounts for the approximation errors and the momentum term of each trajectory. This is the basis of the Hybrid Monte Carlo algorithm/Markov Chain Monte Carlo algorithm (Duane et al., 1987). In this algorithm, each trajectory is updated by approximating the Hamiltonian differential equations by a leapfrog discretisation scheme. Markov Chain Monte Carlo algorithm draws an independent and identically distributed (i.i.d) sample  $\{w^{(i)}; i = 1, 2, \dots, N\}$  from the target distribution  $P(w|s)$ ; in Bayesian sense, posterior probability distribution of target network parameter. Markov process forms a sequence of “state” to draw samples from posterior pdf. It is noted that the chain converges to  $P(w|s)$  given enough space. The states are represented by a particle in the high dimensional network parameter space whose positions are defined by  $q \in R^w$ . So the Eq. (15) can be written in the form of  $\pi(q) \propto \exp\{-E(q)\}$ , where  $E(q)$  is potential energy functions/cost function for optimization problem. By introduction of momentum variables  $p$  with corresponding kinetic energy functions  $V(p) = \frac{1}{2} \sum_{i=1}^N p_i^2$  to efficiently explore large region of phase-space by simulating the Hamiltonian dynamics in fictitious time. The particle mass and Boltzmann constant  $k_B$  can be re-scaled to unity. Full Hamiltonian energy function on a fictitious phase space is

$$H(q, p) = E(q) + V(p) \quad (18)$$

The canonical distribution of Hamiltonian is

$$\pi(q, p) = \frac{1}{Q_H} \exp\{-H(q, p)\} \quad (19)$$

It is clear that if we sample  $(q, p)$  from the distribution  $\pi(q, p) = \frac{1}{Q_H} \exp\{-H(q, p)\}$ , then the marginal distribution of  $q$  is exactly the target distribution  $\pi(q)$ . Here,  $\pi$  is a generic symbol.

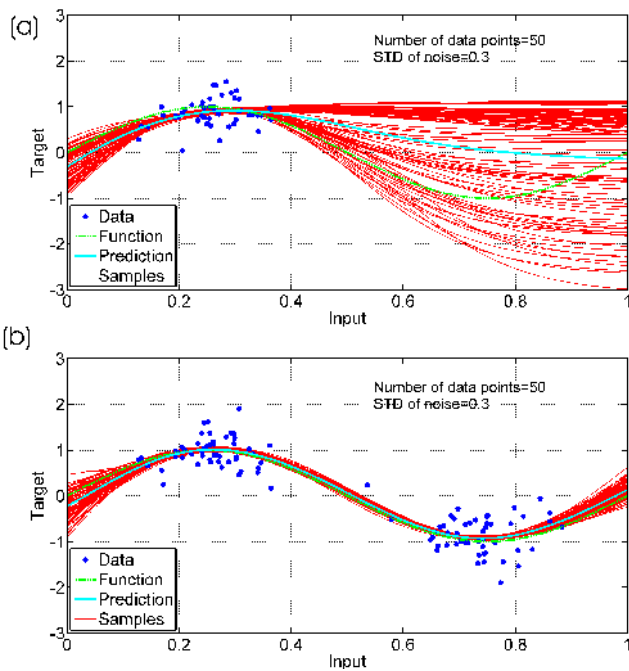
Practically pure Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) is proved to be very slow because the method makes no use of gradient information. The proposed algorithm here is based on Hybrid Monte Carlo (HMC) algorithm for sampling from target distribution which makes the use of the gradient information. The HMC based algorithm is a sampling algorithm that takes into consideration certain gradient information. The algorithm follows the following sequence of steps once a step size  $\theta$  and the number of iterations  $L$  have been decided upon: (i) firstly randomly choose a direction  $\tau$ :  $\tau$  can be either  $-1$  or  $+1$  with the probability 0.5 simulate the dynamics forward or backward in time. The transition probability matrix satisfies microscopic reversibility which implies that probability of these two transitions from  $q_j$  to  $q_i$  or from  $q_i$  to  $q_j$  be same at all times and each pairs of points maintains a mutual equilibrium. (ii) Secondly carry out the iterations: starting with the current state  $[q, p] = [q(0), p(0)]$  of energy  $H$ , where  $p$  is a momentum term which is randomly evaluated at each step, let algorithm perform  $L$  steps with a step size of  $\theta$  resulting in the candidate state,  $[w^*, p^*]$  with energy  $H^*$ . (iii) Thirdly the candidate state is accepted with usual metropolis probability of acceptance,  $\min\{1, \exp[-(H^* - H)]\}$ , where  $H(\cdot)$  is the Hamiltonian energy. If the candidate state is rejected then the new state will be the old state. The above three steps, in essence, describe how the sampling is done from posterior distribution of network parameter so that the summation of Eq. (17) can be accomplished for the optimization of the network. The momentum term  $p$  can be randomly generated or it can be changed dynamically at each step and there are different ways of doing this (Duane et al., 1987; MacKay, 1992; Neal, 1996; Bishop, 1995). The sets of weights are thus selected or rejected according to the three steps above and the numbers of samples that are wished to be retained are the number of weights retained. For each set of weights there is a corresponding network output.

#### 4.2 Synthetic example

We have carried out numerical experiment following Nabney (2004). For example, to illustrate the application of the Bayesian techniques to a “regression” problem, we consider a one-input one-output example involving data generated from the a smooth sine function of the form of

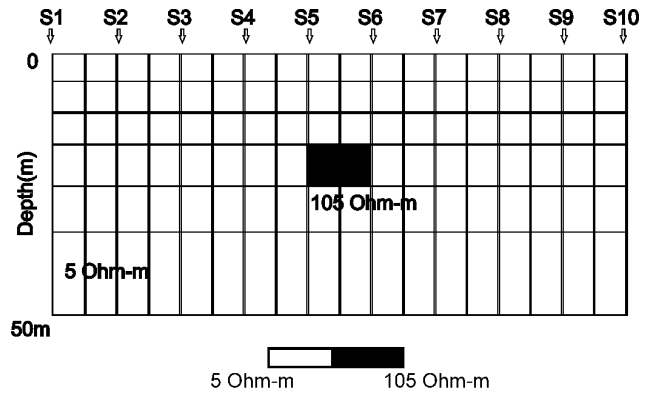
$$f(x) = 0.25 + 0.07 \sin(2\pi x) \quad (20)$$

with additive Gaussian noise having standard deviation of  $\delta = 0.1$ . Value for  $x$  were generated by sampling an isotropic



**Fig. 3.** (a) The graph shows the simulation of a synthetic underlying function using 100 numbers of data points in the presence of noise via Hybrid Monte Carlo (HMC) method when the data are sampled at point 0.25. The standard deviation (STD) of the noise is 0.3. (b) Examples when the data are sampled at point 0.25 and 0.75 and STD of noise distribution is 0.1.

Gaussian mixture distribution. Figure 3a–b shows trained BNN approximation which corresponds to mean of the predictive distributions. Error bar represents one standard deviations ( $\pm 1\delta$ ) of the predictive distribution. This predictive distribution allows us to provide error bars for the network output instead of just a single answer. For the synthetic experiment we have chosen following the parameters: (i) the number of node used in a single hidden layer is 10. (ii) The number of training samples used in the experiments is 50. (iii) The initial prior hyperparameter values are  $\lambda = 0.01$  and  $\mu = 50.0$ . (iv) The tolerance for the weight optimization is set to a very low value ( $10^{-7}$ ). This is because the Gaussian approximation to the weight posterior depends on being at a minimum of the error function (Nabney, 2004). The detailed analysis of sensitivity of the hyperparameters is given in the recent work of Maiti and Tiwari (2010a). The initialization of the model parameters is performed by a distribution of model parameters. The initial values of model parameters of MLP (Multi-Layer Perceptron) (synaptic weight and biases) are formed by Gaussian prior distribution of zero mean and inverse variance  $\lambda$  (also known as regularization coefficient or prior hyperparameters). To define objective function in Bayesian framework, an error model for the data likelihood is required. It is assumed that target data is formed from a smooth function with additive zero mean Gaussian noise.



**Fig. 4.** Forward model used to create the resistivity synthetic training and testing for two-dimensional resistivity model.

Accordingly hyper-parameter  $\mu = 50$  is estimated for both hidden and output layer weights. After defining prior and likelihood functions, posterior has been estimated by Bayes' rule.

## 5 Model initiation and implementation

### 5.1 Hidden layers, connection weights and output

In Bayesian neural network modeling we do not necessarily need to estimate the optimal number of the weights (hidden layer) to have a good generalization (Bishop, 1995). However, in the present problem we find by trial and error that a single hidden layer with twenty five individual nodes is appropriate.

### 5.2 Number of training samples

We considered a total of 1326 representative input/output pairs within the constrained bounds of the plausible geological solutions for 1-D BNN training. Following Van der Bann and Jutten (2000), the number of training samples is kept more than the total internal variables. For two-dimensional inversion, we considered an embedded anomalous mass of resistivity 105 Ohm-m (Fig. 4) to generate a synthetic training set required for training the network. A collinear Schlumberger was set up in sounding mode with half of the current electrode spacing 100 m. The position of the anomalous mass was changed and moved to all the model mesh elements. In this process, we allowed each element in the mesh to be either resistive or conductive. The two-dimensional data set was generated using the finite element forward modeling (Dey and Morrison, 1979; Uchida, 1991) scheme in which a total of eighty training sets were generated for each profile.

### 5.3 Input data scaling and model parameterization

Normalization of the raw data (input/output), before presenting it to the network is crucial to avoid saturation in order to map nontrivial problem. Hence we scaled all the input/output pair values between 0 and 1 [−1 and +1] by using a simple linear transformation algorithm (Poulton, 2001); normalized input =  $2 \times (\text{input} - \text{minimum input}) / (\text{maximum input} - \text{minimum input}) - 1$ . In view of computational simplicity (Bishop, 1995) and to avoid large curvature (Nabney, 2004), the initial values of model parameters (synaptic weight and biases) are formed by following a Gaussian prior distribution function of a zero mean and an inverse variance,  $\lambda$  (also known as regularization coefficient or prior hyperparameter). We use a single hyperparameter  $\lambda$ , for all the weights in a network. In order to define an objective function in the Bayesian framework, an error model for the data likelihood is required. Having assumed that the target data is formed from a smooth function with additive zero mean Gaussian noise, we estimated the hyperparameter  $\mu$  for both hidden and output layer weights.

### 5.4 Data division for model validation and testing

Over-fitting is one of the serious drawbacks in traditionally used Artificial Neural Network (ANN) modeling. In this case, the selected training set is memorized in such a way that ANN fits even the inherent deceptive noise present in the data. Therefore, the performance of the network is excellent only on the data set used for training but not on new data set. However, the Bayesian learning approach control “effective complexity” by considering many adjustable parameters and the parameter uncertainty is considered in form of probability distribution (Bishop, 1995; Nabey, 2004). It may be noted that the HMC-based BNN approach does not require any cross-validation test to prevent “over-fitting” (Bishop, 1995) because the underlying approach is equipped with strong mathematical soundness that takes care of such drawbacks naturally. Thus following the guidelines described earlier in the case of synthetic data modeling, we performed our training and simulation using all data sets.

### 5.5 Network sensitivity to correlated noise

In many geological/geophysical situations, we invariably observe some kind of deceptive/correlated noise, which dominates the field observations and corrupts the actual signal. In the present case, we do not have any precise idea about the “percentage” of noise present in the actual resistivity data. Assuming however, that there is some level of inescapable noise in the data, it would be prudent to test the robustness and the stability of the results. Broadly speaking, there are two types of noise in the data. One is uncorrelated noise or white noise and another is correlated or red noise. The red noise is incredibly deceptive because it correlates with

**Table 1.** Estimation of model parameter of synthetic data set corrupted with 5% correlated red noise.

Model parameter	True value	Estimated value SVD-based conventional method	Estimated value HMC-BNN method	Average uncertainty HMC-BNN method
$\rho_1$	90 Ohm-m	90.50 Ohm-m	89.70	
$\rho_2$	451 Ohm-m	450.00 Ohm-m	450.50	
$\rho_3$	112 Ohm-m	110.00 Ohm-m	112.25	
$\rho_4$	20 Ohm-m	22.10 Ohm-m	19.76	
$\rho_5$	893 Ohm-m	894.00 Ohm-m	892.63	
$\rho_6$	3.00 Ohm-m	4.00 Ohm-m	2.97	~0.30
$h_1$	0.83 m	1.0 m	0.85 m	
$h_2$	1.90 m	2.00 m	1.95 m	
$h_3$	9.10 m	10.00 m	9.00 m	
$h_4$	8.50 m	7.5 m	8.75 m	
$h_5$	10.40 m	12.00 m	10.50 m	

the data rendering unphysical appearances in the data structure. Statistically the red noise is followed by random walk model and often misread as data. The detailed procedure of generation and mixing of red noise for testing and experimental treatments can be found elsewhere (Maiti and Tiwari, 2010b). Here we have tested the trained network in presence of correlated red noise (Table 1). One can see from Table 1 that the network is able to fairly accurately estimate resistivity distribution in presence of correlated red noise (Table 1).

### 5.6 Uncertainty analysis

The uncertainty at the network output arise mainly due to (i) the intrinsic noise in the data embodies in the hyperparameter  $\mu$  and (ii) theoretical error described by the posterior distribution of the weight vector  $w$  (Bishop, 1995). The elements along the main diagonal of output covariance matrix shows the “variances” of the fluctuations about the mean of the Gaussian probability densities that characterizes the uncertainties. The off-diagonal elements show the extent to which these fluctuations are correlated (Tarantola, 1987). The average uncertainty results for the synthetic and real data inversion are given in Tables 1 and 2.

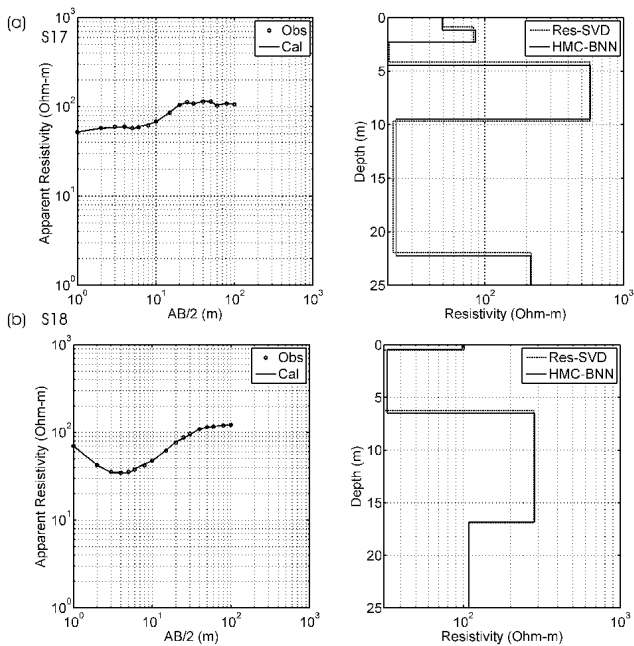
## 6 Modeling resistivity data from the Koyna region

After a successful test of the proposed algorithm on complex synthetic data series, we demonstrate the applicability/efficiency of the proposed inversion method on the actual DC resistivity data collected from the Koyna region, India (Fig. 1). The DC resistivity sounding with Schlumberger array was carried out at 26 VES stations as shown in Fig. 1b using electrode spacing starting from AB = 2 up to 200 m in successive steps. The field sites were chosen on the basis of the accessibility and applicability of the Schlumberger geoelectrical method in order to attain the objective.



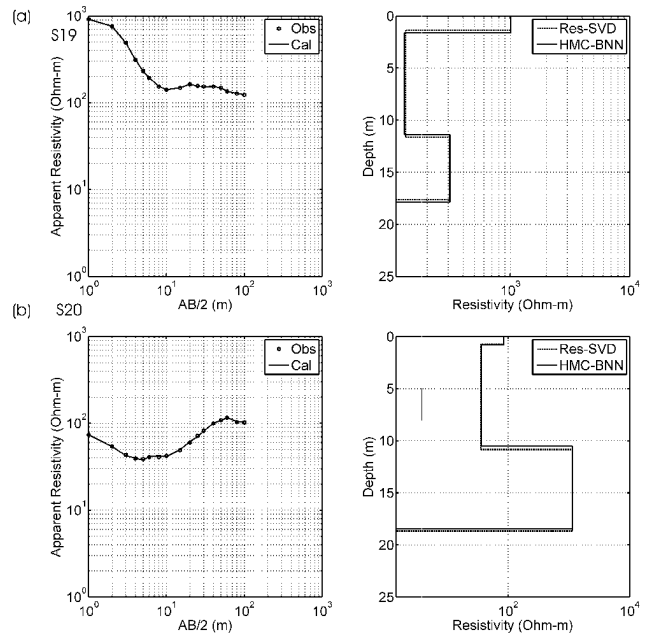
**Table 2.** Estimation of model parameter of real data of sounding point s16.

Model parameter	True value	Estimated value SVD-based conventional method	Estimated value HMC-BNN method	Average uncertainty HMC-BNN method
$\rho_1$	326.00	325.00	325.97	
$\rho_2$	74.60	73.23	75.01	
$\rho_3$	762.00	762.97	761.75	
$\rho_4$	5.82	5.80	5.83	
$\rho_5$	1500.00	1503.71	1498.78	$\sim 0.33$
$h_1$	0.90	0.95	0.93	
$h_2$	7.65	7.01	7.89	
$h_3$	6.71	6.89	6.75	
$h_4$	5.06	5.10	5.03	

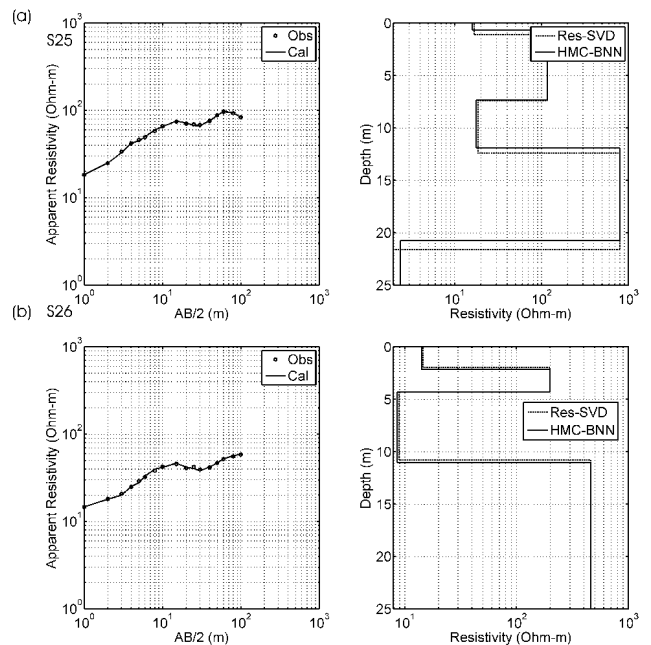


**Fig. 5.** The comparison of Hybrid Monte Carlo (HMC)-based Bayesian neural network (BNN) inversion results with singular value decomposition (SVD)-based conventional resistivity inversion results for soundings (a) s17 and (b) s18.

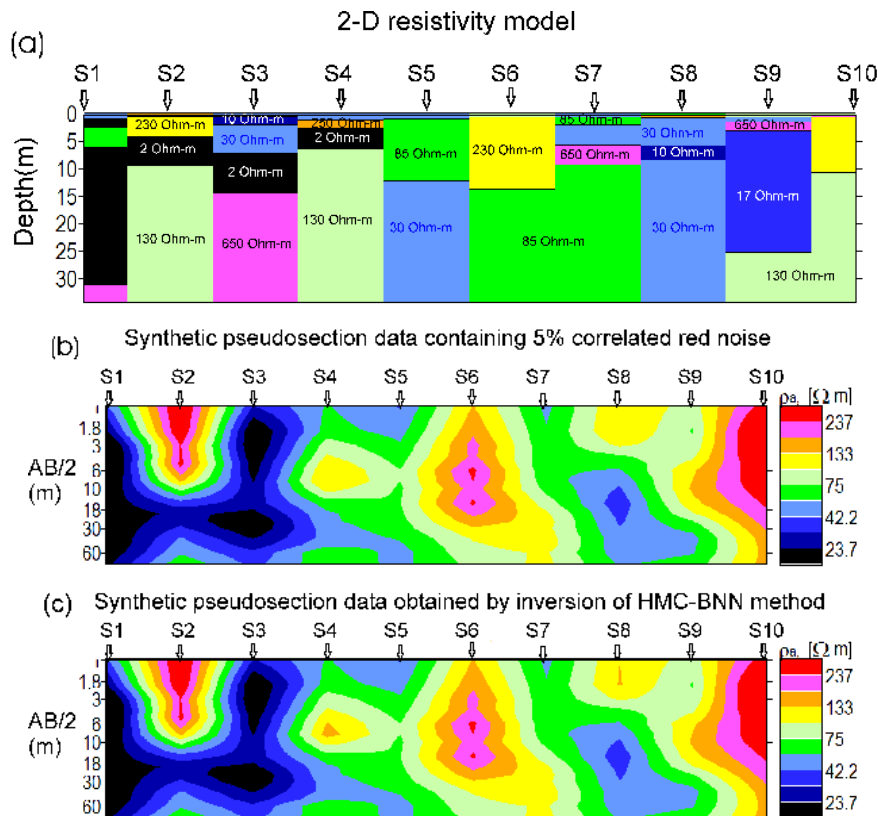
We inverted 26 number of obs DC resistivity sounding data using the new Hybrid Monte Carlo-based Bayesian Neural Network approach. One dimensional inversion results of some selected sounding station are shown (Figs. 5–7). The HMC-based BNN results are compared with the Singular Value Decomposition (SVD)-based conventional DC resistivity inversion results to demonstrate the performance of the new method on real data analysis. In Figs. 5–7, “Obs” denotes observed, and “Cal” denotes the calculated resistivity



**Fig. 6.** The comparison of Hybrid Monte Carlo (HMC)-based Bayesian neural network (BNN) inversion results with singular value decomposition (SVD)-based conventional resistivity inversion results for soundings (a) s19 and (b) s20.



**Fig. 7.** The comparison of Hybrid Monte Carlo (HMC)-based Bayesian neural network (BNN) inversion results with singular value decomposition (SVD)-based conventional resistivity inversion results for soundings (a) s25 and (b) s26.

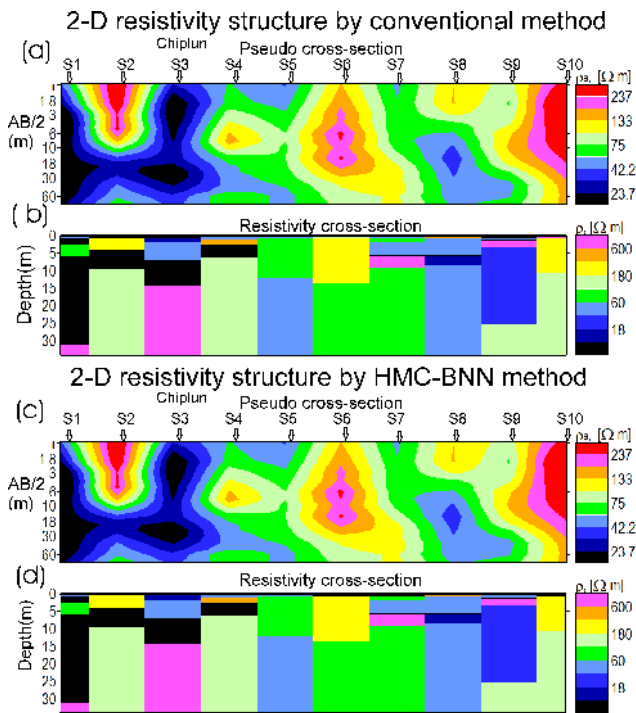


**Fig. 8.** Example to demonstrate two-dimensional (2-D) inversion of resistivity structure: (a) two-dimensional (2-D) model, (b) synthetic pseudosection data containing 5% correlated red noise, (c) synthetic pseudosection obtained by inversion of Hybrid Monte Carlo (HMC) based Bayesian neural network (BNN) approach.

values, “Res-SVD” denotes SVD based conventional resistivity inversion results, “HMC-BNN” denotes Hybrid Monte Carlo-based Bayesian Neural Network results. The new Hybrid Monte Carlo-based Bayesian Neural Network solution is well correlated except some mismatches in the inversion results of sounding point (SP) 25 (Fig. 7a). Overall, our results show that Hybrid Monte Carlo-based Bayesian Neural Network results are stable and well consistent with the SVD-based existing results. In addition to this, however, Bayesian Neural Network approach also provides uncertainty estimation, which is essential in the geophysical interpretations. In the present analysis we found an average uncertainty estimate of about  $\sim 0.30$ . It may be noted that the new BNN based approach avoid “oscillation” problem around the many local minima. The solution (mean of thousands simulations) is chosen from thousands of equally probable model (Fig. 3). The Hybrid Monte Carlo-based Bayesian Neural Network results lies within the 90% confidence interval (CI). We note that majority of the Schlumberger sounding curves in the area are either H-type or A-type. The apparent resistivity ( $\rho_a$ ) vs. half of the current electrode separation ( $AB/2$ ) on log-log graph suggests that there are three to four layered structured in this area (Figs. 5–7).

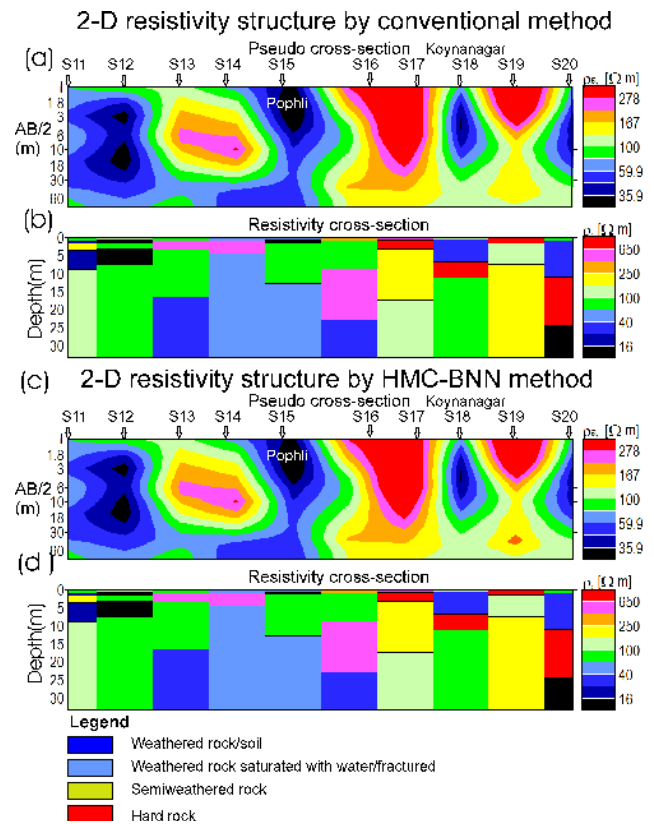
Two-dimensional inversion results suggest that the surface layer is seldom thick, being only in the range of 1–3 m. The surface layer is thinner over elevated land/hill and in some places it is absent too. Its resistivity value is much lower towards Eastern end of the profile (Near Patan, Latitude  $17.36^\circ$  N, Longitude  $73.90^\circ$  E) than the Western side (near Chiplun, Figs. 9–11). The second layer has a corresponding low resistivity which indicates that the layer is composed of weathered rock (mostly, laterite to basalt) possibly saturated with water. Its thickness varies from 10 to 20 m. Third layer is comparatively less resistive than crystalline basement rock. It is also observed that generally the rocks in the lower ground and hill slope are more weathered than those in the high ground, although there are some exceptions too.

One-dimensional sounding result is useful to delineate the layer resistivity and thickness. It is noteworthy that the one dimensional results provided a good starting model for two dimensional inversions. One can also verify two-dimensional results with the help of one dimensional result and vice versa. In principle, 2-D model based interpretation is more appropriate to infer lateral resistivity variation caused by multiple episode of lava flows in this area. Figure 8b shows the synthetic pseudo-section containing 5% correlated red noise.



**Fig. 9.** Resistivity structure of sounding points s1–s10 by (a, b) singular value decomposition (SVD)-based conventional method (c, d) Hybrid Monte Carlo (HMC)-based Bayesian neural network (BNN) approaches.

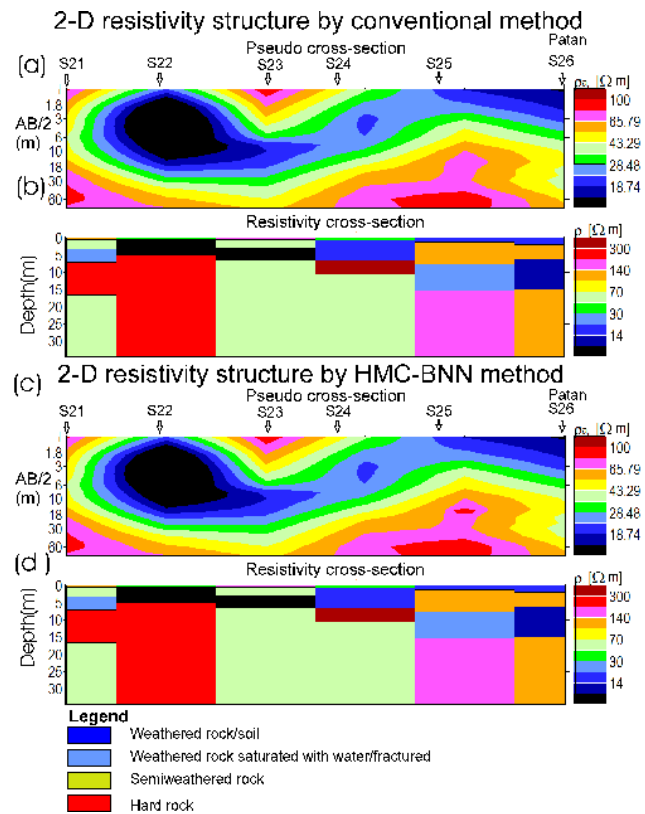
The Fig. 8c shows the synthetic pseudo-section obtained by the inversion of Bayesian Neural Network method. It is worth mentioning here that the present Bayesian Neural Network based resistivity model is well stable in presence of correlated red noise (Fig. 8a–c). One can see from Fig. 9 that there is a prominent resistivity low ( $\sim 2 - 10 \text{ Ohm}\cdot\text{m}$ ) below the station s3, Chiplun (Latitude  $17.51^\circ \text{ N}$ , Longitude  $73.52^\circ \text{ E}$ ). This conductive patch is fractured or weathered basaltic rock saturated with water which is extended in the direction of NE-SW of the profile (profile extends roughly W-E direction) and the profile cuts the fractured zone near Chiplun (Fig. 9). Note that the fractured rocks include jointed rocks, in which chemical weathering is almost absent. The semi weathered and weathered rocks are partly jointed or fractured in this area. In weathered rocks of the area both inter-granular and fracture porosities exist. The presence of feldspar and clay in the fractured rock reduces the permeability of the rock to some extent (Verma et al., 1980). Figure 10 shows the comparative status of resistivity distribution of profile starting from s11 to s20. Results are stable except few (e.g., below SP 19) (Fig. 10). There is a conductive path below the s15 (Pophli, Latitude  $17.43^\circ \text{ N}$ , Longitude  $73.68^\circ \text{ E}$ ) with relatively low resistivity value of the order of  $2-15 \text{ Ohm}\cdot\text{m}$  that could be due to the fracture zone/fault. The origin of the fractured zones/fault seems to be deep seated. The fractured/fault zone is parallel to the di-



**Fig. 10.** Resistivity structure of sounding points s11–s20 by (a, b) singular value decomposition (SVD)-based conventional method (c, d) Hybrid Monte Carlo (HMC)-based Bayesian neural network (BNN) approaches.

rection of flow of the Koyna river. Perhaps, this is due to an abrupt fall in the traps base of roughly  $400 \text{ m}$  at Pophli (Athavale and Mohan, 1976). Below the sounding point s18, it is evident that there is a conductive feature surrounded by high resistive rock which may form isolated ground-water reservoirs. The total porosity and permeability of these isolated reservoirs vary widely resulting in widely variable yield (Verma et al., 1980). In fresh rocks, joints and fissures tend to close at a depth of maximum  $70 \text{ m}$  and there will be practically no circulation of ground water below this depth. The area has relatively high resistivity distribution of the order of  $300 \text{ Ohm}\cdot\text{m}$  (Fig. 10). One can speculate from the Fig. 11 that there is a conductive layer of thickness  $\sim 2 - 5 \text{ m}$  extending from sounding points s22–s26. The conductive layer is exposed to the surface near sounding point s26 (near Patan). The very low resistive path in this station could be due to the very well known Patan Fault lying in the direction of  $\text{N}45^\circ \text{ E}-\text{S}45^\circ \text{ W}$  (see Figs. 1 and 11).

In general, the HMC-based BNN results have been found to be stable and match well with geology of the area. There are some mismatches too (below s19 and s25) (Figs. 10 and 11). Eventually, the uncertainty results estimated as  $0.97$ ,



**Fig. 11.** Resistivity structure of sounding points s21–s26 by (a, b) singular value decomposition (SVD)-based conventional method. (c, d) Hybrid Monte Carlo (HMC)-based Bayesian neural network (BNN) approaches.

is highest below s25. This mismatch could possibly be due to either the poor sampling of the model space in HMC-based BNN approach or the presence of high level of correlated red noise signal in the data. In addition to the finely mapping and demarcating of weathered rocks, lineaments, fractures etc., the present new method is also able to define the basaltic flow layer precisely which is of considerable significance in terms of societal issues like ground water exploration. Further this result is useful to gain a better understanding of the seismicity induced-parallel fault systems.

## 7 Conclusions

1. A novel Hybrid Monte Carlo-based Bayesian neural network inversion scheme is developed for the interpretation of DC resistivity data. This method, on contrary to the other traditional inversion scheme, does not require a prior correct initial model for successful optimization.
2. The Bayesian Neural Network method is a sampling-based global inversion scheme which takes care of non-linearity and non-uniqueness in a natural way. The

method also provides an opportunity to estimate uncertainty measures, which is essential for solving geophysical inverse problem.

3. In many geophysical application (e.g., seismic, electrical tomography), while repeated inversion is required from the same prior information, the new approach developed here could be more prospective and robust for the above purpose.
4. The Hybrid Monte Carlo-based Bayesian Neural Network can be applied as an alternative cost effective means for several other complex geophysical data also.
5. In order to check robustness of the Hybrid Monte Carlo-based Bayesian Neural Network, the method was applied to synthetic data set and then to a large number of actual DC resistive sounding points from the critically dynamic active zone of Koyna region for mapping the sub-surface resistivity structure. Besides providing some new and detailed information, the present new findings also corroborate well with existing results.

*Acknowledgements.* We are thankful to the Directors, Indian Institute of Geomagnetism (DST), New Panvel and National Geophysical Research Institute (CSIR), Hyderabad for their kind permission to publish the work. We are also grateful to K. Mallik, Emeritus Scientist, National Geophysical Research Institute for useful discussions.

Edited by: I. Zaliapin

Reviewed by: two anonymous referees

## References

- Aristodemou, E., Pain, C., Oliveira, C., and Goddard, T.: Inversion of nuclear well-logging data using neural networks, *Geophys. Prospect.*, 53, 103–120, 2005.
- Arnason, K. and Hersir, G. P.: One dimensional inversion of Schlumberger resistivity soundings. Computer Program, Description and User's Guide: The United Nations University, Geothermal Training, Report 8, 59 pp., 1988.
- Athavale, R. N. and Mohan, I.: A technical Report on Integrated Geophysical Studies in the Koyna Hydroelectric Project Area of Maharashtra State, NGRI, India and CEG, Hyderabad, 1976.
- Bescoby, D. J., Cawley, G. C., and Chroston, P. N.: Enhanced interpretation of magnetic survey data from archaeological sites using artificial neural networks, *Geophysics*, 71(5), H45–H53, 2006.
- Bishop, C. M.: *Neural networks for pattern recognition*, Oxford University Press, USA, 1995.
- Coppola Jr., E. A., Rana, A. J., Poulton, M. M., Szidarovszky, F., and Uhl, V. W.: A neural network model for predicting aquifer water level elevations, *Ground Water*, 43(2), 231–241, 2005.
- Devilee, R. J. R., Curtis, A., and Roy-Chowdhury, K.: An efficient probabilistic neural network approach to solving inverse problems: Inverting surface wave velocities for Eurasian

- crustal thickness. *J. Geophys. Res.*, 104(B12), 28841–28857, doi:10.1029/1999JB900273, 1999.
- Dey, A. and Morrison, H. F.: Resistivity modeling for arbitrarily shaped two-dimensional structures, *Geophys. Prospect.*, 27(1), 1020–1036, doi:10.1111/j.1365-2478.1979.tb00961.x, 1979.
- Duane, S., Kennedy, A. D., Pendleton, B., and Roweth, D.: Hybrid Monte Carlo, *Phys. Lett. B*, 195, 216–222, 1987.
- Duncan, R. A. and Pyle, D. G.: Rapid eruption of Deccan flood basalts at the cretaceous-tertiary boundary, *Nature*, 333, 841–843, doi:10.1038/333841a0, 1988.
- Dystart, P. S. and Pulli, J. J.: Regional seismic event classification at the NORESS array: Seismological measurements and use of trained neural network, *B. Seismol. Soc. Am.*, 80, 1910–1933, 1990.
- Ekinci, Y. L. and Demirci, A.: A damped least-squares inversion program for the interpretation of Schlumberger sounding curves, *Journal of Applied Sciences*, 8(22), 4070–4078, doi:10.3923/jas.2008.4070.4078, 2008.
- Feng, X.-T., Seto, M., and Katsuyama, K.: Neural dynamic modeling on earthquake magnitude series, *Geophys. J. Int.*, 128, 547–556, 1997.
- Ghosh, D.: The application of linear filter theory on the direct interpretation of geoelectrical resistivity sounding measurements, *Geophys. Prospect.*, 19, 192–217, 1971.
- Gupta, H. K. and Rastogi, B. K.: Will another damaging earthquake occur in Koyna?, *Nature*, 248, 215–216, doi:10.1038/248215a0, 1974.
- Gupta, H. K., Narain, H., Rastogi, B. K., and Mohan, I.: A Study of the Koyna earthquake of December 10, 1967, *B. Seismol. Soc. Am.*, 59, 1149–1162, 1969.
- Gupta, H. K., Rastogi, B. K., and Narain, H.: Common features of the reservoir associated seismic activities, *B. Seismol. Soc. Am.*, 62, 481–492, 1972.
- Gupta, H. K., Mandal, P., and Rastogi, B. K.: How long will triggered earthquakes at Koyna, India continue?, *Current Science*, 82(2), 202–210, 2002.
- Hastings, W. K.: Monte Carlo sampling methods using Markov Chain and their applications, *Biometrika*, 57(1), 97–109, doi:10.1093/biomet/57.1.97, 1970.
- Horne, S. and MacBeth, C.: Inversion of seismic anisotropy using genetic algorithms, *Geophys. Prospect.*, 42, 953–974, 1994.
- Khan, M. S. and Coulibaly, P.: Bayesian neural network for rainfall-runoff modeling, *Water Resour. Res.*, 42, W07409, doi:10.1029/2005WR003971, 2006.
- Kirkpatrick, S., Gelatt Jr., C. D., and Vecchi, M. P.: Optimization by simulated annealing, *Science*, 220, 671–680, doi:10.1126/science.220.4598.671, 1983.
- Koefoed, O.: A fast method for determining the layer distribution from the raised kernel function in geoelectrical soundings, *Geophys. Prospect.*, 18, 564–570, 1970.
- Lampinen, J. and Vehtari, A.: Bayesian approach for neural networks-review and case studies, *Neural Networks*, 14, 7–24, 2001.
- Levenberg, K.: A method for the solution of certain non-linear problems in least-squares, *Q. J. Appl. Math.*, 2, 164–168, 1944.
- Macias, C. C., Sen, M. K., and Stoffa, P. L.: Artificial neural networks for parameter estimation in geophysics, *Geophys. Prospect.*, 48, 21–47, 2000.
- MacKay, D. J. C.: A practical Bayesian framework for back-propagation networks, *Neural Comput.*, 4, 448–472, 1992.
- Maiti, S. and Tiwari, R. K.: A hybrid Monte Carlo method based artificial neural networks approach for rock boundaries identification: A case study from the KTB bore hole, *Pure Appl. Geophys.*, 166, 2059–2090, doi:10.1007/s00024-009-0533-y, 2009.
- Maiti, S. and Tiwari, R. K.: Automatic discriminations among geophysical signals via the Bayesian neural networks approach, *Geophysics*, 75(1), E67–E78, doi:10.1190/1.3298501, 2010a.
- Maiti, S. and Tiwari, R. K.: Neural network modeling and an uncertainty analysis in Bayesian framework: A Case study from the KTB borehole site, *J. Geophys. Res.*, 115, B10208, doi:10.1029/2010JB000864, 2010b.
- Maiti, S., Tiwari, R. K., and Kumpel, H. J.: Neural network modeling and classification of lithofacies using well log data: A case study from KTB borehole site, *Geophys. J. Int.*, 169, 733–746, doi:10.1111/j.1365-246X.2007.03342.x, 2007.
- Marquardt, D. W.: An algorithm for least-squares estimation of non linear parameters, *SIAM J. Appl. Math.*, 11, 431–441, 1963.
- Meju, M. A.: Geophysical data analysis: Understanding Inverse Problem Theory and Practice, in: *Society of Exploration Geophysics Course Notes Series*, 1st edn., No. 6, SEG publishers, Tulsa, Oklahoma, 296 pp, ISBN 1-56080-027-5, 1994.
- Menke, W.: *Geophysical data Analysis: Discrete Inverse Theory*, Academic Press, Inc., New York, 1984.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equations of state calculations by fast computing machines, *J. Chem. Phys.*, 21, 1087–1091, 1953.
- Morgan, W. J.: Deep Mantle convection plumes and plate motions, *Bulletin of the American Association of Petroleum Geologists*, 56, 203–213, 1972.
- Murat, M. E. and Rudman, A. J.: Automated first arrival picking: A neural network approach, *Geophys. Prospect.*, 40, 587–604, 1992.
- Nabney, I. T.: *Netlab Algorithms for pattern recognition*, Springer, New York, 2004.
- Neal, R. M.: *Bayesian learning for neural networks*, Springer-Verlag, New York Inc., 1996.
- Poulton, M. (Ed.): *Computational neural networks for geophysical data processing*, Pergamon, 2001.
- Qady, G. E. and Ushijima, K.: Inversion of DC resistivity data using artificial neural networks, *Geophys. Prospect.*, 49, 417–430, 2001.
- Raiche, A.: A pattern recognition approach to geophysical inversion using neural nets, *Geophys. J. Int.*, 105(3), 629–648, doi:10.1111/j.1365-246X.1991.tb00801.x, 1991.
- Roy, I. G.: An efficient non-linear least-squares 1D inversion scheme for resistivity and IP sounding data, *Geophys. Prospect.*, 47(4), 527–550, doi:10.1046/j.1365-2478.1999.00142.x, 1999.
- Rubinstein, R. Y.: *Simulation and the Monte Carlo method*, John Wiley and Sons, New York, 278 pp., 1981.
- Singh, U. K., Tiwari, R. K., and Singh, S. B.: One dimensional inversion of geoelectrical resistivity sounding data using neural network – A case study, *Comput. Geosci.*, 31, 99–108, 2005.
- Singh, U. K., Tiwari, R. K., and Singh, S. B.: Prediction of electrical resistivity structure using Artificial neural networks, *Journal of Geological Society of India*, 67, 234–242, 2006.
- Singh, U. K., Tiwari, R. K., and Singh, S. B.: Inversion of 2-D DC resistivity data using rapid optimization and minimal

- complexity neural network, *Nonlin. Processes Geophys.*, 17, 65–76, doi:10.5194/npg-17-65-2010, 2010.
- Spichak, V. and Popova, I.: Artificial neural network inversion of Magnetotelluric data in terms of three-dimensional earth macroparameters, *Geophys. J. Int.*, 142, 15–26, doi:10.1046/j.1365-246x.2000.00065.x, 2000.
- Sri Lakshmi, S. and Tiwari, R. K.: Model dissection from earthquake time series: comparative analysis using modern non-linear forecasting and artificial neural network approaches, *Comput. Geosci.*, 35, 191–204, 2009.
- Talwani, P.: Seismotectonics of the Koyna-Warna area, India, *Pure Appl. Geophys.*, 150(3–4), 511–550, doi:10.1007/s000240050091, 1997.
- Tarantola, A.: *Inverse Problem Theory*, Elsevier, New York, 1987.
- Uchida, T.: Two-dimensional resistivity inversion for Schlumberger sounding, *Geophysical Exploration of Japan*, 44, 1–17, 1991.
- Van der Bann, M. and Jutten, C.: Neural networks in geophysical applications, *Geophysics*, 65, 1032–1047, 2000.
- Verma, R. K., Rao, M. K., and Rao, C. V.: Resistivity investigations for ground water in metamorphic areas near Dhanbad, India, *Ground Water*, 18(1), 46–55, doi:10.1111/j.1745-6584.1980.tb03370.x, 1980.
- Zohdy, A. R.: A new method for automatic interpretation of Schlumberger and Wenner sounding curves, *Geophysics*, 54, 245–253, doi:10.1190/1.1442648, 1989.