# Physics-Guided Deep-Learning Inversion Method for the Interpretation of Noisy Logging-While-Drilling Resistivity Measurements

Kyubo Noh[1], David Pardo[2], and Carlos Torres-Verdín[3, *]

* Corresponding Author

1 Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX 78712 USA; currently Department of Earth Sciences, University of Toronto, Toronto, ON, M5S 3B1, Canada (e-mail: kyubonoh@gmail.com)

2 University of the Basque Country, 48940 Leioa, Spain; Basque Center for Applied Mathematics, 48009 Bilbao Spain; Ikerbasque, 48009 Bilbao, Spain (e-mail: dzubiaur@gmail.com)

3 Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: cverdin@austin.texas.edu)

**Abstract**

Deep Learning (DL) inversion is a promising method for real-time interpretation of logging-while-drilling (LWD) resistivity measurements for well-navigation applications. In this context, measurement noise may significantly affect inversion results. Existing publications examining the effects of measurement noise on DL inversion results are scarce. We develop a method to generate training data sets and construct DL architectures that enhance the robustness of DL inversion methods in the presence of noisy LWD resistivity measurements. We use two synthetic resistivity models to test the three approaches that explicitly consider the presence of noise: (1) adding noise to the measurements in the training set, (2) augmenting the training set by replicating it and adding varying noise realizations, and (3) adding a noise layer in the DL architecture. Numerical results confirm that each of the three approaches enhances the noise-robustness of the trained DL inversion modules, yielding better inversion results—in both predicted earth model and measurements—compared to the basic DL inversion and also to traditional gradient-based inversion results. A combination of the second and third approaches delivers the best results.

**Keywords:** Electrical properties, Machine learning, Downhole methods, Inverse theory, Neural networks

## 1. Introduction

As the demand for drilling complex geological structures and production-efficient well placement increases, well navigation or well geosteering, which dynamically adjusts the well trajectory, has become an essential technique. Among the various geophysical measurements acquired during drilling, logging-while-drilling (LWD) resistivity measurements are pivotal for well geosteering because resistivity contains comprehensive information about the spatial distribution of surrounding rocks and their saturating fluids (Nie *et al.* 2010, Lee *et al.* 2012). Thus, real-time processing and interpretation of LWD resistivity measurements have become important for active well geosteering decisions.

The interpretation of geophysical measurements is most often performed via inversion techniques. Conventional inversion employs deterministic or stochastic methods. Historically, the development and application of inversion techniques have progressed by invoking assumptions in 1.5-dimensional (1.5D—assuming a three-dimensional source and a one-dimensional resistivity structure) (Zhang *et al.* 2004, Bakr *et al.* 2017, Wang *et al.* 2017, Deleersnyder *et al.* 2021, Wang *et al.* 2021), 2.5-dimensional (2.5D) (Tabarovsky & Rabinovich 1998, Thiel *et al.* 2018), and more recently to three-dimensional (Abubakar *et al.* 2006, Puzyrev *et al.* 2018, Marchant *et al.* 2019) spaces. With the recent widespread utilization of machine learning (ML) algorithms, inversion using ML—and in particular, deep learning (DL)—has become a valuable alternative to traditional geophysical inversion methods (Araya-Polo *et al.* 2018, Colombo *et al.* 2021, Wei *et al.* 2022). In electromagnetic logging applications, Jin et al. (2019) invert 1.5D LWD resistivity measurements with a deep neural network (DNN) and a rigorous forward model by invoking a loss function that combines both model- and data-misfit; Shahriari et al. (2021) construct a DNN approximation of both forward and inverse problems; Hu et al. (2020) combine a gradient-based and a ML inversion method to enhance the generalization of training conditions; Noh et al. (2021) and Noh et al. (2022) considers multiple DL inversion architectures designed for varying geological situations for a 2.5D inversion of possibly fractured rocks.

Under real drilling conditions, LWD resistivity measurements are inevitably contaminated by noise. This adverse condition affects the stability and reliability of the inversion. Thus, noise-robustness becomes a key feature of an inversion algorithm. Among the traditional inversion methods, stochastic approaches are naturally designed to evaluate the uncertainty of inversion results in the presence of noisy and inadequate measurements (Wang et al. 2019). In gradient-based methods, alternative approaches have been used, such as introducing linearized uncertainties assuming achievable convergence and noisy measurements (Pardo & Torres-Verdín 2015) and Huber-norm-based inversion to minimize the effect of Gaussian and non-Gaussian noise (Hu & Fan 2018).

However, documented studies examining the robust implementation of DL inversion for noisy LWD resistivity measurements are scarce. Adding noise in the training data set is the most common approach used to enhance the generalization capacity of a DL method (Goodfellow et al. 2016); it has been shown that adding noise during training functions as a traditional Tikhonov regularization (Bishop, 1995). Thus, adding noise in the training data has been widely used to enhance the robustness of geophysical inversion (Côrte *et al.* 2020, Li *et al.* 2020, Chen *et al.* 2021). A few pioneering works on DL inversion of LWD resistivity also show that training with noisy measurements enhances the robustness of the trained DL models (e.g., Jin et al. 2019 and Hu et al. 2020). However, the latter publications only consider the same known noise distributions for both training and test data sets. However, our study aims to investigate the cases in which the noise model for the training data set is different from that used for the test data set. To properly approach the problem of noisy measurements, the following arguments are considered: 1) the adverse effects of noise on inversion results, 2) the effect of adding noise during training in problems where the exact noise level may be unknown, and 3) the development of more effective remedies to invert noisy LWD resistivity measurements using DL.

As part of such efforts, we implement and compare several strategies and their combinations to mitigate the effect of measurement noise on inversion results, namely: (1) to add synthetic measurement noise in the training data set — the standard approach; (2) to incorporate synthetic measurement noise using augmented

training data; and (3) to include neural-network layers that add noise in the DL architecture rather than in the training data. They act as a regularizer of the DL training process. Note that the goal of performing noise-robust inversion is to obtain resistivity models similar to those obtained with noise-free measurements rather than those that maximally fit the noise-contaminated measurements. To numerically assess the performance of these approaches, we employ a 1.5D DL inversion method to compare the results of the aforementioned methods with those obtained from conventional gradient-based inversion. While a 1.5D approach does not provide the final validation of the method, it delivers stable and interpretable results, and is computationally efficient. The comparison includes conditions where different levels or probabilistic distributions of measurement noise are employed for the training and test data sets, and the application of a noise-filtering approach.

## 2. Deep-Learning Inversion for Reconstructing Resistivity Models

### 2.1 Physics-Guided DL inversion using an approximated forward solution

After recent active interest in utilizing DL techniques in geophysical inversion, there has been a series of developments intended to integrate the underlying physics (i.e., physics-driven DL inversion) in training DL inversion modules to increase physical consistency and reliability of the inverse solution. One can rigorously use actual forward solutions or approximated forward solutions. In this study, we take the latter approach, following Shahriari et al., (2020) and Noh et al. (2021 and 2022). Thus, we need to solve the following minimization problem:

$$\left(F_{\phi^*}, I_{\theta^*}\right) := \arg \min_{\phi\in\Phi,\theta\in\Theta} \left\{ \left\| \left(F_\phi \circ I_\theta\right)(\mathbf{D},\mathbf{T}) - \mathbf{D} \right\|_2^2 + \left\| F_\phi(\mathbf{M},\mathbf{T}) - \mathbf{D} \right\|_2^2 \right\}, \quad (1)$$

where **M**, **D**, **T** denote subsurface properties, measurements, and well trajectory, respectively, and $F_\varphi$ and $I_\theta$ are the DL approximations of the forward and inverse functions, respectively. Symbols $\varphi$ and $\theta$ represent sets of trainable weights in the DL approximations. Fig. 1 describes the workflow for training the forward and inverse approximators using the objective function in Eq. (1). The first additive term in Eq. (1) ensures that the composition of the inverse with the forward operator approximates the identity, and it can be interpreted as an encoder-decoder DL architecture, while second term guarantees the accuracy of the forward (neural network) operator. The above loss function produces the two operators $F_{\phi^*}$ and $I_{\theta^*}$. Note that a minimization of ($\|I_\theta(\mathbf{D},\mathbf{T}) - \mathbf{M}\|_2^2$) would lead to an average of inverse solutions that is derived by only minimizing model-misfits–which is not an inverse solution itself because of the non-uniqueness of the geophysical inverse problem–, while the approach employed in this work (Eq. (1)) pursues one of the solutions (branches) of the inverse problem that is derived by minimizing errors in both data- and model-domains. Thus, the above objective function better deals with the possible non-uniqueness of the inverse solution compared to the non-physics-driven DL inversion approach; see Shahriari *et al.* (2020) for details.

### 2.2. Deep Learning Architecture

For solving Eq. (1), we use a deep residual neural network (ResNet) (He et al. 2016a; He et al. 2016b) with convolutional layers (Lecun and Bengio 1998). Our deep ResNet consists of several residual blocks, which include convolutional, batch normalization, and activation layers. The identity mapping between each residual block enables a successful training of the DL architecture by preventing vanishing gradients that can occur when gradient-based optimization and backpropagation are used to train deep architectures (He *et al.* 2016b).

In order to compare their robustness against the measurement noise in the following numerical examples, we build two types of DL architectures: a noise-free ResNet and a ResNet with a hidden noisy layer. We compare the DL inversion results obtained with both architectures. The vanilla form follows the general type of ResNet constituted by residual blocks with batch normalization, ReLU activation (Nair and Hinton 2010), 1D convolution, and Dropout layers (Srivastava *et al.* 2014). For the second architecture, we can apply noise to the hidden layers to act as a regularization and enhance robustness against noise present in the measurements (Goodfellow *et al.* 2016) in a general machine-learning sense. By adding architecture noise, we prevent overfitting and enhance the neural network generalization capability (against "unseen" noisy measurements), which is an explicit form of regularization. This procedure can be implemented in various ways. Neelakantan *et al.* (2015) show that adding noise to the gradient at every training step avoids data overfitting and increases the training performance of deep architectures. In Gulcehre *et al.* (2016), authors propose to use a noisy activation function to enhance randomness in the hidden state. Poole *et al.* (2014) propose injecting noise in the hidden states of the autoencoders—referred to as noisy autoencoders. Following the approach introduced by Poole *et al.* (2014), we add noise to our DL architectures for the inversion of noisy borehole resistivity measurements. Specifically, we add a Gaussian noise component after every convolutional layer implemented using the *GaussianNoise* layer in Keras (Chollet 2015). We select two values of the standard deviation of the Gaussian noise to control the intensity of the hidden noisy layer. Table 1 lists the names of the DL architectures and used standard deviation for the Gaussian noise. Consequently, we have three DL architectures: one noise-free ResNet and two noisy ResNets. The Dropout

layers also act as strong regularizations by functioning as noise in hidden states (Goodfellow *et al.* 2016). Using the selected ResNet architectures, we define both forward and inverse approximators ($F_\phi$ and $I_\theta$ in equation (1)) by setting adequate input and output shapes for each approximator. Detailed shapes of the used DL architectures are shown in Table A1 in Appendix A.

### 2.3. Data Preparation and Training

In what follows, we describe the subsurface and well trajectory parameterizations, measurement system, and measurement noise employed to generate the DL training and validation data sets.

*Subsurface Parameterization*

We consider a piecewise-1D inversion approach, which is commonly used in geosteering applications (Pardo *et al.* 2021). Additionally, the Earth model is assumed to be a three-layer medium, as shown in Fig. 2. The host layer is anisotropic with horizontal and vertical resistivities denoted as $\rho_h$ and $\rho_v$, respectively; the upper and lower layers are isotropic with resistivities $\rho_u$ and $\rho_l$, respectively. We also consider the vertical distance to the upper and lower layers—denoted as $d_u$ and $d_l$, respectively—and the layer dip angle, denoted as $\beta$. According to the physics of our problem, we restrict the resistivity values to the range [0.1, 1000] ohm-m, the anisotropy factor $\rho_v / \rho_h$ to the range [1, 10], the distance values $d_u$ and $d_l$ to the range [0.01, 5] m, and the dip angle $\beta$ to the range [-30, 30] degrees. Furthermore, we sample resistivity and distance values in a log scale, and the dip angle in a linear scale.

*Well Trajectory Parameterization*

For our geosteering application, we consider high-angle well trajectories with a dip angle within the range [80, 100] degrees. We reconstruct each three-layer model from a 2 m-long trajectory segment (with a 0.1-meter sampling interval) given by 21 measurement points. Additionally, a maximum of 0.3-degree deviation of the drilling trajectory is assumed in a 2 meters section; hence, the initial dip angle is within the range [80.3, 99.7] while well deviation from one measurement point to the next is within the range [-0.03, 0.03] degrees.

*Measurement System and Measurement Noise*

We consider two logging instruments: a short-sensing LWD instrument and a deep-sensing instrument—see Fig. 3. The short LWD instrument includes five pairs of transmitters equally spaced from the tool center and operates at 2 MHz; on the other hand, the deep instrument only includes one transmitter-receiver pair separated by 12 meters and operating at 10 kHz. From the five transmitter-receiver pairs of the short LWD instrument and the transmitter-receiver pair of the deep instrument, we calculate attenuation and phase components of one co-axial, two co-planar, and one cross-coupling components (denoted as Geosignal component). Eq. (2) leads to the definition of attenuation and phase difference for the co-axial component:

$$\ln \frac{H_{zz}^{1}}{H_{zz}^{2}} = \ln \frac{|H_{zz}^{1}|}{|H_{zz}^{2}|} + i\left(phase(H_{zz}^{1}) - phase(H_{zz}^{2})\right), \tag{2}$$

In the above equation, the first additive term in the right-hand side is the attenuation, while the second one is the phase difference. The superscript denotes the receiver´ index, while the first and second subscripts denote the orientation of the transmitter and receiver, respectively. For the short LWD instrument having five pairs of two transmitters, the measured real and imaginary components in Eq. 2 were averaged over two transmitters (equally spaced from the tool center) forming attenuation and phase measurements from real and imaginary components, respectively. For the deep instrument, the same equation with $H_{zz}^{2} = 1$ was used. Shahriari *et al.* (2021) describes in detail the corresponding postprocessing implemented to the acquired measurements to replicate the commercial implementations of similar systems. In summary, for each model, the input data consist of 6×4×2 types of measurements (5 pairs of transmitters × 4 directional components × 2 components of attenuation and phase from the short-sensing and 1×4×2 types from the deep-sensing instrument), each one having a length equal to 21.

For the data sets selected in the numerical experiments, we introduce measurement noise in both training and test data sets. We implement two different types of probabilistic distributions to generate noise of a Gaussian distribution and a log-normal distribution. Two different types of distributions were introduced to simulate the case when the inexact measurement noise model is used for the training of DL inversion

modules. Notice that the log-normal distribution only takes positive variables into account, unlike a Gaussian distribution. This difference will likely affect our experiments. Thus, we also set a uniformly random possibility on the sign of the result of the log-normal distribution, so that the noise can have either positive or negative sign in directional measurements. The resulting distribution can be referred to as a domain-extended log-normal distribution, but for simplicity, we will refer to it as log-normal distribution. For both types of noise distributions, we consider two intensities of relatively weak and strong values—see Table 2. Again, each attenuation and phase component corresponds to the first and second terms in the right-hand-side of Eq. (2), respectively.

*Training Data Sets and Training Process*

We generate five training data sets with varying noise types for our comparative study about noise effects on inversion results. Table 3 lists the names and sizes of the five data sets. The first and basic data set contains 100K samples with varying model and trajectory parameters randomly sampled from the aforementioned ranges and their corresponding measurements. We use a semi-analytic numerical method to perform the 1.5D simulations (see Shahriari *et al.* (2018); Pardo *et al.* (2021) and references therein), where the computation takes approximately 20 CPU hours using a computer equipped with Intel© Core i7-8700 CPU. Note that the semi-analytic 1.5D simulator is used as our forward function ($F$) to generate the datasets. A DNN approximation of this function ($F$) is used to train the inverse function in the framework of physics-driven DL inversion by simultaneously training the function approximator--$F_\phi$ in equation (1). Using this first and basic data set, we generate a second and a third data set by adding weak and strong Gaussian noise, respectively. The fourth data set is generated by adding strong log-normal noise to the basic data set. To generate the fifth data set, we first triplicate the basic data set, then we contaminate it with varying weak noise. With this fifth data set, the inversion module is trained to infer identical resistivity models from each set of three measurement samples contaminated with different noise values. By imposing that the trained DL inversion module led to the identical resistivity model from the triplicated and contaminated measurements, we expect the trained DL inversion module to become more robust to the

presence of measurement noise. This fifth data set will be referred to as an augmented weakly noisy data set.

We train each DL architecture independently for each data set. In so doing, we select 80% of the samples for training, while the remaining 20% are equally divided between validation and test data sets. The training processes for each of the first four data sets take approximately 8 hours using a machine equipped with Intel Xeon E5-2620 v4 CPU and four Nvidia 1080-TI GPUs, while that for the augmented fifth data set takes 20 hours with a given set of hyperparameters.

## 3 Numerical Results

To assess the effect of noise contamination on DL inversion results and propose possible remedies, we utilize two synthetic well geosteering problems: a simple three-layer isotropic model and a complex three-dimensional anisotropic model. For the assessment, we compare the inversion results by using varying combinations of noise in the (a) measurements (observed data), (b) training data, and (c) the DL architecture.

### 3.1 Simplified Three-Layer Isotropic Model

We utilize the simple three-layer isotropic model shown in Fig. 4a for DL and conventional inversion and assume a simple three-layered isotropic structure to compare and demonstrate the performance of the inversion algorithms.

*Training with noise-free measurements*

Fig. 4b shows the DL inversion results obtained with noise-free measurements using a noise-free DL architecture trained with a noise-free data set—this situation corresponds to experiment 1 in Table 4. Upper and lower figures of inverted results throughout this paper (Figs. 4, 6, 7, 8, and 11) illustrate horizontal and vertical resistivities, respectively. We observe an adequate reconstruction of formation resistivity: the inversion correctly detects the conductive layer embedded within the resistive host formation; it also detects the bed boundaries nearby the well trajectory (e.g., at horizontal location x within the ranges [35, 40], [50, 55], [60, 65], and [75, 80] m). However, there are some incorrect estimations of bed boundaries located far away from the well trajectory. This occurs due to the limited depth of investigation of the assumed logging instruments, which is slightly above 1 m for short LWD measurements and 5 m for deep-sensing measurements (Fig. 3).

Figs. 4c and 4d show DL inversion results obtained when using the same architecture as in Fig. 4b, but with the weakly and strongly Gaussian-noisy test data, respectively: they correspond to experiments 2 and 3 in Table 4, respectively. We observe a rough approximation of the original formation, where the prediction of layer boundaries is inadequate. The formation model mismatch is greater with the strongly Gaussian-noisy

measurements (Fig. 4d) than with the weakly Gaussian-noisy measurements (Fig. 4c). As a reference to the DL inversion results obtained for the case of strongly Gaussian-noisy measurements in Fig. 4d, we display in Fig. 4e the gradient-based 1.5D inversion results obtained with the strongly noisy measurements, which corresponds to experiment 4 in Table 4. The gradient-based inversion is conducted using the algorithm shown in Pardo & Torres-Verdín (2015). We stop our inversion algorithm when the decrease of the relative data misfit is below 0.5% during the last 20 iterations. For the deterministic DL inversion, we considered a different parameterization: horizontal and vertical resistivities of three-layered model were set to be inversion parameters and distances to neighboring boundaries (corresponding to $d_{\mathrm{u}}$ and $d_{\mathrm{l}}$ in the DL inversion case (Fig. 2)) were fixed—adopting true values—unlike in the DL inversion exercise, because otherwise bed-boundaries would be significantly misplaced. Other types of parameterization for the inversion are possible. For example, over-parameterization subsurfaces combined with spatial smoothing regularization (de Groot–Hedlin, 2000), which was not considered in this work. Inversion results in Fig. 4e lack spatial continuity and exhibit significant artifacts, especially concerning vertical resistivities.

Figs. 5a and 5b compare the original noise-free measurements to those with strong noise, as well as the predictions from the inversion results obtained using DL and gradient-based approaches—they correspond to experiments 3 and 4 in Table 4, respectively. The actual forward simulator (Shahriari *et al.* 2018)—not the trained forward approximator ($F_{\phi^*}$) used to generate training data sets—is used for the calculation of the predictions. Among the existing 24 components, we display the co-axial component of the shortest transmitter-receiver pair included in the LWD instrument. Although both methods reconstruct to some extent the original measurements, gradient-based inversion results exhibit larger differences with respect to those obtained with the noise-free original measurements. To quantitatively evaluate the performance of data reconstruction for the inversion result, we display in Table 4 the root-mean-squared deviation (RMSD) between (1) the predicted measurements from the inverted model and the noise-contaminated measurements (*RMSD–Noisy*) and (2) the predicted measurements from the inverted model and the original noise-free measurements (*RMSD–Noise free*). The next equation shows the definitions of the RMSD:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}\left(d_i^{pred} - d_i^{obs}\right)^2}{\mathrm{N}}}, \tag{3}$$

where $d_i^{pred}$ is the *i*-th predicted datum, $d_i^{obs}$ is the *i*-th observed datum that is either noise-contaminated value for *RMSD–Noisy* or noise-free value for *RMSD–Noise free*.

All 24 components are used to compute the *RMSD–Noisy* and *RMSD–Noise free* values. Comparison of the results obtained from experiments 1 to 3 indicates that both *RMSD–Noisy* and *RMSD–Noise free* increase as measurement noise increases. We also observe lower *RMSD–Noise free* than *RMSD–Noisy* values for all experiments that indicates reconstruction capability of original noise-free measurements of DL inversion. Gradient-based inversion results also exhibit lower *RMSD–Noise free* than *RMSD–Noisy*, but the difference is smaller than that of the DL inversion results. In summary, the gradient-based inversion is less robust to measurement noise than the vanilla DL inversion.

*Training with noisy measurements – Known measurement noise distribution*

In this subsection, we assume that the measurement noise distribution is known, i.e., the same noise distribution is used for the training of DL architectures and for the simulation of measurement noise in the observed measurements. We independently train the DL architectures by using training data sets with measurement noise: “Weak Gaussian Noise” and “Strong Gaussian Noise” in Table 3 and then apply the DL architectures to the test data with strong measurement noise. These conditions correspond to experiments 5 and 6 in Table 4, respectively. Fig. 6a—trained with strong Gaussian noise—shows a more spatially continuous and accurate reconstruction of the distribution of electrical resistivity than Fig. 4d—trained with noise-free data. For example, inversion results become more reliable as the well trajectory penetrates the conductive layer embedded in a resistive host. Fig. 6b is obtained by applying a DL architecture trained with a weaker level of Gaussian measurement noise than the test data set. Thus, this case resembles real-world operations where the actual noise level is uncertain. Compared to the noise-free training data set—Fig. 4d, Fig. 6b shows a more spatially continuous resistivity distribution. Training with

a different noise level than the actual measurement noise level seems helpful in delineating the resistivity structure. Table 4 shows the *RMSD–Noisy* and *RMSD–Noise-free* values associated with the inversion results obtained for experiments 5 and 6. *RMSD–Noisy* and *RMSD–Noise-free* values of experiment 5 are lower than those of experiment 6, which indicate that a better resistivity reconstruction occurs when noise levels in the measurements and training data sets are close to each other. However, even when the noise levels are different, we also observe a better reconstruction of the resistivity distribution when including noise in the training data set than when employing noise-free training data—compare results of experiments 6 and 3.

*Training with noisy measurements – Unknown measurement noise distribution*

For each borehole logging instrument, both the estimated noise level and noise probabilistic distribution are unknown. Thus, it becomes important to compare the effect on inversion results of training over noisy measurements with different noise levels and noise distributions. Fig. 6c and experiment 7 in Table 4 summarize the inversion results obtained from measurements with a strong log-normal distribution noise when using an architecture trained with weak Gaussian noise in the measurements. Compared to the former result of experiment 6 shown in Fig. 6b, the result of experiment 7 shown in Fig. 6c overall displays similar but slightly more diverse inversion results in the horizontal direction and higher *RMSD–Noisy* and *RMSD–Noise free* values. We now observe a degraded inversion result compared to that obtained when using strong Gaussian noise. This behavior is expected because we use two log-normal distributions (with positive and negative values, as mentioned earlier) that have higher mean magnitudes compared to the zero-mean Gaussian distributions, even when the same maximum intensities are employed to generate the noise (see Table 1). However, the similarity in the inversion results shown in Figs. 6b and 6c indicates the generalization capacity of the trained DL inversion architecture even when the selected noise distribution in the measurements is different from the actual one. In other words, the trained architecture generalizes to measurements that contain noise properties different from the ones used to train the network. This behavior

can be attributed to the use of multiple measurements such as directional components or multiple coil spacings.

One simple but efficient approach to deal with noisy measurements is to mitigate the effect of noise by filtering it out from the measurements. We implement a simple moving-average approach to illustrate the latter principle. Fig. 7a compares measurements with/without strong Gaussian-noise contamination and after moving-average filtering. One can notice that the effect of noise contamination is significantly reduced after filtering. Fig. 7b shows the inversion result obtained from the filtered measurements using the DL inversion architecture trained with measurement noise. The reconstructed resistivity model shows a lower variation in the horizontal direction compared to the inversion result obtained without measurement filtering. However, compared to the DL inversion results obtained with the same architecture obtained without filtering in Fig. 6a but with assumed measurement noise during the training, results in Fig. 7b are now more inaccurate and unstable even at and near the central target area. A similar result can also be found in the data domain, as shown in the RMSD results of Table 4. Comparing *RMSD–Noisy* and *RMSD–Noise-free* values from experiments 8 and 3—obtained with and without smoothing filter, respectively—, we conclude that the data-filtering strategy can better fit the inversion measurements (lower *RMSD–Noisy* of experiments 8 and 3), but this does not necessarily improve the original noise-free signal (compare *RMSD–Noise free* results of experiments 8 and 3).

*Training with augmented measurement noise*

We now use a training data set with an augmented measurement noise, corresponding to the "Augmented Weak Gaussian Noise" procedure described in Table 3. We invert strongly Gaussian-noisy measurements, corresponding to experiment 9 in Table 4. Fig. 8a shows the inversion result. Improved inversion results are obtained compared to the result obtained without noise augmentation in Fig. 6b. In particular, we adequately estimate the bed boundaries at x within the range [45, 50] compared to the result shown in Fig. 6b. In addition, the upper and lower bed boundaries become more clearly delineated when the well trajectory is positioned within the conductive layer. Comparison of the *RMSD–Noisy* and *RMSD–Noisefree*

obtained with and without noise augmentation (experiments 6 and 9 in Table 4, respectively), indicates that using the noise-augmentation approach better reconstructs both noise-free and noisy measurements.

*Adding noise layers to the DL architecture*

To enhance the noise robustness of the DL inversion method, we train two DL architectures with a weakly Gaussian-noisy training data set by varying the standard deviation for Gaussian architecture noise: "DL—Strong Gaussian" and "DL—Weak Gaussian" in Table 1. Note that the noisy layers were only activated during the training phase to avoid stochastic features in the actual inversion process. Experiments 10 and 11 in Table 4 are conducted by inverting strongly noisy measurements; inversion results are shown in Fig. 8b and 8c, respectively. Fig. 8b (obtained with DL—Strong Gaussian) shows a spatially smooth resistivity distribution in the horizontal direction and a better reconstruction of the layered structure compared to the inversion result obtained without either measurement noise in the training data or the architecture noise shown in Fig. 4d. Close to the center of the model, inversion results adequately predict a conductive layer embedded within a resistive background. However, we observe inaccurate estimations in other regions in terms of bed boundary locations and horizontal resistivity near layer boundaries. The relatively high architecture noise causes the DL architecture to focus on the general trend rather than on small features. Fig. 8c shows the inversion result corresponding to the strongly Gaussian-noisy test data obtained with the "DL—Weak Gaussian". It exhibits better estimations of bed boundary locations (without losing the resolving power on the three-layered structure) compared to the inversion result obtained with the strongly noisy DL architecture in Fig. 8b. After applying weak architecture noise, inversion results clearly show both bed boundaries and a three-layer structure, without the strong artifact observed in Fig. 6b. The effect of architecture noise can also be observed in *RMSD–Noisy* and *RMSD–Noise free* values in Table 4. The result obtained with weak architecture noise exhibits lower *RMSD–Noisy* and *RMSD–Noise free* values than those corresponding to the strong architecture noise. Important differences between the two architectures indicate that adding architecture noise can only be effective when that the magnitude of the architecture noise is carefully tuned (Neelakantan *et al.* 2015).

*Combined augmented measurement noise and architecture noise*

We combine both augmented measurement noise in the training data set and architecture noise: "Augmented Weak Gaussian" and "DL—Weak Gaussian" in Tables 3 and 1, respectively. Fig. 8d displays the corresponding inversion results obtained with the strongly noisy test data, which corresponds to experiment 12 in Table 4. The inversion result predicts a resistivity distribution similar to that obtained without the augmenting approach in Fig. 8c, but with a smoother resistivity distribution in the horizontal direction. Results shown in Fig. 8d better reconstruct the three-layer structure at the center of the model and bed boundaries compared to the result in Fig. 4d obtained without any treatment for measurement noise. Fig. 8d shows a similar resistivity distribution to that obtained with noise-free measurements; the corresponding predicted measurements (shown in Fig. 5c) also indicate a good reconstruction of the original noise-free measurements. *RMSD–Noisy* and *RMSD–Noise free* values associated with Fig. 8d and shown in Table 4 indicate the lowest performance among the results obtained with strong measurement noise. Specifically, the *RMSD–Noise free* result shows the high capacity of the inversion to predicting denoised original measurements when employing both augmented measurement noise and architecture noise.

In order to quantitatively evaluate the performance of the DL inversion with the combined measures to obtain noise robustness, we calculate the pixel-wise model mismatch of the inversion results using the following equation:

$$model\ mismatch = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{m_i^{true}-m_i^{inv}}{m_i^{true}}\right)^2} \tag{4}$$

where $m_i^{true}$ and $m_i^{inv}$ are pixelwise converted $\log_{10}$ resistivities of the true and inverted models. The pixelwise conversion of the resistivity models was conducted with the size of 0.1 meters for the pixels using 5 meters of threshold in the vertical (depth) direction along the trajectory for both horizontal and vertical resistivity sections. The calculated results for the cases of the strong Gaussian measurement noise are shown in Table 5. The calculated result for the DL inversion result for the noise-free measurements (experiment

1) is also shown for comparison. We observe that the inversion of strong measurement noise degrades the model reconstruction in both DL and gradient-base inversion approaches (see experiments 3 and 4). As the proposed remedies of the augmented measurement noise during the training phase and architectural noise are additionally applied to the DL inversion (see experiments 9 and 12), the model mismatch becomes substantially reduced to a similar level of the model mismatch result of noise-free measurement (experiment 1). The combined results of Tables 4 and 5 show that the proposed remedy of utilizing the augmented measurement noise and architectural noise enhances the noise-robustness of the DL inversion and enhances both model and data reconstruction capabilities.

### 3.2 Three-Dimensional Anisotropic Model

We now apply our DL inversion approach to a complex 3D model. Fig. 9 shows a 3D view of the model along with a 3D well trajectory. We consider a model consisting of ten layers with varying dip angles and two vertical fault planes. Fig. 10 shows plane views of horizontal and vertical resistivity across the model's center in the *y*-direction (i.e., *y*=25 m in Fig. 9) with the 3D well trajectory projected in that plane. Fig. 11a shows the horizontal and vertical resistivity distribution of the corresponding three-layer model along the 3D well trajectory. We also implemented 1.5D simulations to generate the data to be inverted in order not to introduce multi-dimensional effects when analyzing noise effects. Additionally, we decomposed the 3D model into point-by-point 1D models by extracting resistivity models along the well trajectory for every 20 centimeters in order to simulate 1D resistivity measurements and contaminate them using strong Gaussian noise.

Figs. 11b, 11c, and 11d compare inversion results obtained with: (a) DL without any noise treatment—experiment C-1 in Table 6—, (b) DL inversion with augmented weak noise for training and weak architecture noise—experiment C-2 in Table 6—, and (c) gradient-based 1D inversion—experiment C-3 in Table 6. The DL inversion result obtained without noise treatment in Fig. 11b only shows a reliable reconstruction of the host resistivity values along the well trajectory. However, it fails to properly estimate the three-layer structure and the bed boundaries. Fig. 11c, obtained after applying augmented measurement

noise and optimal architecture noise, shows a spatially smooth resistivity distribution in the horizontal direction. Even though some inaccuracies appear in the prediction of bed boundaries, when the latter are close to the well trajectory (e.g., at $x$ values of approximately 120, 140, and 160 m), inversion results estimate the actual location of the bed boundaries. Despite the fact that the gradient-based inversion result shown in Fig. 11d was obtained with the exact bed boundaries, results lack spatial continuity in the horizontal direction, and the estimated host resistivities are degraded compared to those obtained with the DL inversion that is obtained with the proposed remedies of augmented measurement noise and optimal architecture noise (Fig. 11c).

Comparisons of the original noise-free, observed noisy, and predicted measurements obtained from the inversion results shown in Figs. 11b, 11c, and 11d are described in Figs. 12a, 12b, and 12c, respectively. We reproduce the original noise-free measurements of the DL inversion with augmented training data noise and architecture noise. Table 6 compares the *RMSD–Noisy* and *RMSD–Noise free* values of the inversion results shown in Fig. 11. As in the previous numerical model, the DL inversion result with augmented training noise and architecture noise exhibits lower *RMSD–Noisy* and *RMSD–Noise free* values than those associated with the gradient-based and basic DL inversion results. The improvement in *RMSD–Noise free* is remarkable in the DL inversion results when using a noisy training data set and architecture noise. After applying the augmented measurement training noise and a proper level of architecture noise, the DL inversion delivers reliable results.

## DISCUSSION

The DL inversion approach used in this study requires significantly different computational costs in their training and application stages. The low off-line cost for the actual inversion evaluation is often considered as one of the biggest advantages of DL methods over conventional inversion approaches. This is especially true in applications—such as geosteering—which require an instantaneous inversion. However, DL techniques often suffer from a lack of generalization. In the developed approach, the trained DL architectures are constrained to specific logging instruments, resistivity model parameterizations, and trajectory parameterizations. Accordingly, the assumed measurement noise in training can be considered one of the limiting factors of the performance of DL inversion approaches. The objective of the work was to help to acquire or enhance the noise-robustness of the DL inversion approach, especially when the exact measurement-noise models are unknown. Although this study considered application results for numerically simulated responses, the proposed methods can be readily generalized to multi-dimensional cases or even other conditions of geophysical inversion. In such extended studies or in the preceding stages, fair experiments using actual data sets will be helpful in showcasing the advantages of the proposed method under actual and practical application conditions.

## CONCLUSIONS

We investigated the effect of measurement noise in the DL inversion of borehole resistivity measurements using two synthetic models. We then proposed various remedies to improve the inversion results; they consist of training DL networks using a noise-contaminated—and possibly augmented—training data set and adding a noise layer within the architecture. Results show that both gradient-based and DL inversion trained without considering noise are vulnerable to noisy measurements. Adding noise to the training data set enhances the noise robustness of the DL inversion in both model and data spaces. The effectiveness of the noise contamination strategy in the measurements increases when selecting the same noise level and distribution as the one existing in the actual input measurements. Because the exact noise level for a specific acquisition environment is unknown, we propose to use augmented noisy training data sets with different noise levels to improve inversion performance. In addition, even when the exact noise distribution is unknown, it is still possible to construct a DL approximation by implementing different noise distributions. Adding a noise layer in the DL architecture enhances the generalization capacity and noise robustness of the DL inversion. We use a grid-searching strategy to derive the proper level of architecture noise. The combination of augmented measurement noise with a noise layer in the DL architecture provides the best inversion performance for the noisy test data in both the predicted earth model and predicted measurements. Inversion methods developed in this paper are dimension independent. Therefore, we expect the proposed remedies for the DL inversion of noisy measurements to remain valid for multi-dimensional inversion cases.

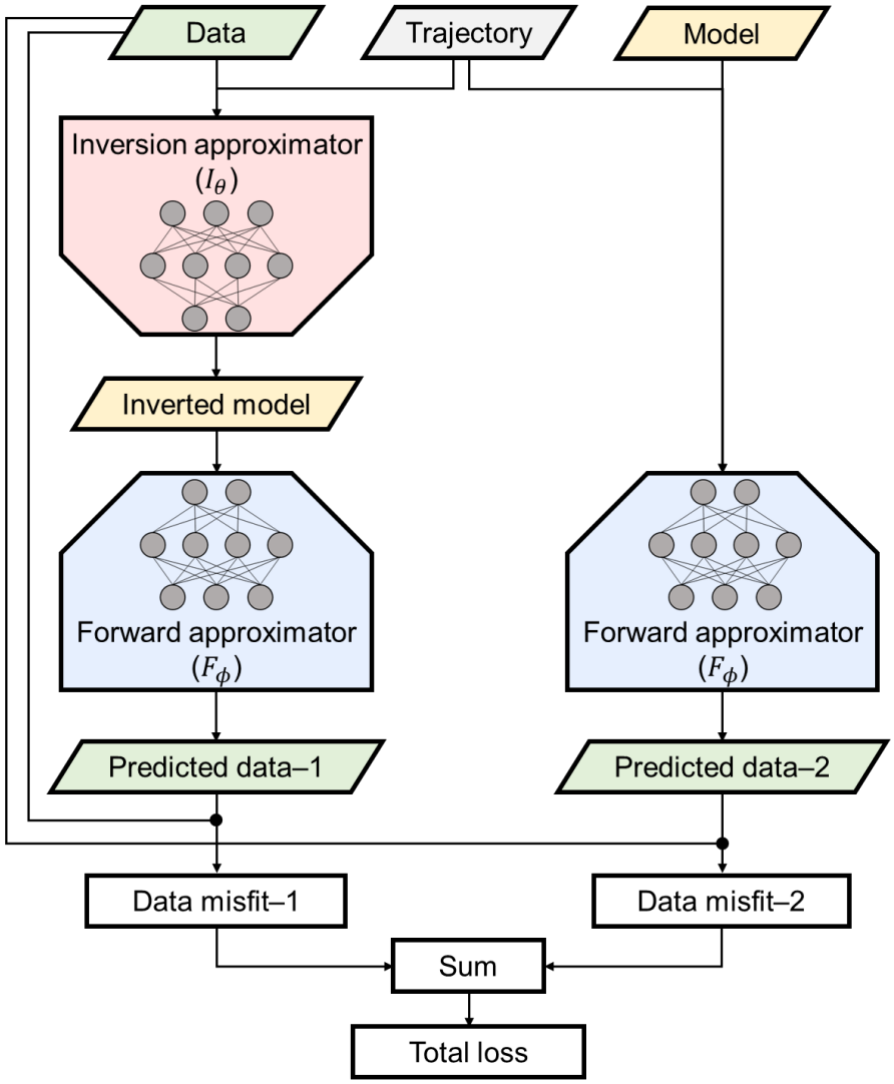


Figure 1. Workflow describing the calculation of the loss function for the physics-guided deep-learning inversion developed in this study. Two data misfit terms: Data misfit−1 and −2 in the flowchart correspond to the first and second terms of the right-hand-side of equation 1, respectively.

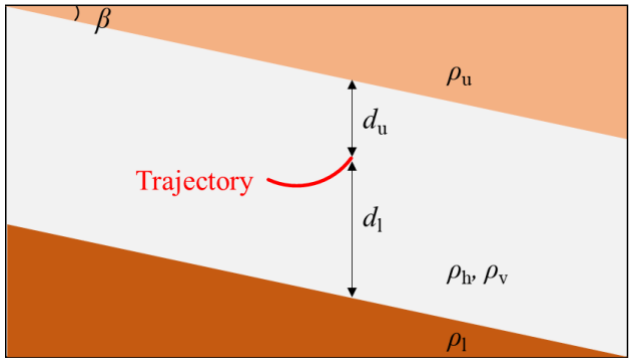


Figure 2. Graphical representation of the one-dimensional resistivity model assumed in this study. The red line indicates the local well trajectory.

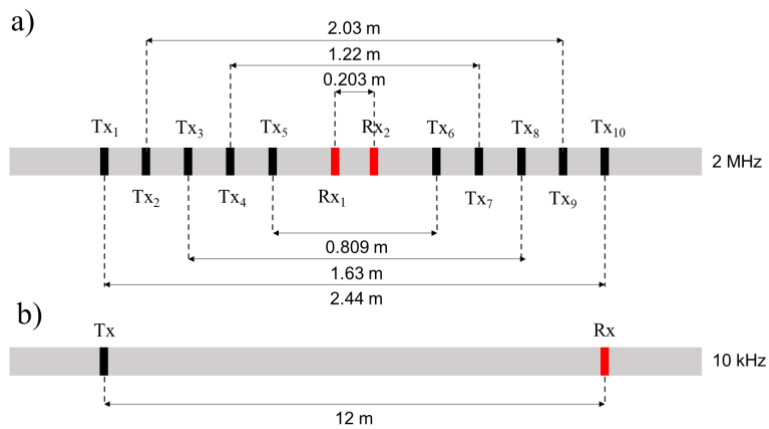


Figure 3. a) Conventional LWD and b) Deep azimuthal logging instruments assumed in this study. Each transmitter and receiver include three mutually orthogonal coils.

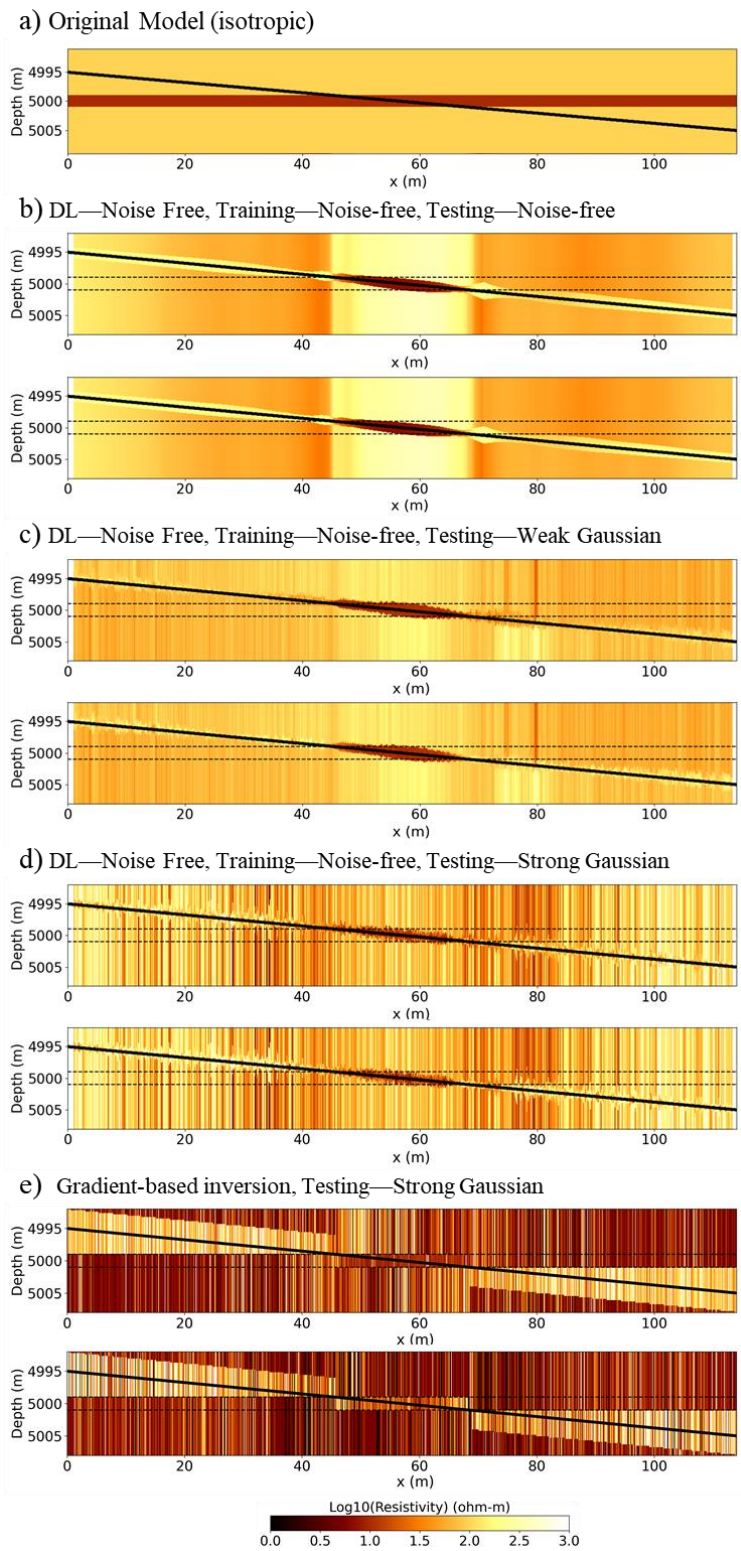
a) Original Model (isotropic)
b) DL—Noise Free, Training—Noise-free, Testing—Noise-free
c) DL—Noise Free, Training—Noise-free, Testing—Weak Gaussian
d) DL—Noise Free, Training—Noise-free, Testing—Strong Gaussian
e) Gradient-based inversion, Testing—Strong Gaussian
Depth (m)
4995
5000
5005
x (m)
0
20
40
60
80
100
Log10(Resistivity) (ohm-m)
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Figure 4. (a) Original isotropic resistivity model for which we implement anisotropic inversion. Resistivity models estimated with: (b) deep learning (DL) inversion of noise-free data, (c) DL inversion of weak Gaussian-noisy data, (d) DL inversion of strong Gaussian-noisy data, and (e) gradient-based inversion of strong Gaussian-noisy data. The upper and lower columns of each inversion result show inverted horizontal and vertical resistivities, respectively. Dashed lines in the inversion results identify the locations of bed boundaries.

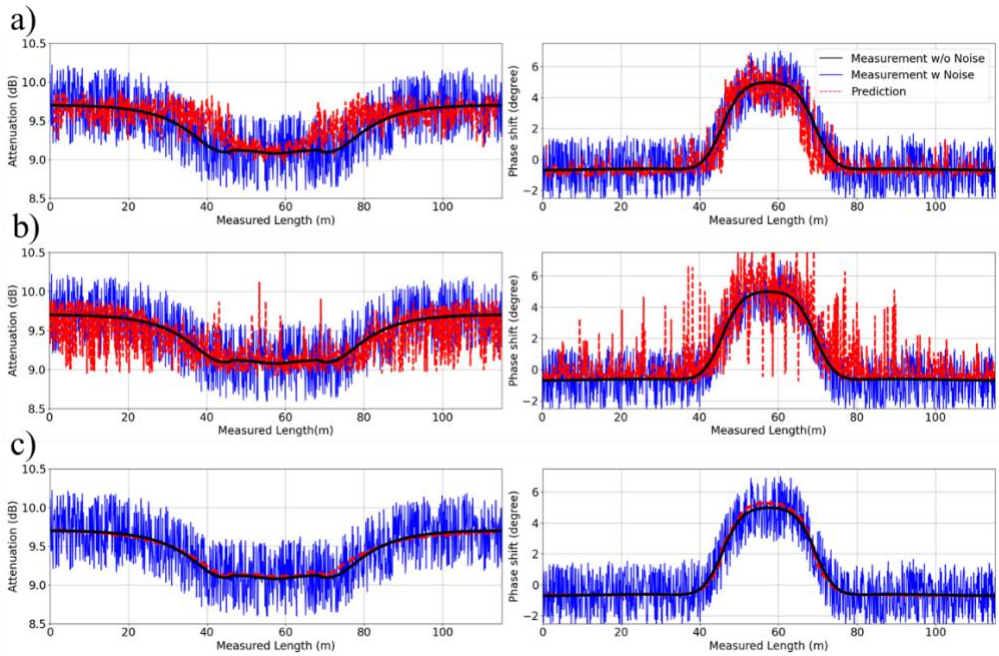


Figure 5. Comparison of noise-free measurements (black solid lines), strongly noise-contaminated measurements (blue solid lines), and predicted measurements obtained with (a) DL inversion without measurement noise during training (corresponding inversion results in Fig. 4d), (b) gradient-based inversion (Fig. 4e), and (c) DL inversion with measurement noise during training but with the proposed remedy of combining augmented measurement noise and architectural noise (Fig. 8d). Left- and right-hand columns display attenuation and phase difference components, respectively, of nearest Tx-Rx pair corresponding to the conventional LWD tool shown in Fig. 3a.

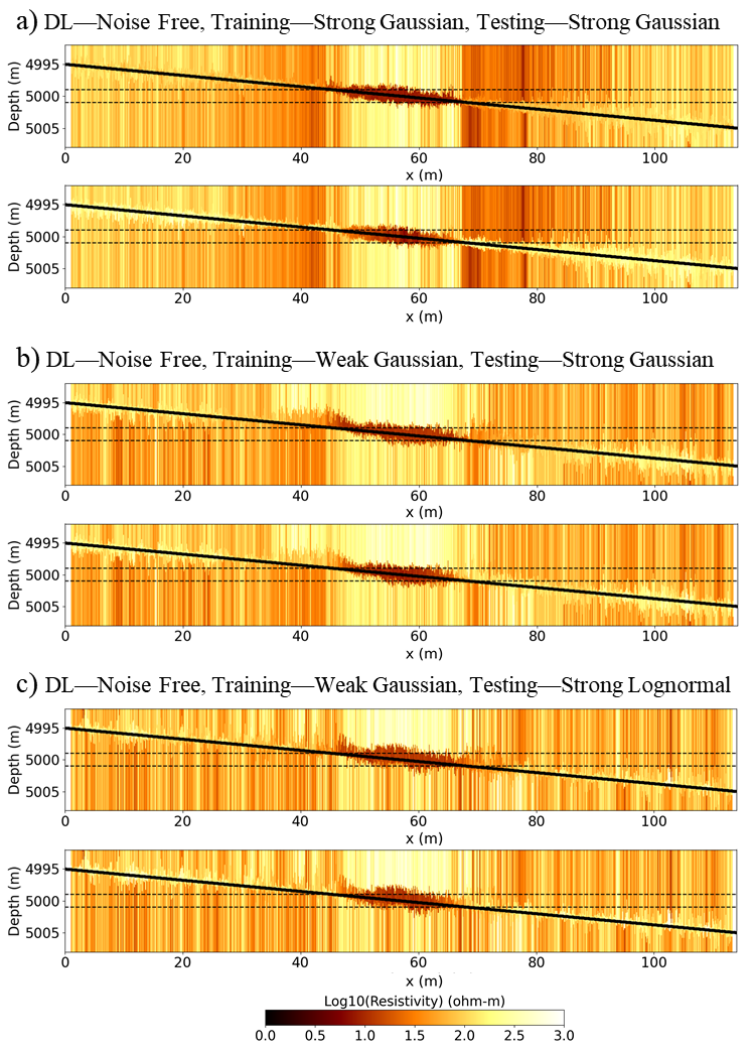


Figure 6. Inverted resistivity models predicted by the DL inversion algorithms of strongly noisy data. Each panels show inversion results obtained from test data contaminated with (a) and (b) strong Gaussian noise and (c) strong Log-normal noise obtained with the DL architecture trained using (a) strong Gaussian noisy measurements, (b) weak Gaussian noisy measurements, and (c) weak Gaussian noisy measurements. No architecture noise was assumed in this experiment. Upper and lower columns of each inversion result show inverted horizontal and vertical resistivities, respectively. Dashed lines in the inversion result identify the location of bed boundaries.

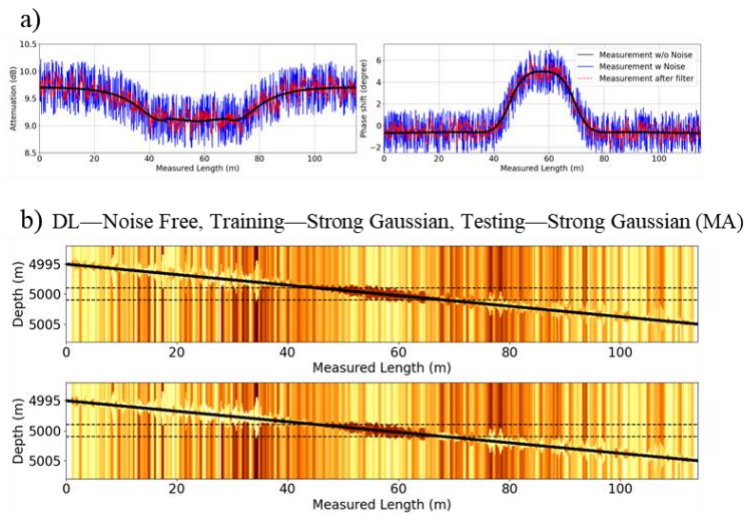


Figure 7. (a) Comparison of noise-free (black solid lines), strong Gaussian noise-contaminated (blue solid lines), and moving average filtered (red dashed lines) measurements. Left- and right-hand columns display attenuation and phase-difference components, respectively, of the nearest Tx-Rx pair corresponding to the conventional LWD tool shown in Fig. 3a. (b) Inverted resistivity models obtained with the deep learning inversion algorithms of the moving-average filtered data. The result is obtained with the deep learning architecture trained with noise-free measurements that are used to obtain the inversion results shown in Fig. 4d. Upper and lower columns of the inversion result show inverted horizontal and vertical resistivities, respectively. Dashed lines in the inversion result identify the location of bed boundaries.

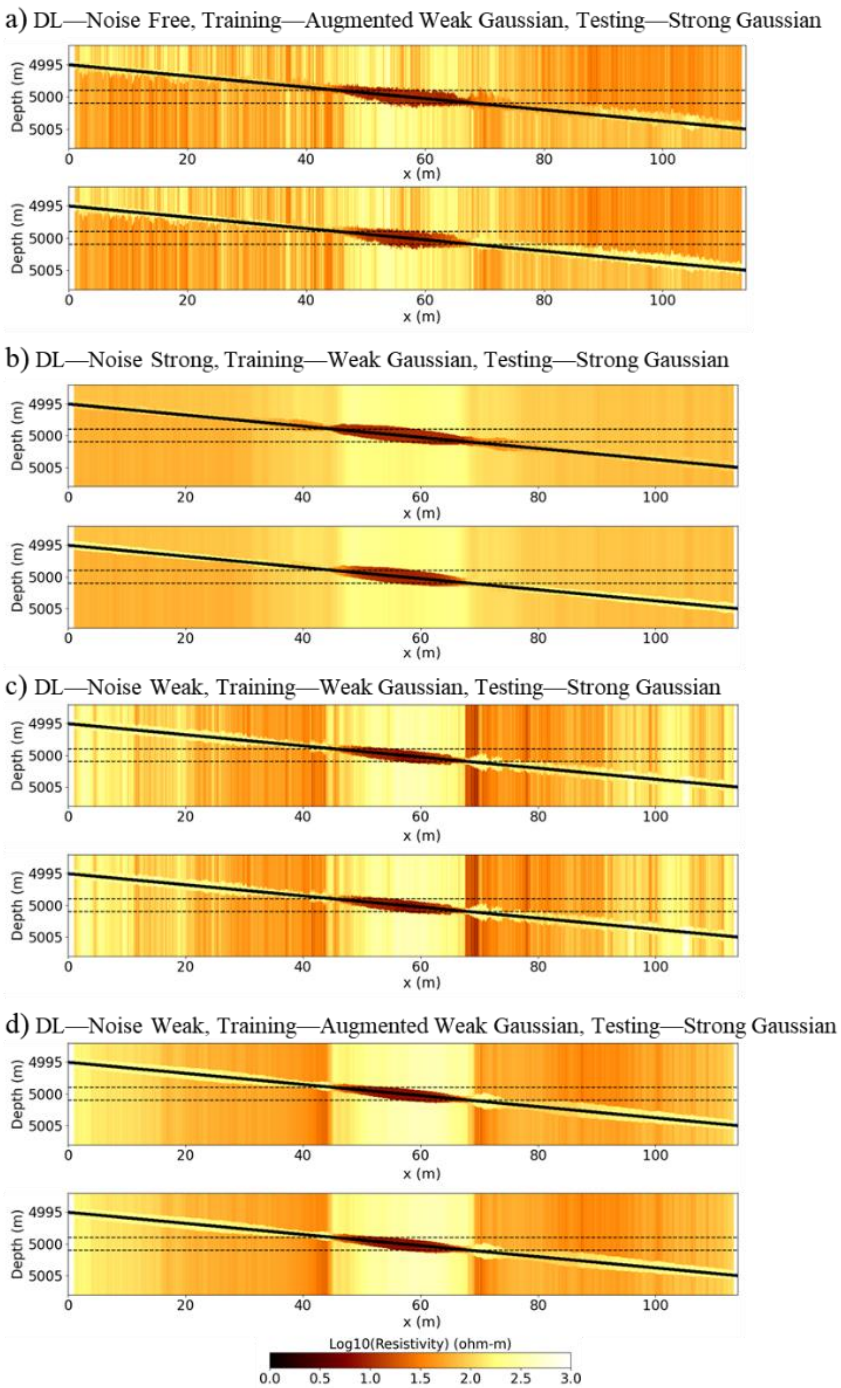


Figure 8. Inverted resistivity models obtained with deep learning inversion algorithms of strongly Gaussian noisy data. We consider four deep learning architectures trained with: (a) augmented weak Gaussian noisy measurements and noise-free architecture, (b) weak Gaussian noisy measurements and strong architecture noise, (c) weak Gaussian noisy measurements and weak architecture noise, and (d) augmented weak

Gaussian noisy measurements and weak architecture noise. Upper and lower columns show inverted horizontal and vertical resistivities, respectively. Dashed lines in the inversion results identify the locations of bed boundaries.

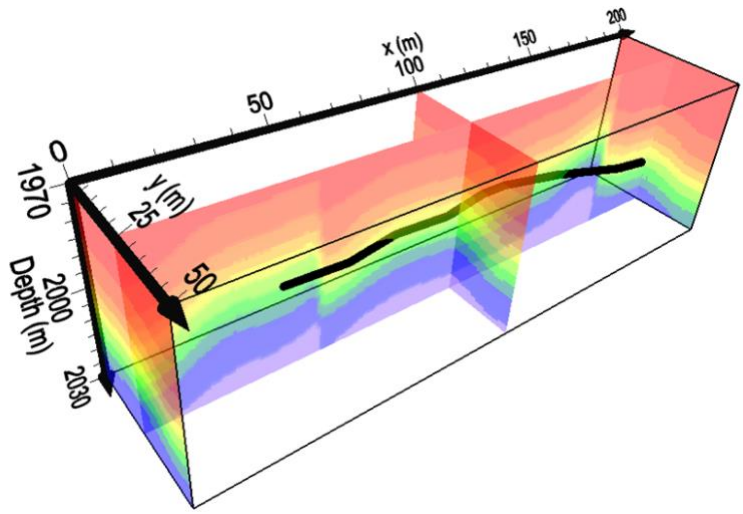


Figure 9. Three-dimensional plot of a complex earth model consisting of ten layers and two vertical faults located at x=50 and x=170 m. The black solid line identifies the well trajectory, while the background color indicates the layer indexes.

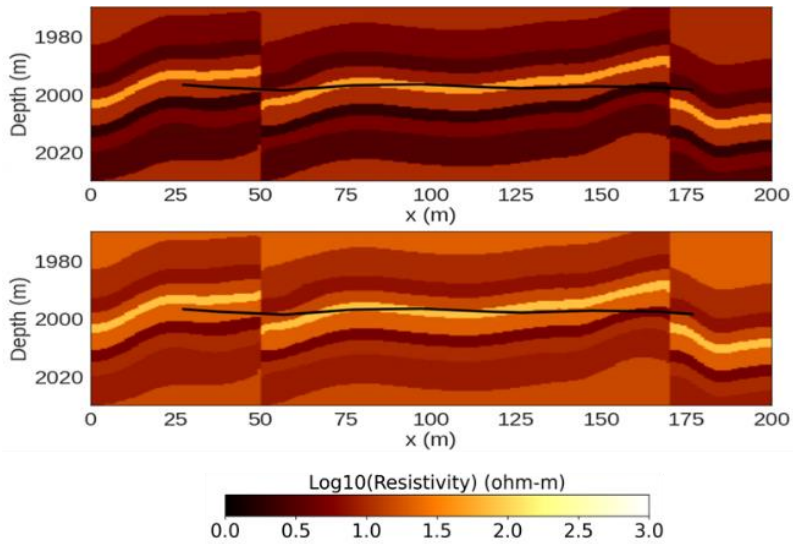


Figure 10. Cross-sections of the horizontal (upper) and vertical (lower) resistivity model in x-depth space at y=25 m shown in Fig. 9. The three-dimensional well trajectory is projected onto the two-dimensional space and displayed with a black line.

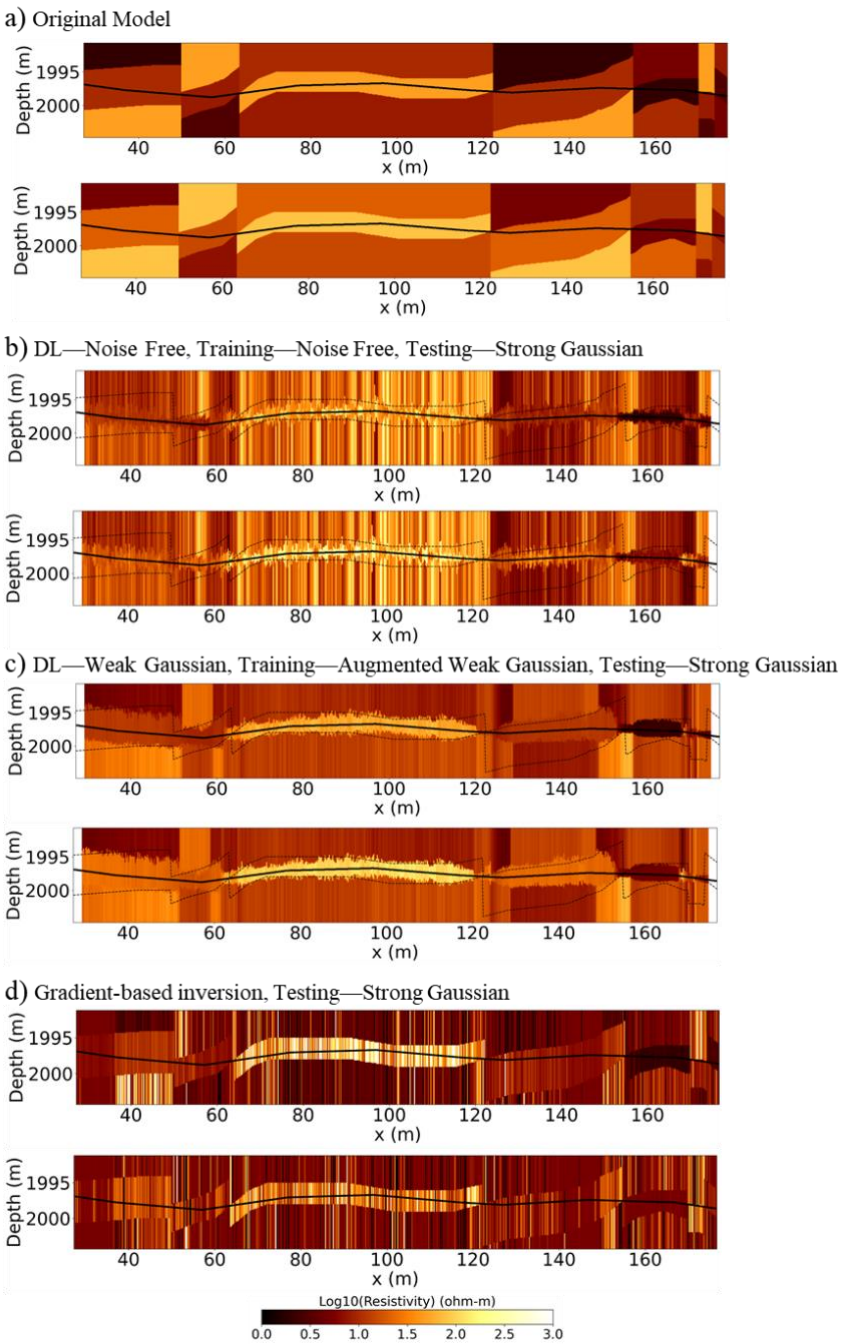


Figure 11. (a) Original three-layer resistivity model along the well trajectory. Panels b), c), and d) display inverted resistivity models from the strongly noisy data obtained with (b) deep learning inversion trained without measurement and architecture noise, (c) deep learning inversion trained with augmented weak measurement noise and optimal architecture noise, and (d) gradient-based inversion. The upper and lower columns show original or inverted horizontal and vertical resistivities, respectively.

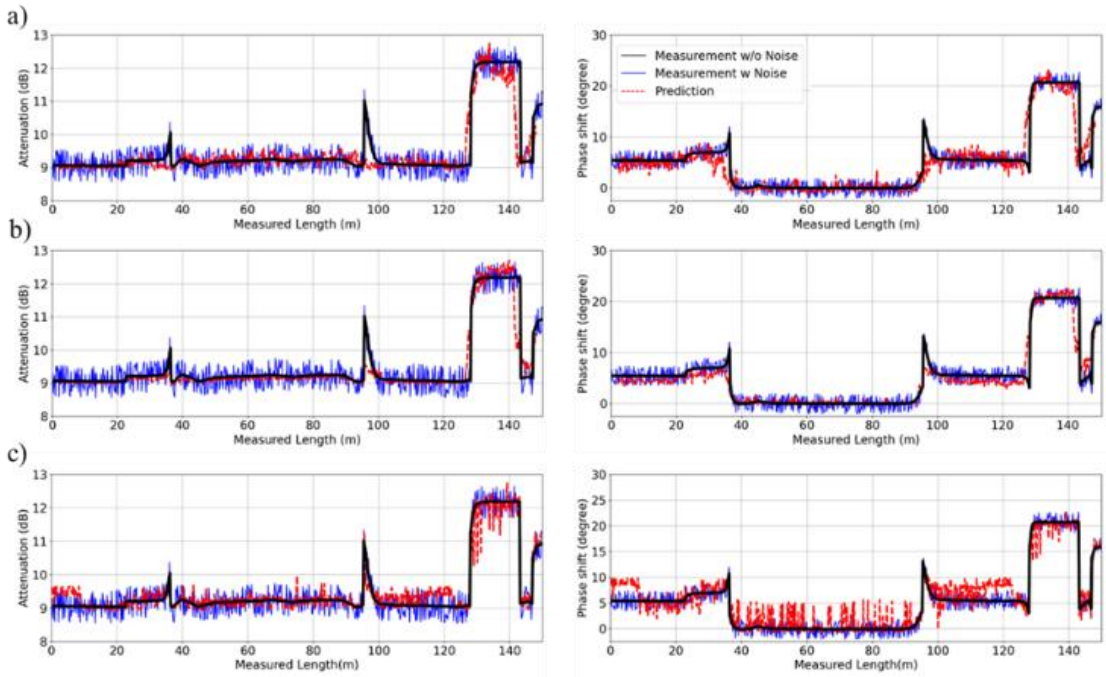


Figure 12. Comparison of noise-free measurements (black solid lines), strongly noise-contaminated measurements (blue solid lines), and predicted measurements obtained with (a) deep learning inversion trained without measurement and architecture noise, (b) deep learning inversion with augmented weakly noisy data and optimal architecture noise, and (c) gradient-based inversion. The corresponding inversion results are shown in Fig. 11b, Fig. 11c, and Fig. 11d, respectively. Left- and right-hand columns display attenuation and phase difference components, respectively, of the nearest Tx-Rx pair corresponding to the conventional LWD tool shown in Fig. 3a.

Table 1. Different deep-learning architectures according to the noise intensity (standard deviation, dimensionless) added to the architecture. Noisy layers are activated only during the training process.

| **Architecture** | **Standard Deviation** |
|---|---|
| DL—Noise Free | - |
| DL—Weak Gaussian | $10^{-5}$ |
| DL—Strong Gaussian | $10^{-1}$ |

Table 2. Maximum measurement-noise intensities used in the training and validation data sets. Intensity values were used for both Gaussian and log-normal distributions.

| | **LWD measurement** [Attenuation, Phase shift] | **Deep measurement** [Attenuation, Phase shift] |
|---|---|---|
| Weak noise | [0.1 dB, 0.4 degrees] | [0.004 dB, 0.4 degrees] |
| Strong noise | [0.5 dB, 2.0 degrees] | [0.02 dB, 2.0 degrees] |

Table 3. Different training data sets with varying measurement noise levels and number of samples. The level of noise intensity is described in Table 1.

| **Dataset** | **# of Samples** |
|---|---|
| Noise Free | 100K |
| Weak Gaussian | 100K |
| Strong Gaussian | 100K |
| Strong Lognormal | 100K |
| Augmented Weak Gaussian | 300K |

Table 4. Root-mean-square deviation of inversion results between predicted and (1) noisy measurement (*RMSD-Noisy*) or (2) noise-free measurements (*RMSD-Noise free*). The corresponding inverted models are also shown in the third column from the left. The minimum *RMSD-Noisy* and *RMSD-Noise free* for the strongly noisy input measurement are shown with bold numbers.

| Ex. # | Inversion Method | Inversion Results | RMSD–Noisy (Attenuation, Phase shift) | RMSD–Noise free (Attenuation, Phase shift) |
|---|---|---|---|---|
| 1 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Noise Free | Fig. 4b | (0.074, 0.469) | (0.074, 0.469) |
| 2 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Weak Gaussian | Fig. 4c | (0.125, 0.716) | (0.092, 0.516) |
| 3 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Strong Gaussian | Fig. 4d | (0.334, 1.377) | (0.194, 0.713) |
| 4 | • Gradient-based<br>• Testing—Strong Gaussian | Fig. 4e | (0.346, 1.655) | (0.282, 1.194) |
| 5 | • DL—Noise Free<br>• Training—Strong Gaussian<br>• Testing—Strong Gaussian | Fig. 6a | (0.290, 1.370) | (0.144, 0.530) |
| 6 | • DL—Noise Free<br>• Training—Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 6b | (0.314, 1.310) | (0.185, 0.706) |
| 7 | • DL—Noise Free<br>• Training—Weak Gaussian<br>• Testing—Strong Lornormal | Fig. 6c | (0.378, 1.913) | (0.238, 0.973) |
| 8 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Strong Gaussian (Filtered) | Fig. 7b | (0.275, 1.287) | (0.207, 0.726) |
| 9 | • DL—Noise Free<br>• Training—Augmented Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 8a | (0.302, 1.274) | (0.125, 0.502) |
| 10 | • DL—Strong Gaussian<br>• Training—Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 8b | (0.315, 1.328) | (0.152, 0.626) |
| 11 | • DL—Weak Gaussian<br>• Training—Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 8c | (0.291, 1.257) | (0.093, 0.499) |
| 12 | • DL—Weak Gaussian<br>• Training—Augmented Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 8d | **(0.285, 1.219)** | **(0.084, 0.483)** |

Table 5. The normalized pixel-wise model mismatch between true and inverted resistivity models calculated using Equation 4. The corresponding inverted models are shown in the third column from the left.

| Ex. # | Inversion Method | Inversion Results | Model mismatch |
|---|---|---|---|
| 1 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Noise Free | Fig. 4b | 0.086 |
| 3 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Strong Gaussian | Fig. 4d | 0.184 |
| 4 | • Gradient-based<br>• Testing—Strong Gaussian | Fig. 4e | 0.289 |
| 9 | • DL—Noise Free<br>• Training—Augmented Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 8a | 0.136 |
| 12 | • DL—Weak Gaussian<br>• Training—Augmented Weak Gaussian<br>• Testing—Strong Gaussian | Fig. 8d | **0.094** |

Table 6. Root-mean-square deviation of inversion results between predicted and noisy measurement (*RMSD-Noisy*) or noise-free measurements (*RMSD-Noise free*). The corresponding inverted models are shown in the third column from the left.

| **Ex. #** | **Inversion Method** | **Inversion Results** | **RMSD–Noisy** (Attenuation, Phase shift) | **RMSD–Noise free** (Attenuation, Phase shift) |
|---|---|---|---|---|
| C-1 | • DL—Noise Free<br>• Training—Noise Free<br>• Testing—Strong Gaussian | Fig. 11b | (0.442, 2.256) | (0.336, 1.636) |
| C-2 | • DL—Weak Gaussian<br>• Training—Augmented Weak Gaussian<br>• Testing—Strong Guassian | Fig. 11c | (0.377, 1.817) | (0.186, 0.974) |
| C-3 | • Gradient-based<br>• Testing—Strong Gaussian | Fig. 11d | (0.346, 1.655) | (0.282, 1.194) |

**ACKNOWLEDGEMENTS**

This work reported in this paper was supported by the University of Texas at Austin's Research Consortium on Formation Evaluation, jointly sponsored by Aramco, BHP Billiton, BP, Chevron, CNOOC International, ConocoPhillips, COSL, ENI, Equinor ASA, Haliburton, INPEX, Lundin-Norway, Oil Search Alaska, Oxy, Petrobras, Repsol, Schlumberger, Todd Energy, Total, and Wintershall. Carlos Torres-Verdín is grateful for the financial support provided by the Brian James Jennings Memorial Endowed Chair in Petroleum and Geosystems Engineering. The authors also acknowledge the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing HPC resources that have contributed to the research results reported in this paper. URL: http://www.tacc.utexas.edu

**DATA AVAILABILITY**

Data or code relating to this work is available upon request.

## REFERENCES


Abubakar, A., Habashy, T.M., Druskin, V., Knizhnerman, L. & Davydycheva, S. (2006) A 3D parametric inversion algorithm for triaxial induction data. *Geophysics*, **71**. doi:10.1190/1.2168009

Araya-Polo, M., Jennings, J., Adler, A. & Dahlke, T. (2018) Deep-learning tomography. *Leading Edge*, **37**. doi:10.1190/tle37010058.1

Bakr, S.A., Pardo, D. & Torres-Verdín, C. (2017) Fast inversion of logging-while-drilling resistivity measurements acquired in multiple wells. *Geophysics*, **82**. doi:10.1190/GEO2016-0292.1

Bishop, C. M. (1995) Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, **7**

Chen, H., Gao, J., Zhang, W. & Yang, P. (2021) Seismic Acoustic Impedance Inversion via Optimization-Inspired Semisupervised Deep Learning. *IEEE Transactions on Geoscience and Remote Sensing*. doi:10.1109/TGRS.2021.3107257

Chollet, F., 2015, Keras, https://github.com/fchollet/keras

Colombo, D., Turkoglu, E., Li, W., Sandoval-Curiel, E. & Rovetta, D. (2021) Physics-driven deep-learning inversion with application to transient electromagnetics. *Geophysics*, **86**. doi:10.1190/geo2020-0760.1

Côrte, G., Dramsch, J., Amini, H. & MacBeth, C. (2020) Deep neural network application for 4D seismic inversion to changes in pressure and saturation: Optimizing the use of synthetic training datasets. *Geophysical Prospecting*, **68**. doi:10.1111/1365-2478.12982

Deleersnyder, W., Maveau, B., Hermans, T. & Dudal, D. (2021) Inversion of electromagnetic induction data using a novel wavelet-based and scale-dependent regularization term. Geophys J Int, 226. doi:10.1093/gji/ggab182

Goodfellow, I., Bengio, Y. & Courville, A. (2016) Regularization for deep learning *in Deep Learning*, MIT Press, ch. 7, sec. 4, pp. 236-238.

Gulcehre, C., Moczulski, M., Denil, M. & Bengio Y. (2016) Noisy activation functions, *in Proc. PMLR*, New York, USA. pp. 3059-3068.

de Groot–Hedlin, C. (2000) Smooth inversion of induction logs for conductivitymodels with mud filtrate invasion. Geophysics, 65, pp.1648-1475. doi: 10.1190/1.1444835

He, K., Zhang, X., Ren, S. & Sun, J. (2016a) Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2016-December. doi:10.1109/CVPR.2016.90

He, K., Zhang, X., Ren, S. & Sun, J. (2016b) Identity Mappings in Deep Residual Networks. arXiv preprint. arXiv:1603.05027v3

Hu, X. & Fan, Y. (2018) Huber inversion for logging-while-drilling resistivity measurements in high angle and horizontal wells. *Geophysics*, **83**. doi:10.1190/geo2017-0459.1

Hu, Y., Guo, R., Jin, Y., Wu, X., Li, M., Abubakar, A. & Chen, J. (2020) A Supervised Descent Learning Technique for Solving Directional Electromagnetic Logging-While-Drilling Inverse Problems. *IEEE Transactions on Geoscience and Remote Sensing*, **58**. doi:10.1109/TGRS.2020.2986000

Jin, Y., Wu, X., Chen, J. & Huang, Y. (2019) Using a physics-driven deep neural network to solve inverse problems for LWD azimuthal resistivity measurements. *SPWLA 60th Annual Logging Symposium 2019*. doi:10.30632/T60ALS-2019_IIII

Lecun Y. & Bengio Y. (1998) Convolutional networks for images, speech, and time series. *in the handbook of brain theory and neural networks*, MIT Press, pp. 276-278

Lee, H.O., Teixeira, F.L., San Martin, L.E. & Bittar, M.S. (2012) Numerical modeling of eccentered LWD borehole sensors in dipping and fully anisotropic earth formations. *IEEE Transactions on Geoscience and Remote Sensing*, **50**. doi:10.1109/TGRS.2011.2162736

Li, S., Liu, B., Ren, Y., Chen, Y., Yang, S., Wang, Y. & Jiang, P. (2020) Deep-Learning Inversion of Seismic Data. *IEEE Transactions on Geoscience and Remote Sensing*, **58**. doi:10.1109/TGRS.2019.2953473

Marchant, D., Clegg, N., Rawsthorne, L. & Kunnas, J. (2019) 3D inversion of electromagnetic logging-while-drilling data. *ASEG Extended Abstracts*, **2019**. doi:10.1080/22020586.2019.12073097

Nair, V. & Hinton, G. (2010) Rectified linear units improve restricted Boltzmann machines. *In Proc. ICML*. Haifa, Israel, Jun., pp. 807-814.

Neelakantan, A., Vilnis, L., Le, Q. V., Sutskever, I., Kaiser, L., Kurach, K. & Martens, J. (2015) Adding Gradient Noise Improves Learning for Very Deep Networks, *arXiv preprint*, arXiv:1511.06807

Nie, X.C., Yuan, N. & Liu, C.R. (2010) Simulation of LWD tool response using a fast integral equation method. *IEEE Transactions on Geoscience and Remote Sensing*, **48**. doi:10.1109/TGRS.2009.2027112

Noh, K., Pardo, D. & Torres-Verdin, C. (2021) Real-time 2.5D inversion of LWD resistivity measurements using deep learning for geosteering applications across faulted formations. *IEEE Geoscience and Remote Sensing Letters*, **19**. doi:10.1109/LGRS.2021.3128965

Noh, K., Pardo, D. & Torres-Verdin, C. (2022) 2.5-D Deep Learning Inversion of LWD and Deep-Sensing em Measurements Across Formations with Dipping Faults. *Petrophysics*, **63**. doi:10.30632/PJV63N4-2022a2

Pardo, D. & Torres-Verdín, C. (2015) Fast 1D inversion of logging-while-drilling resistivity measurements for improved estimation of formation resistivity in high-angle and horizontal wells. *Geophysics*, **80**. doi:10.1190/GEO2014-0211.1

Pardo, D., Matuszyk, P. J., Puzyrev, V., Torres-Verdín, C., Nam, M. J. & Calo, V. M. (2021) Chapter 2.9: 1.5D Variational Formulations *in Modeling of Resistivity and Acoustic Borehole Logging Measurements Using Finite Element Methods*, Elsevier, 2021

Poole, B., Sohl-Dickstein, J. & Ganguli, S. (2014) Analyzing noise in autoencoders and deep networks, *arXiv preprint* arXiv:1406.1831v1

Puzyrev, V., Torres-Verdín, C., and Calo, V. (2018) Interpretation of deep directional resistivity measurements acquired in high-angle and horizontal wells using 3D inversion. *Geophysical Journal International*, **213**. doi: 10.1093/gji/ggy047

Shahriari, M., Rojas, S., Pardo, D., Rodríguez-Rozas, A., Bakr, S. A., Calo, V. M., Muga, I. (2018) A Numerical 1.5D Method for the Rapid Simulation of Geophysical Resistivity Measurements. *Geosciences*, **8**(6). https://doi.org/10.3390/geosciences8060225

Shahriari, M., Pardo, D., Rivera, J.A., Torres-Verdín, C., Picon, A., Ser, J. del, Ossandón, S., *et al.* (2021) Error control and loss functions for the deep learning inversion of borehole resistivity measurements. *Int J Numer Methods Eng*, **122**. doi:10.1002/nme.6593

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014) Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, **15**, pp. 1929–1958

Tabarovsky, L.A. & Rabinovich, M.B. (1998) Real time 2D inversion of induction logging data. *J Appl Geophy*, **38**. doi:10.1016/S0926-9851(97)00034-7

Thiel, M., Bower, M. & Omeragic, D. (2018) 2D reservoir imaging using deep directional resistivity measurements. *Petrophysics*, **59**. doi:10.30632/PJV59N2-2018a7

Wang, G.L., Barber, T., Wu, P., Allen, D. & Abubakar, A. (2017) Fast inversion of triaxial induction data in dipping crossbedded formations. *Geophysics*, **82**. doi:10.1190/GEO2015-0610.1

Wang, L., Li, H. & Fan, Y. (2019) Bayesian Inversion of Logging-While-Drilling Extra-Deep Directional Resistivity Measurements Using Parallel Tempering Markov Chain Monte Carlo Sampling. *IEEE Transactions on Geoscience and Remote Sensing*, **57**. doi:10.1109/TGRS.2019.2917839

Wang, L., Liu, Y., Wang, C., Fan, Y. & Wu, Z. (2021) Real-time forward modeling and inversion of logging-while-drilling electromagnetic measurements in horizontal wells. *Petroleum Exploration and Development*, **48**. doi:10.1016/S1876-3804(21)60012-5

Wei, N., Yang, D., Wang, Z. & Lu, Y. (2022) Joint 3D inversion of gravity and magnetic data using deep learning neural networks. Second International Meeting for Applied Geoscience & Energy, 1457-1461

Zhang, Z., Yu, L., Kriegshäuser, B. & Lev, T. (2004) Determination of relative angles and anisotropic resistivity using multicomponent induction logging data. *Geophysics*, **69**. doi:10.1190/1.1778233

## Appendix

*Neural Network Architecture*

For the approximation of forward and inverse functions, we employ the neural network architectures described in Table A1. We consider the same structure of ResNet for both forward and inverse modules with proper input and output shapes for each purpose. We also consider fully connected layers for the forward module. The fully connected layers with ten series of dense, batch normalization, ReLU, and Dropout layers are implemented as a feature extractor to properly feed to the ResNet the input resistivity structure. Following the method in the original work on ResNet (He *et al.* 2016), down-sampling is performed at the beginning of the groups of residual blocks starting from the second block (i.e., Conv 2_1, Conv 3_1, Conv 4_1, Conv 5_1, Conv 6_1 in Table A1) with a stride of two. The use of Gaussian noise layers after every 1D convolutional layer in the ResNet architecture is intended to appraise the approach's effectiveness and its robustness in the presence of measurement noise.

Table A1. Description of our neural network architecture. We consider independent ResNet architectures for the forward and inverse modules. Numbers inside curly, angle, and square brackets denote the shape of vectors or matrices, the number of nodes for dense layers, and the shape and numbers of convolution filters, respectively. Dense, BN, 1D Conv., and Gaussian noise in sublayers refer to the regular deeply connected layer, batch normalization layer, one-dimensional convolution layer, and zero-centered Gaussian noise layer, respectively. The possible inclusion of Gaussian noise layers after the 1D convolutional layers are described in parentheses.

| **Layers** | **Forward** | | **Inverse** | | **Sub layers** |
|---|---|---|---|---|---|
| Input | Model $\{7\}$ | Trajectory $\{21\}$ | Measurement $\{21\times24\}$ | Trajectory $\{21\}$ | - |
| FC 1 | $\langle 21\rangle\times10$ | - | - | - | Dense, BN, ReLU, Dropout |
| Concatenation | $\{21\times1\}, \{21\} \rightarrow \{21\times2\}$ | | $\{21\times24\}, \{21\} \rightarrow \{21\times25\}$ | | - |
| Conv 1_x | $\begin{bmatrix}1, 32\\3, 32\\1, 64\end{bmatrix}\times2$ | | | | 1D Conv., (Gaussian), BN, ReLU |
| Conv 2_x | $\begin{bmatrix}1, 64\\3, 64\\1, 128\end{bmatrix}\times3$ | | | | 1D Conv., (Gaussian), BN, ReLU |
| Conv 3_x | $\begin{bmatrix}1, 64\\3, 64\\1, 128\end{bmatrix}\times4$ | | | | 1D Conv., (Gaussian), BN, ReLU |
| Conv 4_x | $\begin{bmatrix}1, 128\\3, 128\\1, 256\end{bmatrix}\times4$ | | | | 1D Conv., (Gaussian), BN, ReLU |
| Conv 5_x | $\begin{bmatrix}1, 128\\3, 128\\1, 256\end{bmatrix}\times4$ | | | | 1D Conv., (Gaussian), BN, ReLU |
| Conv 6_x | $\begin{bmatrix}1, 128\\3, 128\\1, 256\end{bmatrix}\times4$ | | | | 1D Conv., (Gaussian), BN, ReLU |
| FC 2 | $\langle 462\rangle$ | | $\langle 7\rangle$ | | Dense |
| Output | Measurement $\{21\times24\}$ | | Resistivity Model $\{7\}$ | | - |