Estimating subsurface properties using a semisupervised neural network approach

Haibin Di¹ and Aria Abubakar¹

ABSTRACT

Estimating static rock properties (e.g., density and porosity) from seismic and well logs is one of the essential but challenging tasks in subsurface interpretation and characterization. To compensate for the sparsity of well logs and the limited bandwidth of seismic data, a semisupervised learning workflow is used for efficiently integrating seismic and logs and simultaneously estimating multiple subsurface properties. It consists of two components: (1) unsupervised seismic feature engineering and (2) supervised seismic-well integration, each of which is implemented as a convolutional neural network. Compared to most existing methods, it is advanced in three aspects. First, it allows the use of local 3D seismic patterns for building an optimal nonlinear mapping function with 1D logs, which is more noise robust and significantly improves the lateral consistency of machine prediction throughout the entire seismic survey. Second, it is capable of automatically bridging the gap of vertical resolution between seismic and well logs, which simplifies the workflow of data preparation, such as log upscaling. In addition, it enables Monte Carlo dropout-based epistemic uncertainty analysis. The performance of our solution is evaluated through two examples: relative acoustic impedance and porosity estimation in a synthetic PreSDM data set of 36 pseudowells and sonic and density estimation in the Groningen data set of 375 wells. The good match between the machine predictions and the actual measurements demonstrates the capability of our semisupervised learning method in providing reliable seismic and well integration and delivering robust estimation of subsurface properties, including those of a relatively weak physical link with seismic, such as density and porosity.

INTRODUCTION

Robust estimation of static rock properties in the subsurface, including but not exclusively, density and velocity, is crucial to the success of geologic characterization and evaluation. Seismic surveying and well logging are two important approaches for measuring subsurface rock properties, and both have pros and cons. Specifically, well logging measures the rocks by directly drilling into the subsurface and thus is considered most accurate; however, due to the high expense of drilling, these measurements are usually in one dimension and are moreover limited in amount as well as spatial distribution. Contrarily, seismic surveying provides a large spatial coverage of the study area; however, it reflects the subsurface geology in an indirect way by stimulating seismic wave propagation through the rocks. Given the current seismic acquisition technologies, the collected seismic data have a limited bandwidth, which introduces a nonunique-ness into rock property inversion when purely from seismic data (Di et al., 2020a). A reliable integration of seismic and well logging is expected for more robust property estimation (Caron, 1989).

With the recent emergence of machine learning (ML), the popular convolutional neural network (CNN), recurrent neural network, and generative adversarial network (GAN) have been introduced to the seismic domain and implemented for building an optimal nonlinear mapping function between seismic signals and the rock properties measured at the wells (e.g., Alfarraj and AlRegib, 2019; Das et al., 2019; Mustafa et al., 2019, 2020; Wang et al., 2019; Cai et al., 2020; Di et al., 2020a). To the best of our knowledge, there exist four major challenges for robust ML-based property estimation. The first challenge is how to link 1D well logs with 3D seismic patterns, instead of single seismic traces. Otherwise, even subtle lateral variations in seismic signals would greatly disturb the link and are likely to increase the risk of unstable prediction, causing
artifacts such as spikes or blobs. The second challenge is how to bridge the gap in the vertical resolution between seismic and well logs. As a traditional treatment, downsampling well logs to the same scale as seismic not only requires human judgment and manual effort, but it also undesirably runs the risk of eliminating useful information available in well logs. Another major challenge is how to quantify the uncertainty in machine prediction, which comes from the nonuniqueness between the actual rock properties and the observed data (Tarantola and Valette, 1982; Mavko and Mukerji, 1998; Claerbout, 1999; Takahashi, 2000). Recently, Gal and Ghahramani (2016) groundbreakingly prove the use of dropout (Srivastava et al., 2014) in prediction to be an approximated variational solution to the Bayesian deep learning model. Due to its simple implementation, the Monte Carlo (MC) dropout method has recently been implemented into various applications, such as biomedical image processing (Tanno et al., 2017; Nair et al., 2020) and seismic image segmentation (Zhao and Chen, 2020; Pham and Fomel, 2021). However, it has not been adopted into deep neural network-based property estimation. The fourth challenge in seismic-well integration (SWI) through ML is how to deal with the sparsity of well logs, which limits the amount of training data and causes the learning to be highly prone to overfitting (Di et al., 2020a).

This work aims at addressing the preceding challenges by developing a semisupervised learning workflow for subsurface property estimation from seismic and well logs. First, we describe how such semisupervised learning works by integrating two CNNs, with the first CNN learning the features present in the given seismic volume in an unsupervised way and the second CNN mapping seismic data with well logs on the basis of the common features prelearned in the first CNN, respectively. The next section demonstrates how the proposed workflow performs on two example data sets: PreSDM and Groningen.

WORKFLOW DESCRIPTION

Semisupervised learning

As shown in Figure 1, the proposed semisupervised learning workflow for subsurface property estimation consists of two key components, each of which is formulated as a CNN and illustrated next in detail.

1) Seismic feature engineering (SFE), which aims at understanding the target seismic data in an unsupervised or self-supervised way. Correspondingly, the trained SFE CNN can be considered knowledgeable of the common features throughout the entire seismic survey at certain levels. There exist various approaches for building such an SFE CNN, and some examples are the autoencoder, dynamic filtering, Bayesian-based methods (Li et al., 2019), and CNN (Di et al., 2020b). In this work, the SFE CNN is in the architecture of a denoising convolutional autoencoder (CAE), in which the encoder and decoder consist of six layers (Figure 2). In both examples shown in the next section, the seismic tiles used for training are of size 128 samples (height) by 128 adjacent traces (width) by five adjacent sections (depth), they are randomly retrieved within the given 3D seismic, and 10% Gaussian noise is added for denoising. The loss function is formulated as the mean-square error (MSE) between the denoised seismic and the actual seismic.

2) SWI, which aims at constructing optimal nonlinear mapping between well logs and local seismic patterns at the well locations. In this work, as shown in Figure 3, the SWI CNN consists of a feature engine, an encoder-decoder, and a rescaler, which are connected through operations such as up-/downsampling along the height axis, a dropout rate of 30%, concatenation along the channel axis, and averaging along the width axis. Each component contains one or more blocks of three convolutional layers followed by one dropout layer. Specifically, starting from a set of adjacent 2D seismic images, the feature engine, which is implanted from the pretrained SFE CNN and set nontrainable while training the SWI CNN, first extracts 2D features from the input images. Then, the encoder-decoder analyzes these 2D features further at multiple scales through down- and upsampling along the height axis. Next, the rescaler upsamples...
the 2D features further along the height axis to match the scale of the target well logs. Finally, the rescaled 2D features are averaged along the width axis as the predicted logs, each of which corresponds to one of the actual logs. The loss function is formulated as the MSE between the prediction logs and the actual logs.

After the SWI CNN has been trained, it is then applied to the entire seismic volume trace by trace and generates the corresponding property volume(s). Each volume corresponds to one of the properties in the target well logs, and it has the same spatial coverage as the input seismic, but it is vertically at the same scale as the well logs.

Uncertainty analysis

The uncertainty can be classified into two categories: aleatoric and epistemic (Matthies, 2007; Der Kiureghian and Ditlevsen, 2009). Given the nature of our seismic and well-log data, the uncertainty in property estimation falls into the category of epistemic uncertainty, which is often analyzed under the framework of Bayesian deep learning (Neal, 1995). In this study, starting from the approximation between enabling dropout in the prediction and estimating the posterior distribution of parameters in a neural network proposed by Gal and Ghahramani (2016), we implement the MC dropout into the uncertainty analysis of our proposed property estimation workflow. In practice, we first deliberately enable the dropout in prediction by setting the dropout rate the same as in the training stage and running it $N$ times repeatedly to generate $N$ sets of machine predictions. Then, the approximative property estimation $p$ and absolute epistemic uncertainty $v$ are estimated as the mean and variance of the $N$ predictions, respectively:

$$p(y_i = c_i/X) \approx \frac{1}{N} \sum_{t=1}^{N} p(y_{i,t} = c_i/X, W^t)$$

(1)

and

$$v^2(y_i/X) = \frac{1}{N-1} \sum_{t=1}^{N} \left[ p(y_{i,t} = c_i/X, W^t) - p(y_i = c_i/X) \right]^2,$$

(2)

where $X$ is the data within the effective receptive field in an input; $y_i$ is the prediction at $i$; $y_{i,t}$ and $W^t$ are the true value and model weights from the $t$th prediction, respectively; and $c_i$ is the true value at $i$. Furthermore, we evaluate the uncertainty band of 95% confidence as $[p - 1.96v, p + 1.96v]$. In both examples shown in the next section, we heuristically

1) set the dropout rate to 0.3 at the training and prediction stages for the SWI CNN, meaning that 30% of the data passing through a dropout layer are set to zero
2) repeat the prediction 30 times, which is tested empirically in both examples subsequently.

APPLICATIONS

In this study, we use two examples for demonstrating the performance of the proposed semisupervised learning on estimating various properties as well as the corresponding uncertainties, including

![Figure 3](image)

Figure 3. The architecture of the SWI CNN, which consists of a feature engine, an encoder-decoder, and a rescaler. Note that the feature engine is implanted from the trained SFE CAE and is locked from updating during training the SWI CNN.

![Figure 4](image)

Figure 4. (a) A 3D view of the synthetic PreSDM data set used in this study for demonstrating the performance of the proposed semisupervised learning for simultaneous RAI and porosity estimation. (b) Map view of the 36 pseudowells in the area, with 33 for training and 3 for validation (as annotated).
(1) the relative acoustic impedance (RAI) and porosity in a synthetic PreSDM data set and (2) the sonic and density in the Groningen data set.

**PreSDM data set**

Figure 4a and 4b displays a 3D view of a synthetic PreSDM data set, which is in depth and consists with 397 inlines, 327 crosslines, and 700 samples per trace, with 1 ft as the sampling interval, and the available 36 pseudwells, respectively. This data set simulates a fluviodeltaic Triassic gas field that is contained within an extensional fault-bound trap and sealed by an overlying regional postrift unconformity and is well imaged (Daly et al., 2020). The target property is the RAI and porosity. In this experiment, 33 of the 36 wells are used for training, whereas the rest three (w8, w21, and w24) are reserved for validation (Figure 4b). Figure 5 displays the ground truth RAI and porosity in three dimensions for result analysis next.

Figure 6 displays the estimated RAI and porosity, the error from the ground truth, and the corresponding AEU, respectively. As compared to the ground truth (Figure 5), the RAI and porosity estimation are of a good match, low error, and high lateral consistency, even across the two major faults. In addition, it is the relatively...
higher AEU observed near the reflection boundaries. The results in the section of validation wells w21 and w24 (Figure 7) and the comparison at the three validation wells (Figure 8) lead to the same observations. 

Furthermore, it is noted that compared to porosity, the RAI is better estimated of higher correlation and narrower uncertainty band, which may result from the stronger physical link between seismic and RAI. Meanwhile, larger errors observed near the

![Figure 7. The estimated (a) RAI and (b) porosity in the section of validation wells w21 and w24. The number beside each well represents the correlation coefficient between the machine prediction and the ground truth.](image)

![Figure 8. Result analysis of the (a) seismic, (b) RAI, and (c) porosity at three validation wells with the ground truth in red, the estimated in blue, and the corresponding uncertainty band of 95% confidence in cyan. The number on the top-left corner of each subplot represents the correlation coefficient between the machine prediction (in blue) and the ground truth (in red).](image)

![Figure 9. (a) The 3D view of the Groningen data set used in this study for demonstrating the performance of the proposed semisupervised learning for simultaneous sonic and density estimation. (b) The map view of the 375 wells available with 350 for training and 25 for validation as annotated.](image)
two major faults indicate the necessity of adding near-fault training data for further improvement.

**Groningen data set**

Our second example is the Groningen data set, which is in depth, covers an area of approximately 2700 km², and consists of 1961 inlines, 2181 crosslines, and 501 samples per trace, with 8 ft as the sampling interval (Figure 9a). The field is featured with thick Zechstein salt layers and complex fault networks particularly in the Rotliegend Formation below the salt (van Thienen-Visser et al., 2015; van Thienen-Visser and Breunese, 2015). For the experiment of property estimation in this study, 375 wells are sorted out after log quality screening, which are sampled by 1 ft (Figure 9b). The target properties in this example are sonic and density. Among them, all of the 375 wells have a gamma-ray log, and 351 wells have a density log, whereas only 148 wells have a sonic log. Although using 350 wells for training, we reserve 25 wells (annotated in Figure 9b) for validating the predicted properties.

First, we display the estimated sonic and density in three dimensions in Figure 10a and 10b, respectively, both of which demonstrate high lateral continuity and good correlation with the seismic patterns (Figure 8a). More specifically, the Zechstein salt layer is clearly delineated of low density and low sonic (high velocity), whereas the top of Rotliegend formation (RO_T) is of high density and low sonic (high velocity). Figure 11 further demonstrates how both properties laterally vary across the RO_T surface.

Next, for further evaluating the machine predictions at the validation wells, we clip the estimated sonic, density, sonic AEU, and density AEU to two sections: inline 8354 of validation well TBR-4 and inline 8813 of validation well SWO-1 in Figures 12 and 13, respectively. Figure 14 displays the sonic and density prediction as well as the corresponding uncertainty band at 10 of the 25 validation wells. From the comparisons, the observations include (1) a good match between the machine prediction and the actual logs, (2) consistent propagation from the well logs toward the rest of the seismic survey, (3) low uncertainty within the salt layer but relatively high uncertainty near the boundaries, and (4) the sonic estimation appears better than the density one that may result from the stronger physical link and correlation between the sonic and seismic. In addition, larger errors are observed near the top and bottom salt boundaries, which may result from the lack of training data that penetrate the salt layer. The results are expected to improve when some additional information, such as the salt boundaries or bodies, can be used to constrain how the machine learns in the salt layer.

After the two examples, we next investigate the impact of two factors on our proposed method, including the use of SFE and the number of training wells.
SFE

As shown in Figure 1, one major innovation of the proposed workflow is to build the SWI CNN from a pretrained SFE CNN, instead of starting from scratch. For demonstrating the associated superiorities, we use the Groningen data set and build two SWI CNNs, one from the SFE CNN and the other from scratch. Our observations are twofold. First, by inheriting the prior knowledge of the target seismic signals from the SFE CNN, training the SWI CNN starts with a lower loss and converges faster, as shown in Figure 15. Second, and more importantly as indicated by the higher correlation coefficients of machine prediction against the ground truth logs at the validation wells (Figure 16), the use of SFE CNN leads to better sonic and density prediction, even though both CNNs converge after training with the loss reduced to a similar level (Figure 15).

Dependency on the number of training wells

The other key factor to any supervised learning algorithm is the amount of training data, including the proposed CNN-based property estimation. For analyzing its dependency on the number of training wells, we use the PreSDM data set and repeat the workflow with the same hyperparameters but different amounts of training wells. Figure 17 displays the corresponding porosity prediction from 7, 16, and 33 training wells. The use of seven wells makes it possible for delineating the general porosity patterns; with the amount of training wells increasing, the machine prediction captures more subtle porosity variations, particularly in the high-porosity layers where channels exist.

DISCUSSION AND FUTURE WORK

Although the proposed workflow is capable of providing consistent property estimation from seismic and well logs, we identify the following, which may require more effort in the future.

First, as shown in Figure 3, the final layer of the current SWI CNN is in one dimension, which learns from the input 1D well logs and predicts the target properties trace by trace. Therefore, such trace-wise training and prediction require the wells to be vertical or slightly deviated, so that the mapping between seismic and wells can be indexed by a pair of \(x-y\) coordinates. When turning to the scenario of highly deviated or even horizontal wells, however, such...
indexing becomes invalid; thus, the proposed trace-wise SWI CNN is no longer applicable. One possible way to revise the SWI CNN is to conduct sample-wise training and prediction, which fits in all types of well logs, but it is of reduced computational efficiency.

Second, although the proposed workflow implements the MC dropout scheme for uncertainty analysis, the estimated is the model uncertainty, which indicates how the prediction of a model varies with any changes in the model parameters. However, such uncertainty seems to differ from what geoscientists are expecting, causing the machine estimation to be interrogated from the perspective of reliability and interpretability. For example, a relatively larger uncertainty is usually expected in a zone of poor seismic quality and/or far away from training wells than that in a zone of well-imaged seismic and nearby training wells. One potential solution is to introduce seismic quality index and distance to wells into the SWI CNN as future work.

Other future work could be more noise-oriented SFE, instead of the simple random noise attenuation in the current SFE CNN. In the presence of coherent noise such as acquisition footprints in the target seismic data set, SFE CNN can be customized for attenuating this noise, and the training could be either supervised or unsupervised. By doing so, the disturbance of noise on SWI would be reduced and the machine prediction would be further improved.

Figure 15. The comparison of loss and correlation while training the SWI CNN from scratch (in red) and the SFE CNN (in blue) in the Groningen data set. Note the lower initial loss and faster convergence by the SFE-based training.

Figure 14. The comparison of (a) seismic, (b) sonic, and (b) density at 10 of the 25 validation wells, with the ground truth in red, the machine prediction in blue, and the corresponding uncertainty band of 95% confidence in orange. The number in each subplot represents the correlation coefficient between the machine prediction (in blue) and the ground truth (in red).
CONCLUSION

For robust subsurface property inversion, we have presented a semisupervised learning workflow capable of efficiently integrating the 3D seismic data and 1D logs at a few sparsely distributed wells. As tested by the PreSDM and the Groningen data sets, the proposed method not only successfully generates highly consistent and accurate rock property estimation for the given seismic survey, but it also is capable of providing epistemic uncertainty analysis for the estimated properties, which can help us better understand the machine prediction and guide the following property modeling procedures.

ACKNOWLEDGMENTS

We would like to thank editor in chief J. Etgen, associate editor S. Chávez-Pérez, and four anonymous reviewers for their insightful comments and suggestions on improving the quality of the paper. Our thanks go to C. Daly and Nederlandse Aardolie Maatschappij for providing the PreSDM data set and the Groningen data set, respectively. The authors appreciate the Digital Subsurface Intelligence Team (especially X. Chen and H. Maniar) at Schlumberger for the fruitful discussions and suggestions on the project. The neural networks are developed using the open-source software library TensorFlow developed by the Google Brain Team.
DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES


Carron, D., 1989, High resolution acoustic impedance cross-sections from wireline and seismic data: Presented at the Annual Logging Symposium, SPWLA.


Takahashi, I., 2000, Quantifying information and uncertainty of rock property estimation from seismic data: Ph.D. dissertation, Stanford University.


Biographies and photographs of the authors are not available.