Automatic well log baseline correction for rapid characterization of potential CO_2 storage sites in the Gulf of Mexico using deep learning

Misael M. Morales, Carlos Torres-Verdín, and Michael J. Pyrcz, The University of Texas at Austin; Murray Christie and Vladimir Rabinovich, S&P Global*

Summary

We propose a deep neural network-based framework for automatic baseline correction (ABC-Net) of spontaneous potential (SP) logs. To overcome the challenge of SP log deviation and trend accumulation with depth due to salinity and temperature effects, we propose an automatic correction algorithm using deep learning. The proposed framework utilizes a deep convolutional U-Net model to estimate the baseline-corrected SP log from the raw SP log and a set of attributes. The benefit of this approach is its ability to compress and denoise the raw SP log and its attributes into a latent representation and to efficiently predict the baseline-corrected SP log without manual interpretation. We validate this approach against manually corrected SP logs, and test with unseen wells in the Gulf of Mexico. Finally, we use the trained model to estimate the volumetric concentration of shale in order to detect potential CO_2 storage sites in the Gulf of Mexico.

Introduction

The spontaneous potential (SP) log is one of the earliest borehole measurements in the energy industry and has a significant role in formation evaluation for determining lithology and permeable zones (Asquith & Krygowski, 2004). SP logs, along with Gamma Ray (GR) logs, are often the only measurements available in old hydrocarbon wells. However, temperature and salinity have significant effects on the SP measurements (McConnell, 1983, 1988). These effects result in a trend accumulation along the depth of the well, and require correcting, or shifting, the log to the baseline trend for accurate interpretation of lithology and permeable zones.

Automatic baseline correction algorithms have been widely explored, and often depend on an iterative approach for data shifting based on a precomputed attribute or filter (Gan et al., 2006). McConnell (1983, 1988) is the first to apply a baseline correction method for the SP log using a linear correction term first and later a combination of potential, environmental, and salinity correction terms. Later, Bautista-Anguiano & Torres-Verdín (2015) develop a robust mechanistic modeling framework for the interpretation of SP logs, including a physics-based correction based on reservoir topology.

Peyret et al. (2019) compare deep learning methods for automatic well log interpretation from lithology-specific logs. Shan et al. (2021) develop a deep learning method for well log generation that consistent with reservoir topology. Tang et al. (2021) introduce an ensemble machine learning framework for sweet spot detection using a suite of well logs. Simoes et al. (2022) develop a deep learning-based multi-well automatic log correction workflow for imputation and generation of missing logs. However, none of these approaches combine the concepts of automatic baseline correction specifically for SP logs and the prediction of sweet spots for potential CO_2 storage locations along a well and across a regional scale.

We propose a deep learning-based framework for automatic baseline correction of SP logs and sweet spot detection to identify potential CO_2 storage sites in the Gulf of Mexico. The deep learning method, named ABC-Net for automatic baseline correction network, exploits the latent representation of the raw SP log for compression and denoising and uses a combination of attributes to estimate the baseline-corrected SP log. The baseline-corrected SP log is used to estimate the volumetric concentration of shale along the well to detect local sweet spots for CO_2 injection. We validate the framework on a field dataset with over 300 wells in the Gulf of Mexico.

Method

We first process the data by filtering a large library of well logs in the Gulf of Mexico and select the ones with the SP log. The next step is to impute any missing values with a zero mask and zero-padding the SP log for all wells to create a tabular training set. The masked and padded values are flagged so that the deep learning method do not account for those values during the training. Normalization is applied to the well logs to aid the training process of the neural network model.

We compute eight attributes of the SP log to improve the prediction and exploit the latent structure in the data. Let $\boldsymbol{X}$ represent the raw SP log, the attributes are calculated as follows:

The derivative with respect to depth is given by the centered finite difference formula:

$$\boldsymbol{\nabla X} = \frac{x_{i+1} - x_{i-1}}{2h}, \quad (1)$$

where h is the sampling rate in depth of the well log, typically 0.25 or 0.5 ft.

The autocorrelation of $\boldsymbol{X}$ is given by:

$$\boldsymbol{R_{XX}} = \sum_{i=0}^{\|x\|-1} x_i x^*_{i-k+N-1}\,, \qquad (2)$$

where $\cdot^*$ is the complex conjugate operator, $N = \|x\|$, and $k = 0,1,\dots,2\|x\| - 2$.

The linear detrend attribute is given by:

$$\boldsymbol{L_X} = \boldsymbol{X} - m_{\boldsymbol{X}}, \qquad (3)$$

where $m_{\boldsymbol{X}}$ is the slope coefficient obtained from a least-squares fit, $\|\boldsymbol{X} - (m_{\boldsymbol{X}}\boldsymbol{X} + b)\|_2^2$.

The Fourier transform of $\boldsymbol{X}$ is given by:

$$\boldsymbol{\mathcal{F}_X} = \sum_{n=0}^{N-1} \boldsymbol{X} e^{\frac{-2\pi i}{N}kn}. \qquad (4)$$

The Hilbert transform of $\boldsymbol{X}$ is given by:

$$\boldsymbol{\mathcal{H}_X} = \mathcal{F}_{\boldsymbol{X}}^{-1}(\mathcal{F}_{\boldsymbol{X}} 2U), \qquad (5)$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform and U is the unit step function.

The symmetric infinite impulse response (IIR) filter is given by:

$$\boldsymbol{\mathcal{I}_X} = \frac{c_0}{(1 - z/x)(1 - zx)}, \qquad (6)$$

where c_0 and z are parameters of the transfer function.

The Savitzky-Golay filter is given by:

$$\boldsymbol{Y_X} = \sum_{k=-m}^{m} c_k x_{i+k}\,, \qquad (7)$$

where c_k are the coefficients of a polynomial fit of the raw data and m is half the window size.

The cubic spline coefficients is given by:

$$\boldsymbol{\mathcal{S}_X} = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, \quad (8)$$

where a_i, b_i, and c_i are coefficients that solve a system of equations that ensure continuity and smoothness.

Let $\boldsymbol{\xi}$ represent the set of a raw SP log and its corresponding set of attributes. Figure 1 shows a raw SP log and its corresponding attributes, $\boldsymbol{\xi}$, for a randomly selected well from the dataset.

Deep convolutional U-Net neural networks have been widely used for computer vision and signal processing tasks, including translation, segmentation and denoising (Chang et al., 2021). The proposed ABC-Net is a deep convolutional U-Net neural network for automatic baseline correction. ABC-Net is trained to estimate the baseline-corrected SP log, namely $\hat{\boldsymbol{\xi}}$, from the raw SP log and its set of attributes, $\boldsymbol{\xi}$. The encoder portion of ABC-Net, Enc, compresses the inputs into a latent representation, $\boldsymbol{z}$, such that $\boldsymbol{z} = Enc(\boldsymbol{\xi})$. The decoder portion, Dec, is a mirror image of the encoder and estimates the baseline-corrected

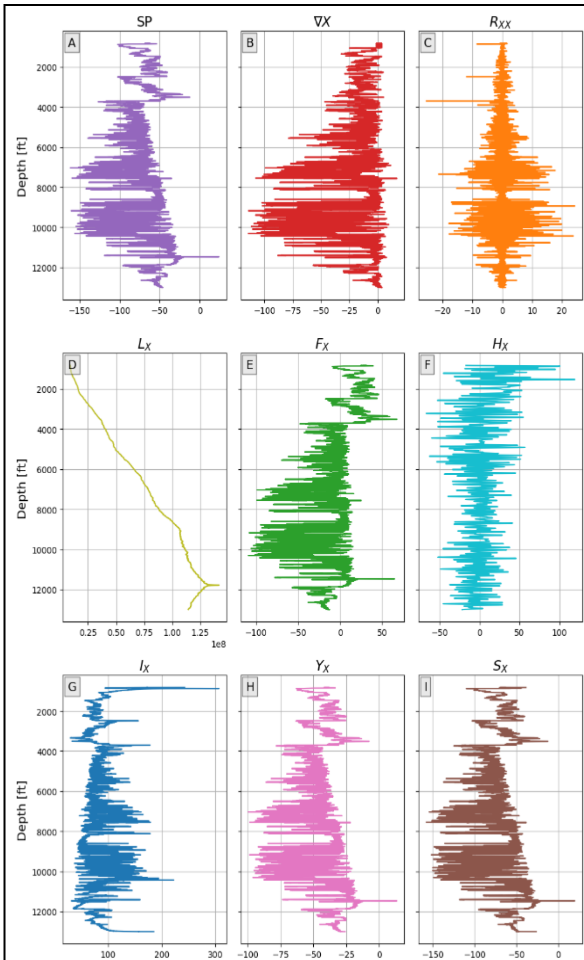

Figure 1: A randomly selected raw SP log (A) and its corresponding features: (B) derivative with respect to depth, $\nabla\boldsymbol{X}$, (C) autocorrelation, $\boldsymbol{R_{XX}}$, (D) linear detrend, $\boldsymbol{L_X}$, (E) Fourier transform, $\boldsymbol{\mathcal{F}_X}$, (F) Hilbert transform, $\boldsymbol{\mathcal{H}_X}$, (G) IIR filter, $\boldsymbol{\mathcal{I}_X}$, (H) Savitzky-Golay filer, $\boldsymbol{Y_X}$, and (I) cubic spline coefficients, $\boldsymbol{\mathcal{S}_X}$.

SP log, $\hat{\boldsymbol{\xi}}$, from the latent representation, such that $\hat{\boldsymbol{\xi}} = Dec(\boldsymbol{z}) = Dec\big(Enc(\boldsymbol{\xi})\big)$. Residual concatenations, also known as skip connections, connect the layers of the Encoder and Decoder to enhance data flow, retrain fine-grained details and spatial information, and reduce information loss. The Encoder is composed of three hidden layers, each with a 1D convolution, batch normalization, ReLU activation, Dropout, and MaxPooling. Similarly, the Decoder is composed of three hidden layers each ending with an UpSampling operator instead of MaxPooling. Figure 2 shows the model architecture and a description of the internal structure of each layer in the Encoder and Decoder portions of the model.

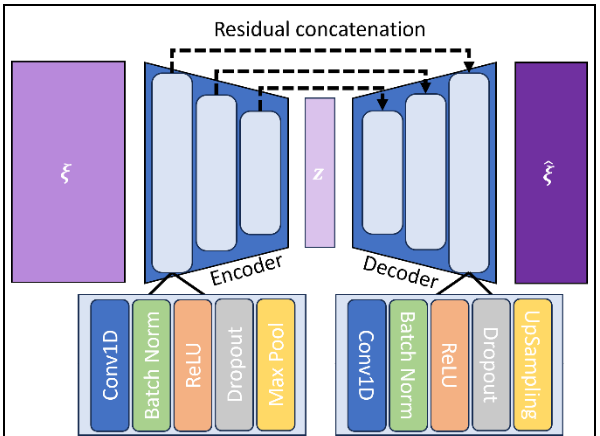

Figure 2: Architecture of the proposed ABC-Net. The raw SP log and its corresponding set of attributes, $\boldsymbol{\xi}$, are compressed into a latent representation, $\boldsymbol{z}$, through the encoder portion of the network, Enc. The Decoder, Dec, then predicts the baseline-corrected SP log, $\hat{\boldsymbol{\xi}}$, such that $\hat{\boldsymbol{\xi}} = Dec(\boldsymbol{z}) = Dec(Enc(\boldsymbol{\xi}))$. The Encoder and Decoder are mirror images of each other, and each have 3 layers, with residual concatenations used to connect the corresponding layers.

Results

A subset of 389 is wells is selected for training. Each well is processed and then partitioned into a random training and testing set with 300 and 89 wells, respectively. The ABC-Net is trained using the AdamW optimizer with learning rate 0.01 and weight decay 1e-5 and batch size of 30 for 100 epochs using an NVIDIA RTX 3080 GPU. For each batch, a random subset of 20% the batch size is used for validation. The Mean Squared Error (MSE) is used as the loss function, such that $\mathcal{L} = \|\hat{\boldsymbol{\xi}}^* - \hat{\boldsymbol{\xi}}\|_2^2$, where $\hat{\boldsymbol{\xi}}^*$ is the manually labeled baseline-corrected SP log. The model has a total of 89,681 parameters and requires approximately 1.64 hours to train.

Comparing the baseline-corrected SP log predicted from ABC-Net, $\hat{\boldsymbol{\xi}}$, against manually labeled SP logs, $\hat{\boldsymbol{\xi}}^*$, the average training and testing error is 12.9% and 13.6%, respectively. Figure 3 shows the raw and baseline-corrected SP logs for 3 randomly selected train and test wells. Once trained, each test prediction takes approximately 420 milliseconds, providing a significant advantage for rapid estimation of lithology and sweet spots compared to traditional or manual techniques.

We observe that ABC-Net is capable of estimating the baseline-corrected SP log accurately and rapidly. Due to the lossy compression of the Encoder-Decoder architecture, there are differences in terms of amplitude at several locations along the well; however, the baseline-corrected trend is accurately captured everywhere.

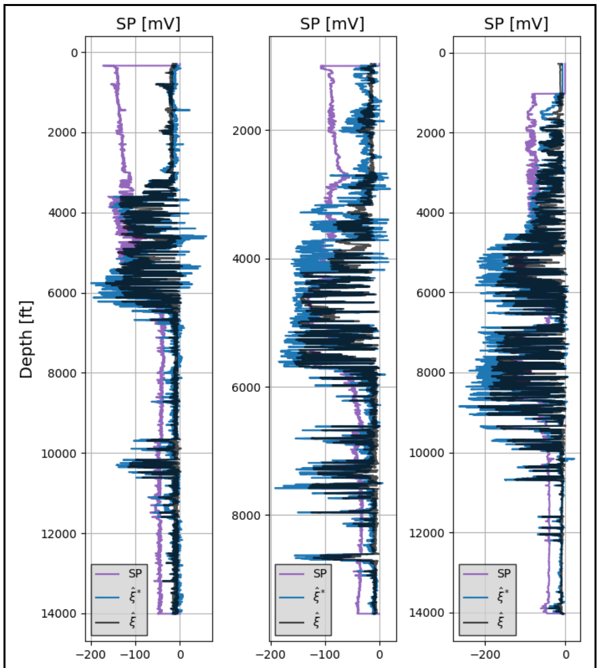

Figure 3: The raw SP log (purple), manually labeled baseline-corrected SP log (blue), and the predicted baseline-corrected SP log (black) using ABC-Net for three randomly selected wells in the Gulf of Mexico.

The predicted baseline-corrected SP logs from ABC-Net are used to calcualte the volumetric concentration of shale, C_{sh}, as a function of depth along the well as follows:

$$C_{sh}(\hat{\boldsymbol{\xi}}) = \frac{\hat{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}_{10}}{\hat{\boldsymbol{\xi}}_{90} - \hat{\boldsymbol{\xi}}_{10}}, \tag{9}$$

where $\hat{\boldsymbol{\xi}}_{90}$ and $\hat{\boldsymbol{\xi}}_{10}$ represent the 90^{th} and 10^{th} percentile of the baseline-corrected SP log.

Given that the SP log is a lithology-dependent well log, the estimation of C_{sh} from the baseline-corrected SP log provides a quick interpretation of permeable and impermeable zones as sweet spots for potential CO_2 storage. To interpret the sweet spots along the well, we compute the moving window average of the estimated C_{sh} as

$$(C_{sh} * U)_n = \sum_{m=-\infty}^{\infty} C_{sh_m} U_{n-m}, \tag{10}$$

where $*$ is the convolution operator and U is the unit step function. The window is defined to have a size $n = 400$, corresponding to 200 ft for a well log with sampling rate of 0.5 ft. A cutoff, κ, is defined to mask the sweet spots such that $(C_{sh} * U) < \kappa \Longrightarrow C_{sh}^{sweet}$. Here we define $\kappa = 0.6$. Figure 4 shows the estimated C_{sh} for 3 randomly selected wells and their corresponding sweet spots. However, it is recommended to apply further petrophysical analysis to improve the sweet spot identification for more accurate interpretation.

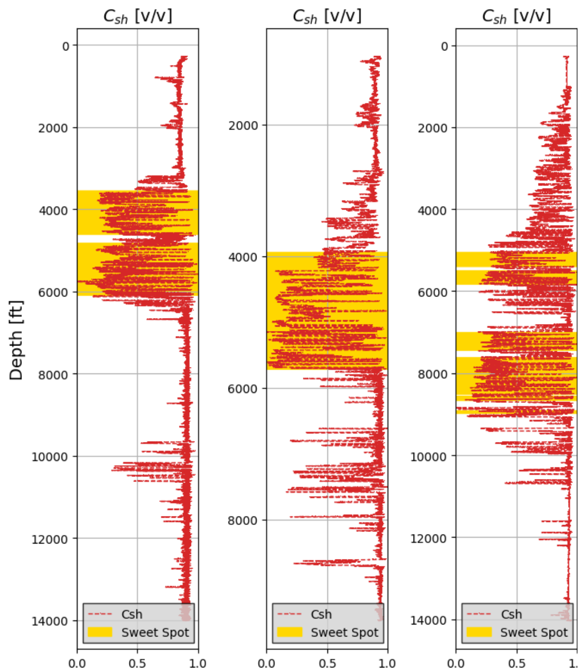

Figure 4: Estimated C_{sh} and sweet spots from the baseline-corrected SP log from ABC-Net, $\hat{\xi}$, for three randomly selected wells in the Gulf of Mexico. The yellow mask shows the estimated sweet spots for CO_2 injection along each well.

The implemented workflow with ABC-Net is capable of rapid estimation for a large number of wells with varying degrees of baseline trends, as well as robust to varying logging interval, multiple runs, and noise levels. A single training session is required for ABC-Net to obtain the optimal weights and biases. Once trained to estimate the baseline-corrected SP log, ABC-Net can be deployed for a very large number of wells to obtain predictions of permeable and impermeable zones very rapidly. Furthermore, by mapping the spatial distribution of wells along a region, (e.g., the Gulf of Mexico), the ABC-Net worfklow can be deployed to estimate regional sweet spots at a basin scale within reasonable accuracy and at very low computational costs.

ABC-Net provides rapid estimation of baseline-corrected SP logs, which are used to estimate the volumetric concentration of shale along the well as a tool to identify permeable and impermeable zones for CO_2 storage. However, ABC-Net is only trained for SP log baseline correction, and would require more data and retraining in order to estimate other well logs for lithology identification. Also, the attributes used to train ABC-Net require pre-computing and can be time consuming, so further sensitivity and feature selection for the best attributes should be performed.

Conclusions

We presented a deep learning-based workflow for automatic baseline correction of well logs, namely ABC-Net. This method allows for efficient petrophysical evaluation of well logs without the need for manual correction or interpretation which would be extremely time-consuming and subjective.

Using SP logs as a lithology-dependent measurement, we can estimate permeable and impermeable zones along a well for possible CO_2 storage zones. We validate our method with data from 389 wells from the Gulf of Mexico and obtain predictions for each well within 420 milliseconds at only 13.6% error on average without the need for user interpretation or manual corrections.

Ultimately, the ABC-Net workflow can be implemented at a regional or basin scale for a sufficiently large number of well logs to estimate the distribution in space and depth of possible sweet spots for CO_2 injection based on lithology from the baseline-corrected SP logs.

Acknowledgements

This work is supported by the Digital Reservoir Characterization (DiReCT) consortium at the University of Texas at Austin.

Automatic well log baseline correction for rapid characterization of potential CO_2 storage sites

References

Asquith, G., and D. Krygowski, 2004, Chapter 2: Spontaneous Potential, *in* AAPG Methods in Exploration, No. 16: AAPG Special Volumes, p. 21–30.

Bautista-Anguiano, J., and C. Torres-Verdín, 2015, MECHANISTIC DESCRIPTION, SIMULATION, AND INTERPRETATION OF SPONTANEOUS POTENTIAL LOGS.

Chang, J., J. Li, Y. Kang, W. Lv, D. Feng, and T. Xu, 2021, SegLog: Geophysical logging segmentation network for lithofacies identification: IEEE Transactions on Industrial Informatics, v. 18, no. 9, p. 6089–6099.

Gan, F., G. Ruan, and J. Mo, 2006, Baseline correction by improved iterative polynomial fitting with automatic threshold: Chemometrics and Intelligent Laboratory Systems, v. 82, no. 1, p. 59–65, doi:https://doi.org/10.1016/j.chemolab.2005.08.009.

McConnell, C. L., 1988, A general correction for spontaneous potential well logs in fresh water: Journal of Hydrology, v. 101, no. 1–4, p. 1–13, doi:10.1016/0022-1694(88)90024-8.

McConnell, C. L., 1983, Spontaneous potential corrections for groundwater salinity calculations — Carter County, Oklahoma, U.S.A.: Journal of Hydrology, v. 65, no. 4, p. 363–372, doi:https://doi.org/10.1016/0022-1694(83)90087-2.

Peyret, A. P., J. Ambía, C. Torres-Verdín, and J. Strobel, 2019, Automatic interpretation of well logs with lithology-specific deep-learning methods, *in* SPWLA 60th Annual Logging Symposium 2019: Society of Petrophysicists and Well-Log Analysts (SPWLA), doi:10.30632/T60ALS-2019_SSSS.

Shan, L., Y. Liu, M. Tang, M. Yang, and X. Bai, 2021, CNN-BiLSTM hybrid neural networks with attention mechanism for well log prediction: Journal of Petroleum Science and Engineering, v. 205, doi:10.1016/j.petrol.2021.108838.

Simoes, V., H. Maniar, A. Abubakar, and T. Zhao, 2022, Deep Learning for Multiwell Automatic Log Correction, *in* Petrophysics: Society of Petroleum Engineers (SPE), p. 724–747, doi:10.30632/PJV63N6-2022a10.

Tang, J., B. Fan, L. Xiao, S. Tian, F. Zhang, L. Zhang, and D. Weitz, 2021, A new ensemble machine-learning framework for searching sweet spots in shale reservoirs, *in* SPE Journal: Society of Petroleum Engineers (SPE), p. 482–497, doi:10.2118/204224-PA.