

Sonic Well-Log Imputation Through Machine-Learning-Based Uncertainty Models

Eduardo Maldonado-Cruz¹, John T. Foster^{1,2}, and Michael J. Pyrcz^{1,3}

ABSTRACT

Sonic well logs provide critical information to calibrate seismic data and support geomechanical characterization. Advanced subsurface data analytics and machine learning enable new methods and workflows for property estimation, regression, and classification for geoscience and subsurface engineering applications. However, current applications for imputation of well-log values rely only on model accuracy and low error predictions.

Traditional model validation techniques are not enough to validate models and account for the substantial uncertainty in the subsurface. Well-log imputation estimates and their associated uncertainty models are essential to the field development planning and decision-making workflows, such as reservoir modeling, volumetric

resource assessment, predrill prediction with uncertainty, remaining resource mapping, and production allocation.

When performing subsurface feature imputation with machine learning, we must expand our machine-learning model training and complexity tuning workflows to check the entire uncertainty model to ensure uncertainty distributions are precise and accurate. We propose a workflow that integrates the goodness metric to calculate accurate and precise uncertainty models of sonic well-log predictions based on ensembles of the machine-learning estimates. Our workflow combines model evaluation and visualization of the estimates and the uncertainty model with respect to measured depth. Our proposed method provides intuitive diagnostics and metrics to evaluate estimation accuracy and uncertainty model goodness.

INTRODUCTION

Well logging offers invaluable information to characterize and evaluate reservoirs and aquifers. When integrated, well-log information is used to construct reservoir models for production forecast (Eskandari et al., 2004; Alexeyev et al., 2017; Esmaeli et al., 2022; Dong et al., 2022; Ye et al., 2022). Sonic-logging tools transmit compressional and shear waves through the formation to obtain information about the matrix and the fluid content (Wyllie et al., 1958). From the responses in the tool receiver, we compute the compressional slowness (DTC) and shear slowness (DTS) logs. The difference between these measurements contains information about the petrophysical properties of the rock and fluid content, especially gas. (Souder, 2002; Zhang et al., 2009; Misra et al., 2021; Jamali et al., 2022; Li et al., 2022).

Frequently, sonic well-log information is missing due to economic or mechanical reasons. Missing well-log information is detrimental to reservoir characterization. Relogging wells come with high costs, making well-log imputation an attractive alternative. Multiple solutions exist to impute missing well-log data based on physical models (Asquith et al., 2004; Rolon et al., 2005; Du et al., 2008; Bateman, 2012). However, these solutions often rely on assumptions and simplifications that do not necessarily capture the elastic properties of the rock available in the multivariate well-log information.

Using machine-learning models is an alternative to estimating synthetic well logs with existing collocated conventional log data (Eskandari et al., 2004; Du et al., 2008; Tan et al., 2015). Machine-learning models are widely used in the petroleum industry, such as drilling (AlBahrani and Morita, 2021; Abdelaal et al., 2021; Olukoga and Feng,

Manuscript received by the Editor July 30, 2022; revised manuscript received October 11, 2022; manuscript accepted October 11, 2022.

¹Hildebrand Department of Petroleum and Geosystem Engineering, Cockrell School of Engineering, The University of Texas at Austin, emaldonadocruz@utexas.edu; mpyrcz@austin.utexas.edu; john.foster@utexas.edu

²Department of Aerospace Engineering and Engineering Mechanics, Cockrell School of Engineering, The University of Texas at Austin, john.foster@utexas.edu

³Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, mpyrcz@austin.utexas.edu

2021), production (Alarifi and Miskimins, 2021; Liu et al., 2021; Vanegas et al., 2021), petrophysics (Deng et al., 2019; Kazak et al., 2021; Neff et al., 2021), and well-log imputation (Rolon et al., 2005; Eshkalak et al., 2013; Maleki et al., 2014; Long et al., 2016; Tariq et al., 2016; Li and Misra, 2017; Kim et al., 2020; Tatsipie and Sheng, 2021).

Current machine-learning models for well-log imputation rely on model accuracy even when uncertainty is significant due to spatial feature heterogeneity, limited data veracity, and sparsity. Therefore, the prediction of a single accurate estimate must be replaced with an accurate and precise uncertainty distribution.

Bayesian neural networks can capture real-world uncertainties, but their application results in high computational costs (Blundell et al., 2015; Kwon et al., 2020; Jia et al., 2021). Uncertainty estimation within the standard Bayesian ensemble-based framework has been extensively studied for neural network models (Hinton et al., 2012; Srivastava et al., 2014; Gal and Ghahramani, 2016; Malinin et al., 2021); tree-based models (Chipman et al., 2010; Lakshminarayanan et al., 2017; Linero, 2017), and gradient boosting based on decision trees (GBDT) (Duan et al., 2020; Malinin et al., 2021; Zhang et al., 2021).

The gradient boosting algorithm is a robust method for small data sets and multiple features (Friedman, 2001). The gradient boosting algorithm creates multiple weak models sequentially to construct a robust model M . These weak learners are shallow decision trees that recursively partition the feature space into mutually exclusive regions, and the sum of shallow decision trees is the final GBDT model.

The gradient boosting decision tree, $M(x)$, is updated at each iteration as:

$$M^{(t)}(x) = M^{(t-1)}(x) + \epsilon h^{(t)}(x) \quad (1)$$

- M , updated model
- $M^{(t-1)}$, model constructed at the previous iteration
- $h^{(t)}$, shallow decision tree at iteration t
- ϵ , learning rate

Ensemble approaches separate the total uncertainty into data and knowledge uncertainty (Gal and Ghahramani, 2016; Depeweg et al., 2017; Malinin et al., 2021). Data uncertainty arises due to noise in the data, while knowledge uncertainty is caused by sparse sampling (Malinin et al., 2021). Ensembles are often more accurate than individual estimators (Dietterich, 2000; Anifowose et al., 2017).

By constructing an ensemble of accurate estimators, the algorithm can average their predictions and reduce the risk of choosing the wrong estimate.

Maldonado-Cruz and Pyrcz (2021) applied the goodness metric and accuracy plots (Deutsch, 1997) as efficient tools to diagnose and tune hyperparameters for deep-learning-based uncertainty models. Consider an indicator function $\xi(\mathbf{u}_i; p)$ for every testing data point \mathbf{u}_i , $i = 1, \dots, n$ as;

$$\xi(\mathbf{u}_i; p) = \begin{cases} 1, & \text{if } F_y(\mathbf{u}_i; y(\mathbf{u}_i)) \in (p_{low}, p_{upp}], \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where $F_y(\mathbf{u}_i; y(\mathbf{u}_i))$ are the cumulative probabilities associated with the withheld testing values $y(\mathbf{u}_i)$, and p_{low} and p_{upp} are symmetric p -probability intervals. The proportion of testing data within the probability intervals is calculated as the average of $\xi(\mathbf{u}_i; p)$ over all data locations. The uncertainty model is accurate and precise when the average of the indicator function is equal to the probability of the interval, $\xi(p) = p, \forall p$. Following this concept, Fig. 1 shows an example of accuracy plots as a diagnostic to evaluate the predictions of uncertainty models. Both models have a high correlation coefficient, yet only model Fig. 1a captures the total uncertainty due to data and knowledge uncertainty.

While current methods for well-log imputation achieve high correlation coefficients during testing (Al-Bulushi et al., 2007; Elkatatny et al., 2016; Akinnikawe et al., 2018; Liang et al., 2021; Yu et al., 2021), there are no diagnostics to evaluate the quality of the resulting uncertainty model. In addition, these models are limited to deterministic predictions without uncertainty, even when high uncertainty exists in the subsurface (Maldonado-Cruz and Pyrcz, 2022).

We propose a workflow that integrates the modified goodness metric for tuning deep-learning-based uncertainty models (Maldonado-Cruz and Pyrcz, 2021) to impute accurate and precise sonic well-log ensemble-based nonparametric uncertainty models. We extend the use of the uncertainty model goodness metric for GBDT applications. The use of our workflow results in the generation of accurate and precise GBDT uncertainty models of sonic well-log predictions, \hat{y} , based on ensembles of the estimate represented by nonparametric cumulative distribution functions, $F_y(\mathbf{u}_i)$, over testing data, $y(\mathbf{u}_i)$. Our workflow integrates model evaluation and visualization of the estimates and the uncertainty model with respect to measured depth. This results in intuitive diagnostics and metrics to evaluate accuracy and uncertainty model accuracy and precision.

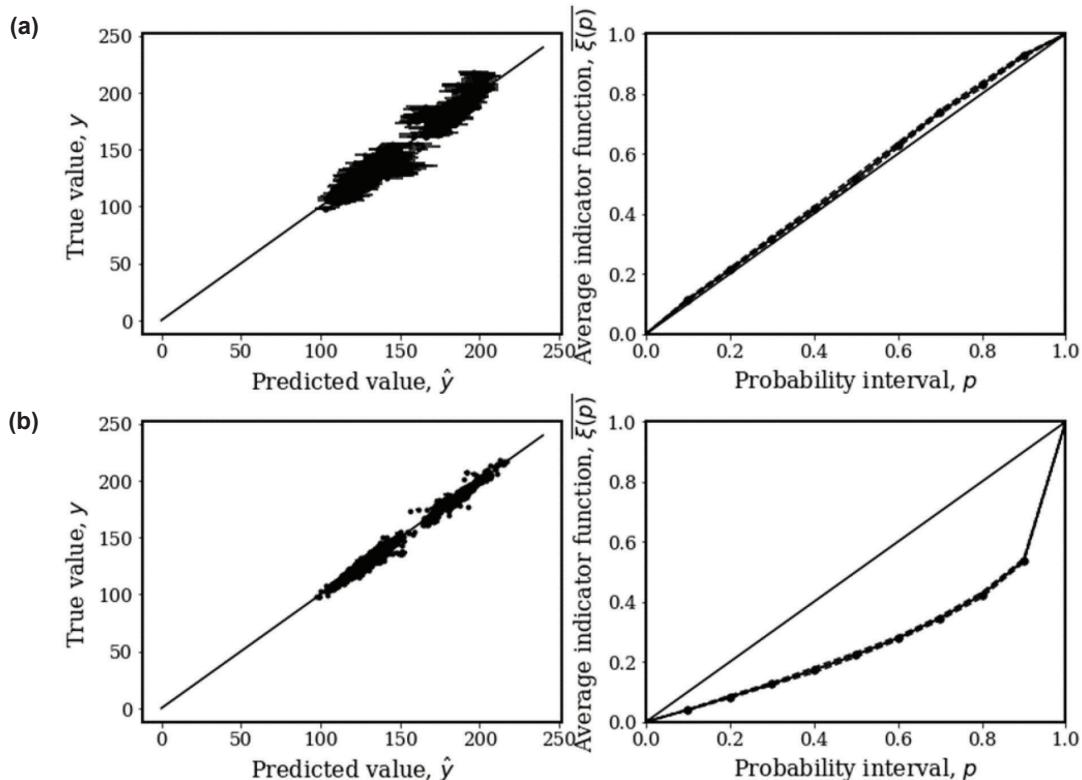


Fig. 1—Application of accuracy plots to diagnose uncertainty in machine-learning models. The left plots show the crossplot of the true values vs. those predicted by the ensemble model, and the plots on the right show the accuracy plot. (a) shows an accurate and precise probabilistic model, and (b) shows a probabilistic model that is inaccurate and imprecise.

The workflow section describes the algorithm for the imputation of DTS and DTC logs with uncertainty. We present a series of steps to calculate robust GBDT-based uncertainty models essential for subsurface imputation of sonic well logs. The results section shows the application of the proposed workflow for a regression problem with the prediction of missing sonic logs for two wells. In addition, we include the uncertainty of the predictions and the accuracy plot as diagnostics for the uncertainty model performance.

METHODOLOGY

We use the publicly available Volve Field data set (Equinor, 2018). Volve Field is located in the central part of the North Sea and has an areal extension of 6 km². The average permeability is 1,000 md, and the average porosity is 0.21. We select three wells with DTS and DTC logs. The data set contains conventional well logs (CALI, GR, NPHI, PEF, RT, RHOB) containing around 40,000 DTS and DTC data points. In Fig. 2, we plot conventional well logs in Tracks 1 to 6 and the DTC and DTS logs in Tracks 7 and 8.

We propose the following algorithm for the imputation and uncertainty quantification of DTS and DTC logs with gradient boosting decision trees.

Input

Conventional well logs selected from mutual information at selected depths and m hyperparameter combinations.

Output

Uncertainty model for DTS and DTC well-log predictions:

1. **Quality control:** visualization, petrophysical analysis, signal processing, and normalization to clean and process raw training data
2. **Feature selection:** using mutual information at selected depths to measure the general pairwise mutual dependence between features
3. Sample m hyperparameter combinations for the models to test
4. **Model training and hyperparameter tuning:**
while $i < m$ **do**

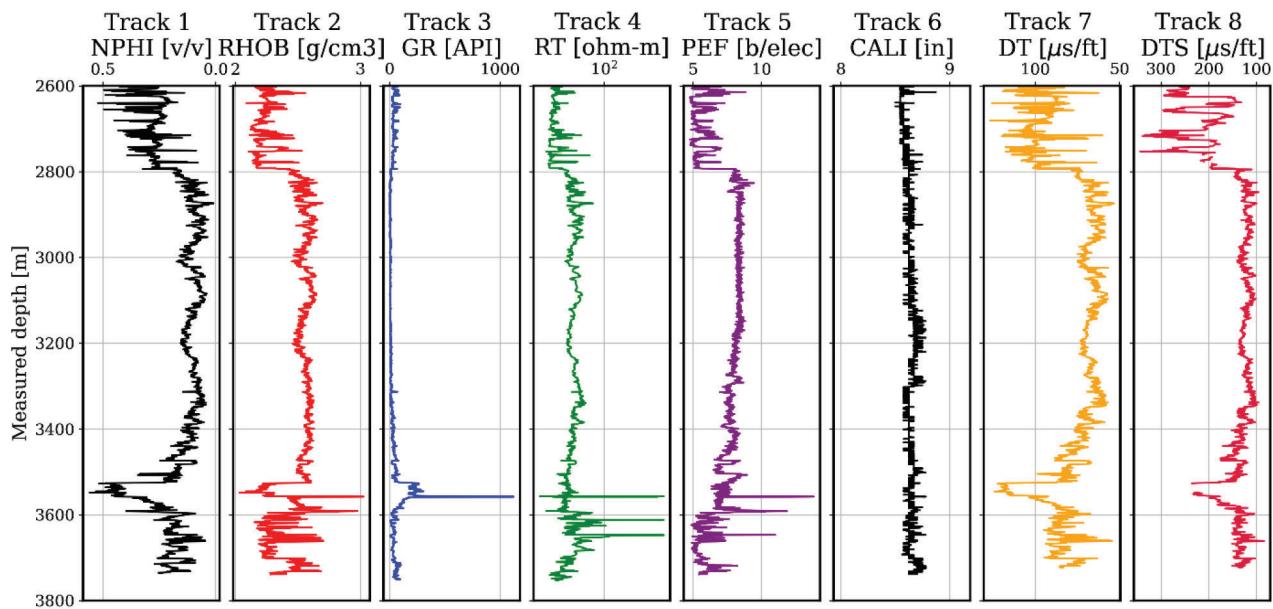


Fig. 2—Example data set used for training. Track 1 shows neutron porosity log (fraction), NPHI. Track 2 shows a bulk density log (gr/cc), RHOB. Track 3 shows the gamma ray log (API), GR. Track 4 shows a true resistivity log ($\Omega \cdot \text{m}$), RT. Track 5 shows a photoelectric factor log (b/elec), PEF. Track 6 shows a caliper log (in.), CALI. Tracks 7 and 8 show the sonic logs ($\mu\text{s}/\text{ft}$), DTC and DTS, respectively.

- a. Train models with the selected combination of hyperparameters
- b. For the trained ensemble model, calculate prediction realizations over the testing data and estimate the cumulative distribution function
- c. Evaluate the average indicator function and estimate the accuracy and precision of the uncertainty model
- d. Use Eq. 4 to evaluate the goodness of the uncertainty model

Return model hyperparameters.

Data Preprocessing

We first preprocess the raw data from the well logs to increase the quality of our predictions. Preprocessing includes the integration of domain expertise and all data sources for depth selection, signal processing, and outlier removal. For signal processing, we use the Savitzky-Golay filter (Savitzky and Golay, 1964) to smooth the measurements while maintaining the shape and magnitude. For outlier removal, we use a one-class support vector machine (SVM) (Schölkopf et al., 1999). One-class SVM identifies data in the limits to be anomalies.

Zones with extreme rugosity and washout, excessive density correction (DRHO), and sonic semblance intervals were removed. All well logs used as predictor features to develop the data-driven model are scaled using a min-max normalization following Eq. 3:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}, \quad (3)$$

where X_{max} is the maximum value of the predictor feature, X_{min} is the minimum value of the predictor feature, and X' is the normalized predictor feature value. We perform min-max normalization after outlier removal to remove the scale effect over feature importance (Singh and Singh, 2020).

Feature Selection

We perform feature selection over defined depths using mutual information (Peng et al., 2005; Peng and Fan, 2017) to measure the general pairwise mutual dependence between features and select the conventional well logs (predictor features) with the highest mutual information with the sonic DTS and DTC logs (response features). We favor mutual information over correlation analysis for a more flexible measure of information sharing between features, including patterns such as nonlinearity and heteroscedasticity that would negatively impact correlation.

Model Training and Hyperparameter Tuning

We separate two wells for training and one for testing from the data set. Our workflow considers only a train-test approach (Beleites and Salzer, 2008; Beleites et al., 2013; Xu and Goodacre, 2018). The next step is to train a GBDT and tune the hyperparameters using the loss function in Eq. 4 to evaluate the combined prediction accuracy and uncertainty model accuracy and precision. To select which hyperparameter to tune, we first quantify hyperparameter importance (Hutter et al., 2014). Based on this study, we choose the learning rate and the number of estimators as they are critical hyperparameters for building GBDT ensembles (Malinin et al., 2021; Yin and Li, 2022) while maintaining the default values for the rest of the hyperparameters. Common regression problems are evaluated using an L1 or L2 norm and aim to minimize error over withheld testing data. Traditional model validation techniques such as cross validation and quantification of model accuracy of model predictions and withheld testing data are not enough to account for the substantial uncertainty in the subsurface.

We use the loss function proposed by Maldonado-Cruz and Pyrcz (2021) in Eq. 4. The loss function includes the mean absolute error (MAE) and a measure of goodness, where $y(\mathbf{u}_i)$ are the training values and are the estimates over n testing data values, $\bar{\xi}(p)$ is an average indicator function, and $a(p)$ is a measure of accuracy. We use Eq. 4 as the objective function to tune model hyperparameters. This equation considers both prediction accuracy and uncertainty model accuracy and precision and is bounded between 0 and 1, where 0 is a very inaccurate and imprecise model. As we approach 1, we get a model that is both accurate and precise, so we aim to maximize the objective function,

$$L(\hat{y}) = 1 - \frac{1}{2} \left[\frac{1}{n} \sum_{i=1}^n |y(\mathbf{u}_i) - \hat{y}(\mathbf{u}_i)| + \left[\int_0^1 [3a(p) - 2] [\bar{\xi}(p) - p] dp \right] \right]. \quad (4)$$

RESULTS AND DISCUSSION

We demonstrate our workflow for predicting DTS and DTC logs using well-log information from the Volve Field (Equinor, 2018). We use the information from three wells in interval depths of 2,600 to 4,800 m to train and test the uncertainty model following the algorithm listed in the “Methodology” section. Next, we use the resulting model to predict the DTS and DTC logs at two wells that do not have sonic logs. The uncertainty model $F_y(\mathbf{u}_i)$, is required to predict DTS and DTC response features at every measured depth, \mathbf{u}_i , using the information from the other available well logs. We want to capture the uncertainty of DTS and DTC values at every measured depth for our regression problem. First, we preprocess our data set to improve model training.

Figure 3a shows the matrix scatterplot of the data set before data preprocessing, and Fig. 3b shows the matrix scatterplot after data preprocessing. Data preprocessing includes min-max normalization, outlier removal, and signal processing to remove noise.

Next, we calculate mutual information at selected measured depths for quantitative feature selection. Figure 4 shows the results of using mutual information (a) to impute DTS and (b) to impute DTC log values at a measured depth between 3,287 to 3,400 m over training data. We select the first five logs from mutual information.

The training and testing set for the DTC and DTS log prediction split is 70% and 30%, respectively. We conduct a grid search with cross validation using fivefolds to demonstrate the inclusion of Eq. 4 for GBDT models. We search a space between 200 and 2,000 trees and learning rates from 0.001 to 0.08. In total, we have 3,900 trials to maximize Eq. 4. Our calculations demonstrate the absence of overfitting by obtaining similar performances between training and testing data sets (Beleites and Salzer, 2008; Beleites et al., 2013; Xu and Goodacre, 2018). Table 1 is a summary of the models selected by hyperparameter grid search. In the table, we include two models that minimize MSE and two that aim to maximize the uncertainty model goodness metric from Eq. 4. The average standard deviation for all fivefolds using MSE as prediction error for DTS prediction is $0.6741 \frac{\mu\text{s}}{\text{ft}}^2$ for training and $0.5741 \frac{\mu\text{s}}{\text{ft}}^2$ for testing.

We try hyperparameter tuning using only mean squared error (MSE) as prediction accuracy and show the grid search results in Fig. 5. While common approaches in well-log imputation aim to reduce MSE, we show that minimizing MSE results in inaccurate and imprecise uncertainty models with low model goodness. We incorporate the accuracy plot for two trials as a diagnostic plot in Fig. 5. We select two trials that correspond to low MSE values without overfit, as shown in Table 1. The performance between training and testing data is similar. These plots show that MSE as prediction accuracy results in low uncertainty model goodness.

To illustrate the performance of Model 3277, we construct Fig. 6, where we show predictions of DTS and DTC from this model with respect to measured depth in the interval from 3,475 to 3,900 m. The predictions from this model are inaccurate and imprecise, as observed from the accuracy plot. The average MSE of Model 3277 is $3.29 \mu\text{s}/\text{ft}$ for DTC predictions and $8.82 \mu\text{s}/\text{ft}$ for DTS predictions. While this model obtains a low prediction error, the accuracy plot shows that this model does not capture the total uncertainty.

Sonic Well-Log Imputation Through Machine-Learning-Based Uncertainty Models

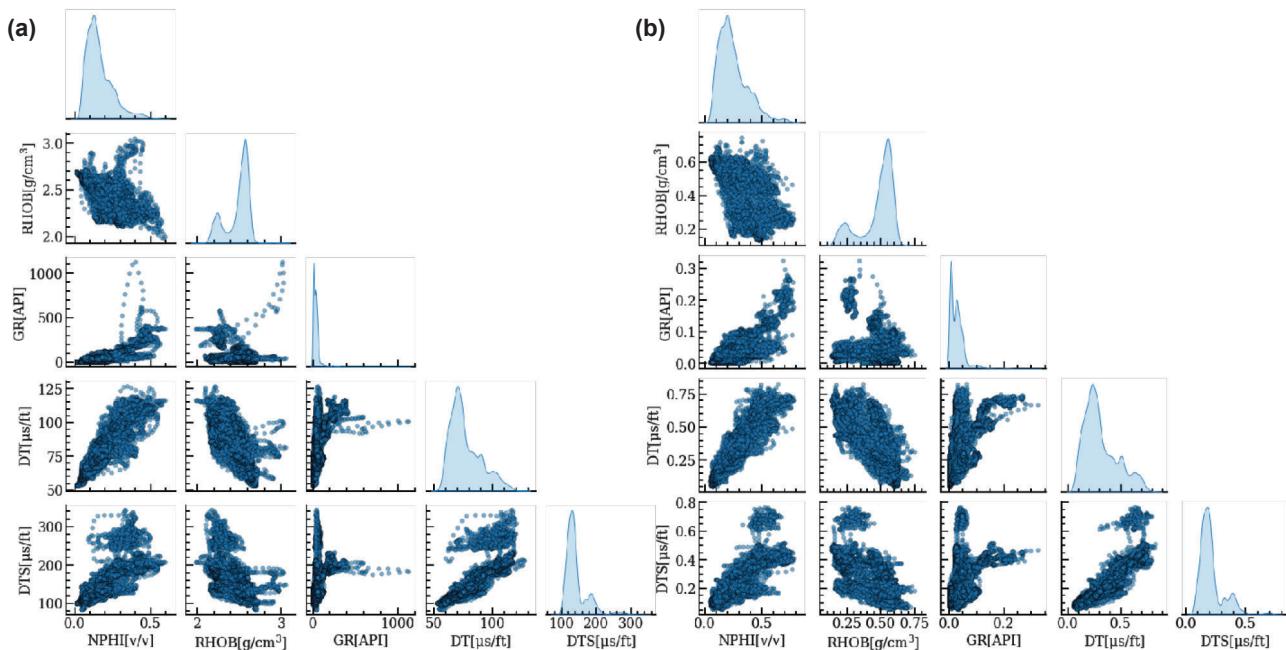


Fig. 3—Data set matrix scatterplots for the training data; the diagonal shows the kernel density estimator. We show (a) a matrix scatterplot before data preprocessing and (b) a matrix scatterplot after data preprocessing.

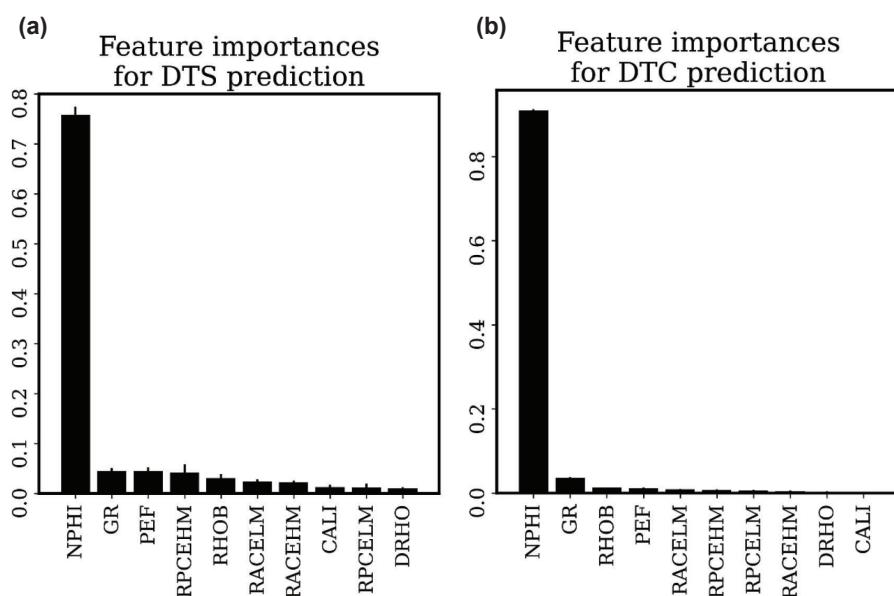


Fig. 4—Mutual information results for measured depths between 3,287 to 3,400 m over training data. (a) Mutual information to predict DTC log values and (b) mutual information to predict DTC log values. We select the predictor features with the highest mutual information with the response feature.

Table 1—Summary of Selected Models From the Grid Search*

| Model Number | Tree Number | Learning Rate | Accuracy | Precision | Goodness | DTS MSE Testing [$\mu\text{s}^2/\text{ft}$] | DTS MSE Training [$\mu\text{s}^2/\text{ft}$] | Normalized Log Likelihood |
|--------------|-------------|---------------|----------|-----------|----------|---|--|---------------------------|
| Trial 3543 | 1940 | 0.07 | 0.03 | 0.98 | 0.69 | 22.43 | 21.64 | 0.72 |
| Trial 3277 | 1880 | 0.064 | 0.03 | 0.99 | 0.65 | 23.78 | 22.24 | 0.68 |
| Trial 225 | 260 | 0.036 | 0.45 | 0.98 | 0.99 | 66.31 | 60.64 | 0.33 |
| Trial 301 | 280 | 0.028 | 0.5 | 0.99 | 1 | 58.49 | 57.64 | 0.37 |

*Models with low MSE result in low model goodness.

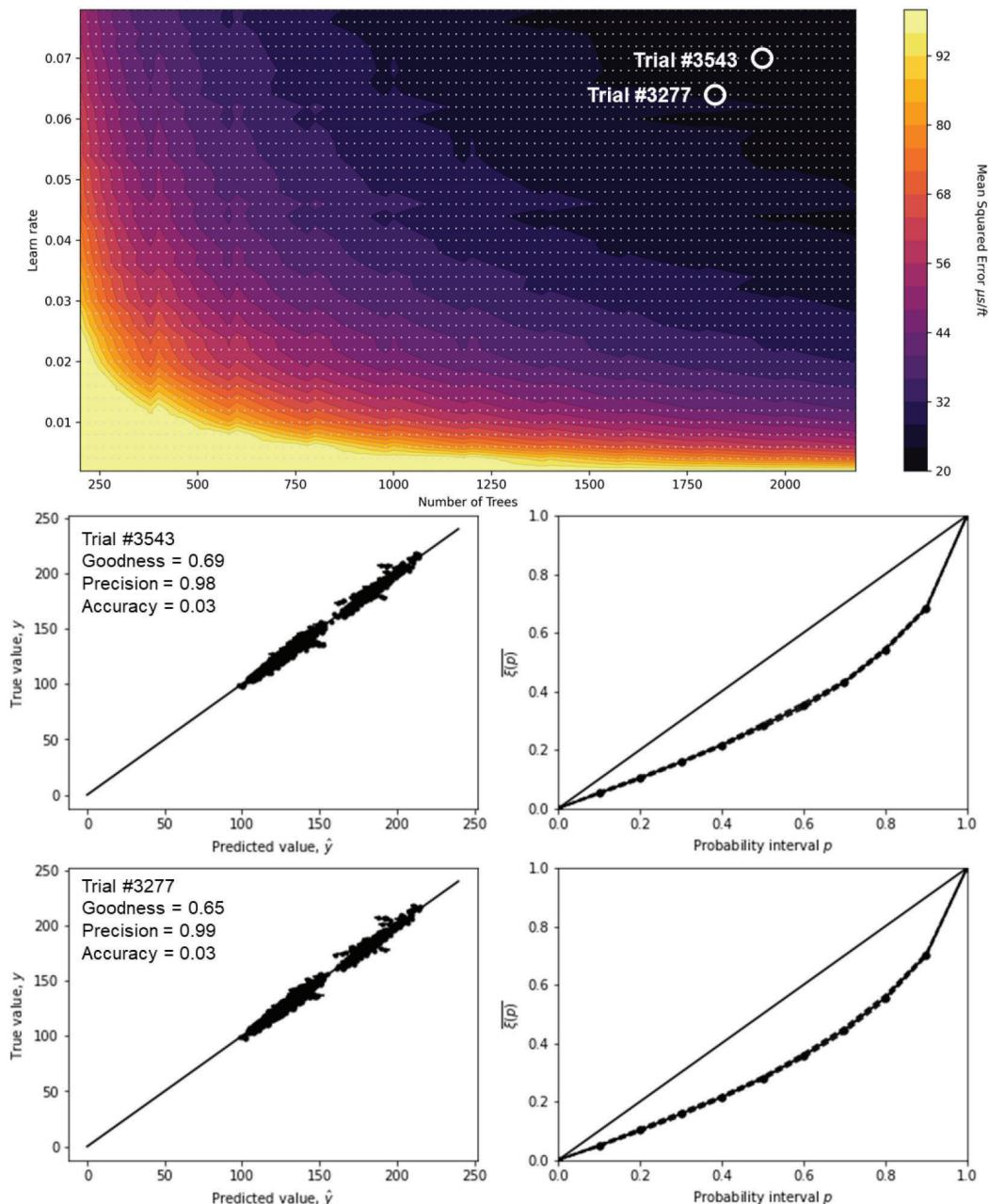


Fig. 5—We color the trial search space using MSE as a metric for prediction accuracy over testing data. We select two trials from the grid search and display the crossplot of the true values vs. those predicted by the ensemble model and their respective accuracy plot.

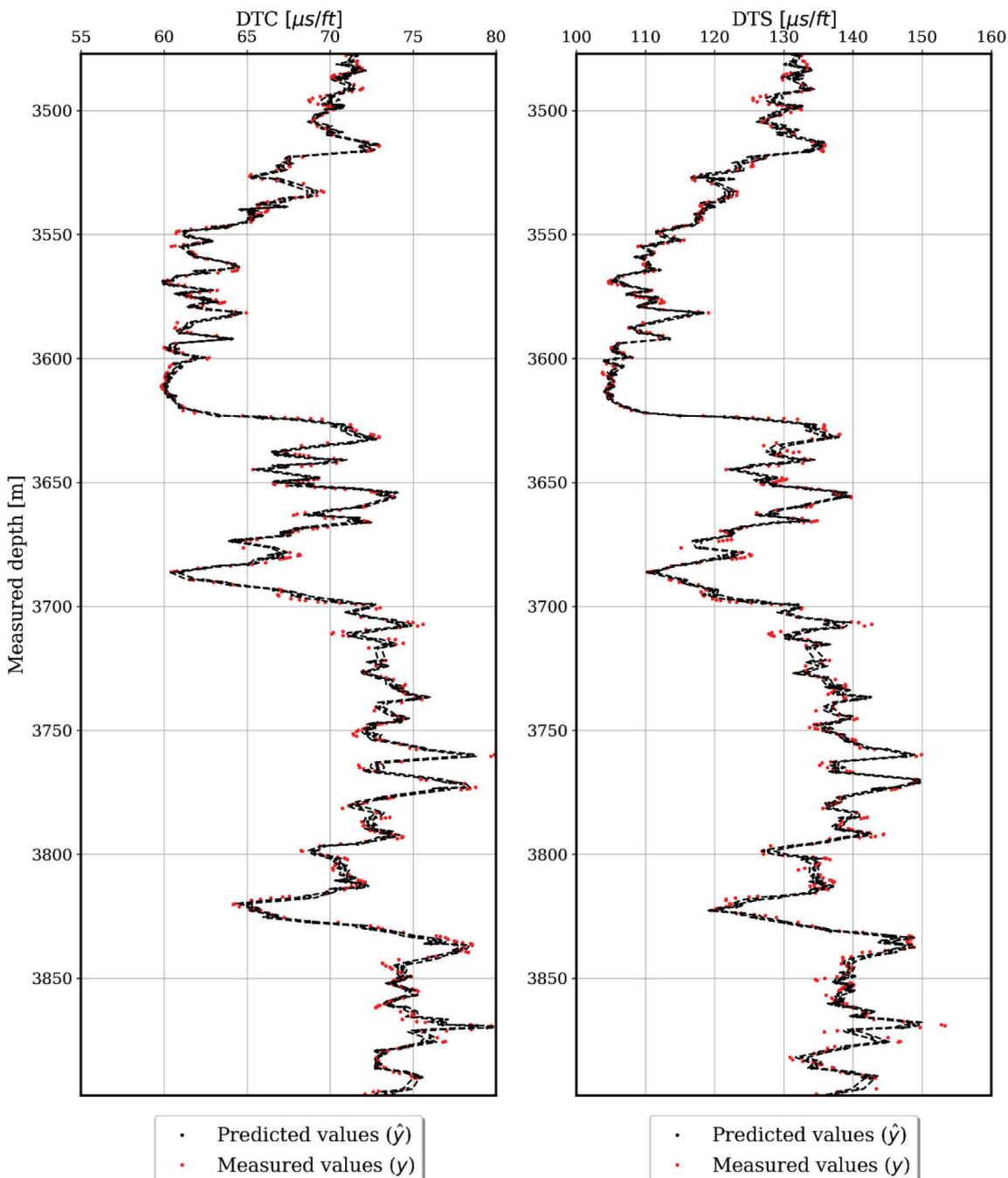


Fig. 6—DTC and DTS predictions over testing data using Model 3543 in the interval from 3,475 to 3,900 m.

The maximum likelihood estimate (MLE) is often used for geomechanical property prediction (Ledesma et al., 1996; Hartigan, 1998), where the best estimation of parameters is found by maximizing the likelihood, L , of a hypothesis, θ , given an error measurement between measurements, y ,

and model predictions, \hat{y} . Using the likelihood function, we color the hyperparameter space in Fig. 7. Maximizing the likelihood function results in low MSE models. However, the accuracy plot shows that these models are inaccurate and imprecise, as indicated by the accuracy plot.

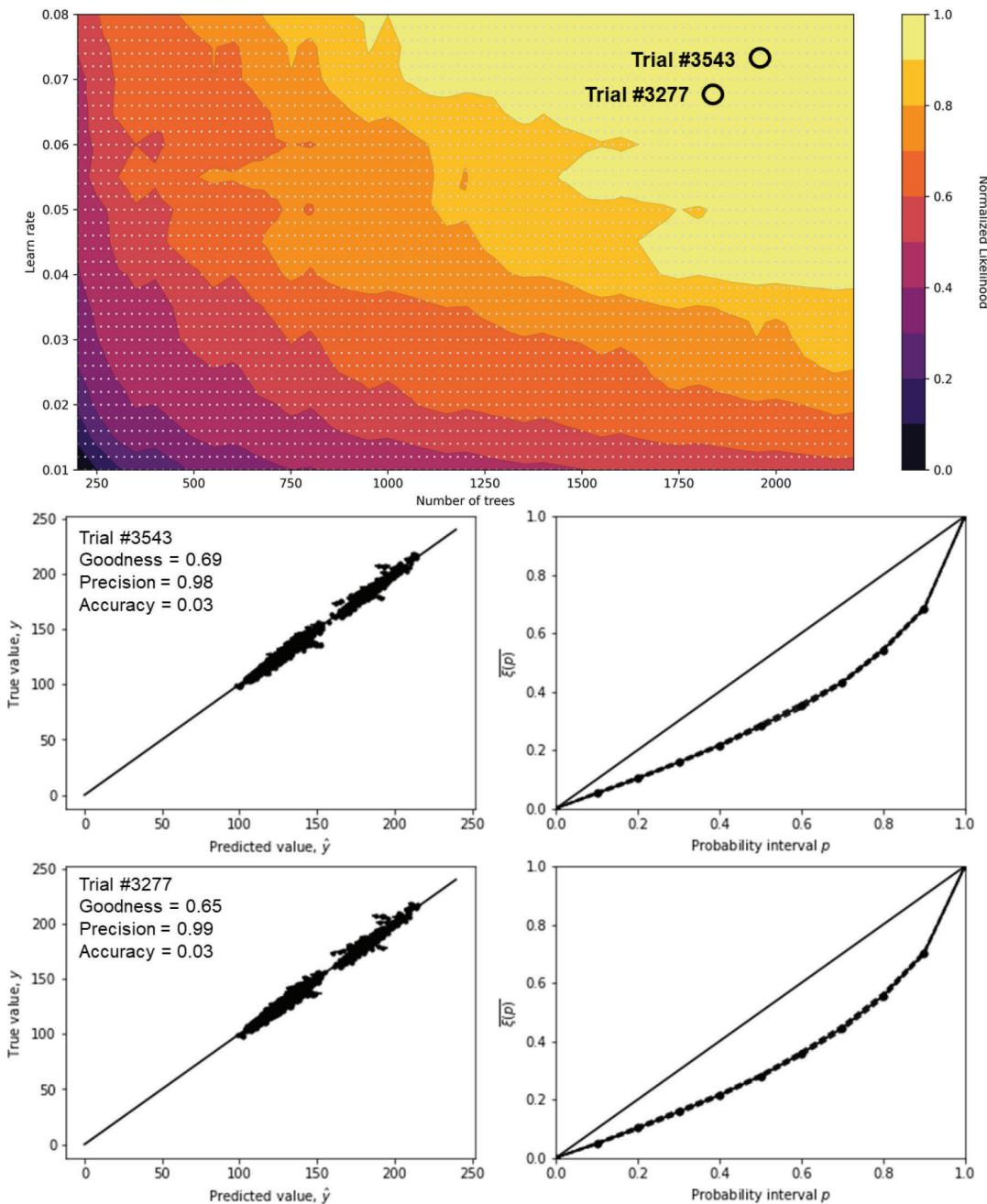


Fig. 7—Hyperparameter space colored by the likelihood function over testing data. We select Trials 3543 and 3277 from the grid search and display the crossplot of the true values vs. those predicted by the ensemble model and their respective accuracy plot.

Figure 8 shows the hyperparameter search space colored by the uncertainty model goodness in Eq. 4. We select two trials that maximize Eq. 4 from the search space and plot the true values vs. predicted by the ensemble and the corresponding accuracy plots as diagnostic. We choose Trial 301 to construct Fig. 9. This figure shows predictions

of the DTS and DTC logs in the interval of 3,475 to 3,900 m. The model from Trial 301 maximizes Eq. 4, indicating that our ensemble captures the total uncertainty of the training data correctly. The model from Trial 301 has an average standard deviation of $6.82 \frac{\mu s}{ft}$ for DTS predictions and $3.18 \frac{\mu s}{ft}$ for DTC predictions.

Sonic Well-Log Imputation Through Machine-Learning-Based Uncertainty Models

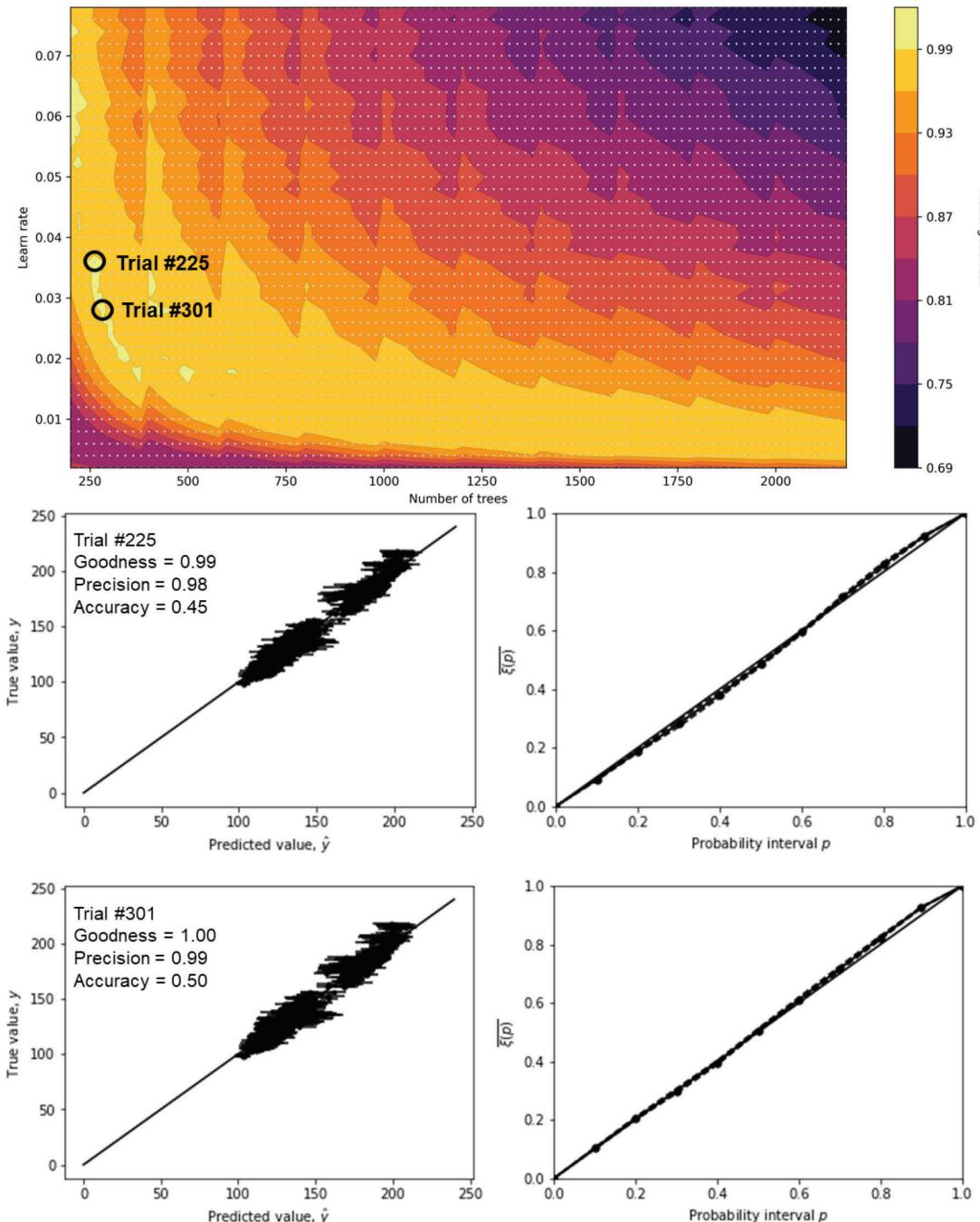


Fig. 8—Using Eq. 4, we color the hyperparameter space over testing data. Yellow colors represent the maximum goodness values. Similarly, we select two trials from the grid search and display the crossplot of the true values vs. those predicted by the ensemble model and their respective accuracy plot.

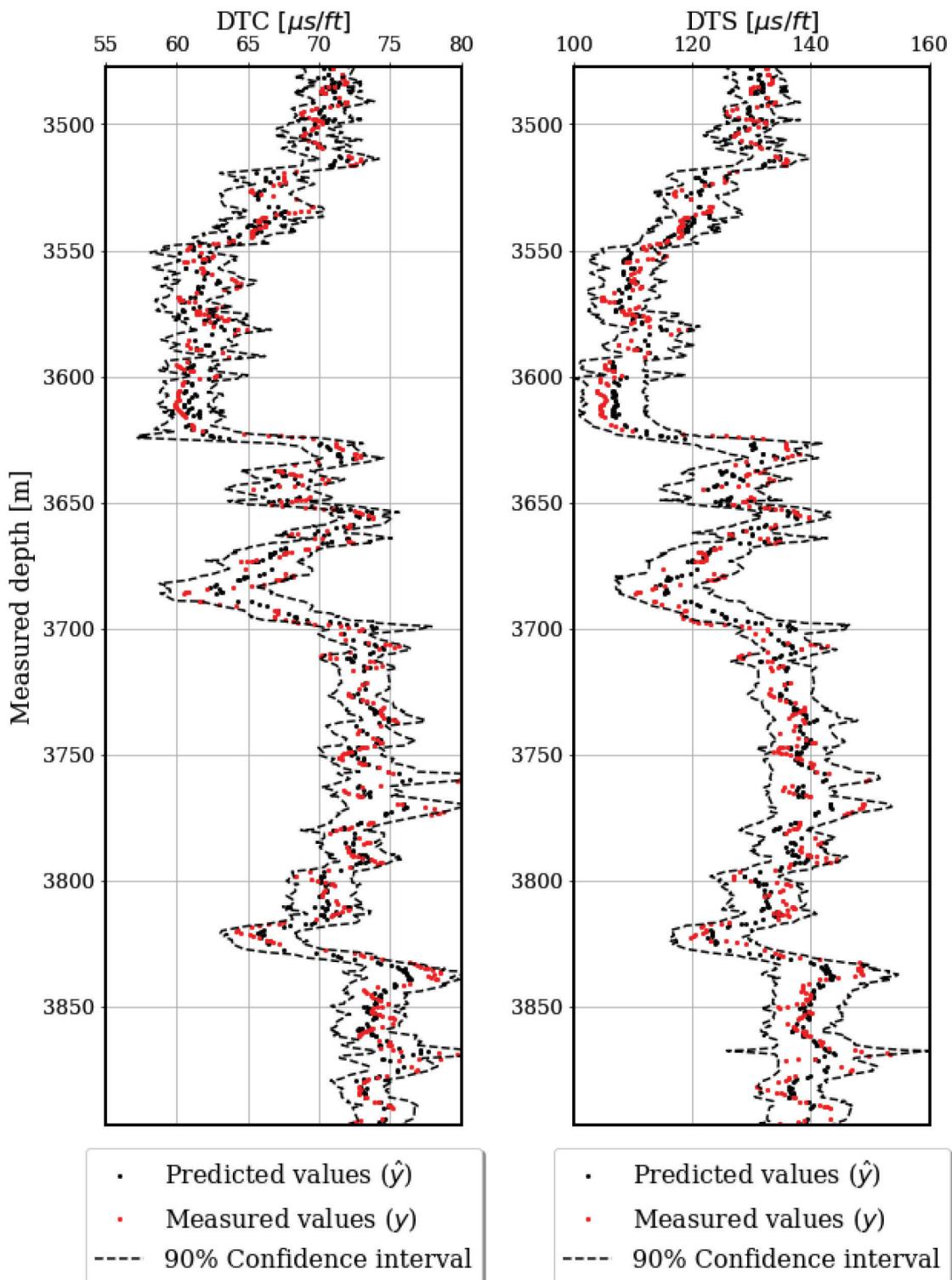


Fig. 9—Testing data comparison of the single predictions, estimation ensemble 90% confidence interval, and measured values with respect to measured depth for Model 302, where hyperparameters were optimized using Eq. 4.

To show the reduction in uncertainty with our proposed workflow as we include additional information from correlated wells, we separate the training and testing data into Well 1, Well 2, and Well 3 and repeat the training and hyperparameter tuning using Eq. 4. We obtain three different models. Model 1 is trained using the data from Well 1, Model 2 is trained using the data from Well 1 and 2, and Model 3 is trained using the data from Well 1, 2, and 3.

Next, we predict DTC values and show in Fig. 10 the reduction in uncertainty as we integrate information from Well 2 and Well 3 into the training data set. Our models are subject to the constraints of accuracy and precision, and we want to account for all relevant data to reduce uncertainty in our predictions.

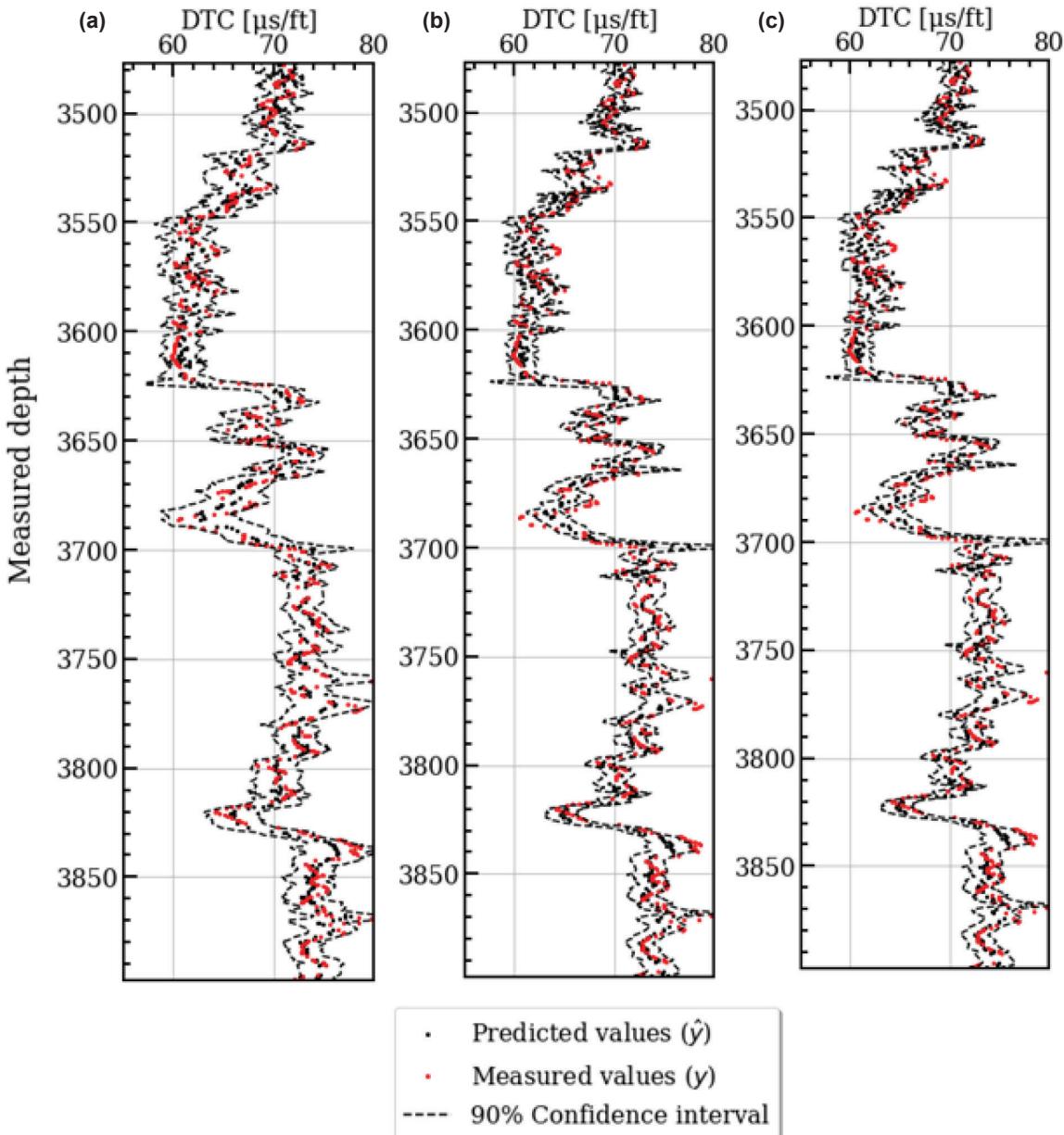


Fig. 10—Example of reduction in uncertainty in testing data as we include more information from correlated wells in the DTC log. (a) Predictions using the information from one well, (b) predictions using the information from two wells, and (c) predictions using the information from three correlated wells.

Using Model 3, trained with the information from Wells 1, 2, and 3, we predict the sonic logs from two wells with missing information and prepare Fig. 11. We explore the interval of 3,475 to 3,900 m. Model 3 is subject to

the constraints of accuracy and precision with the least uncertainty and maximum model goodness.

In Fig. 12, we include the complete set of well logs used to predict DTS and DTC values.

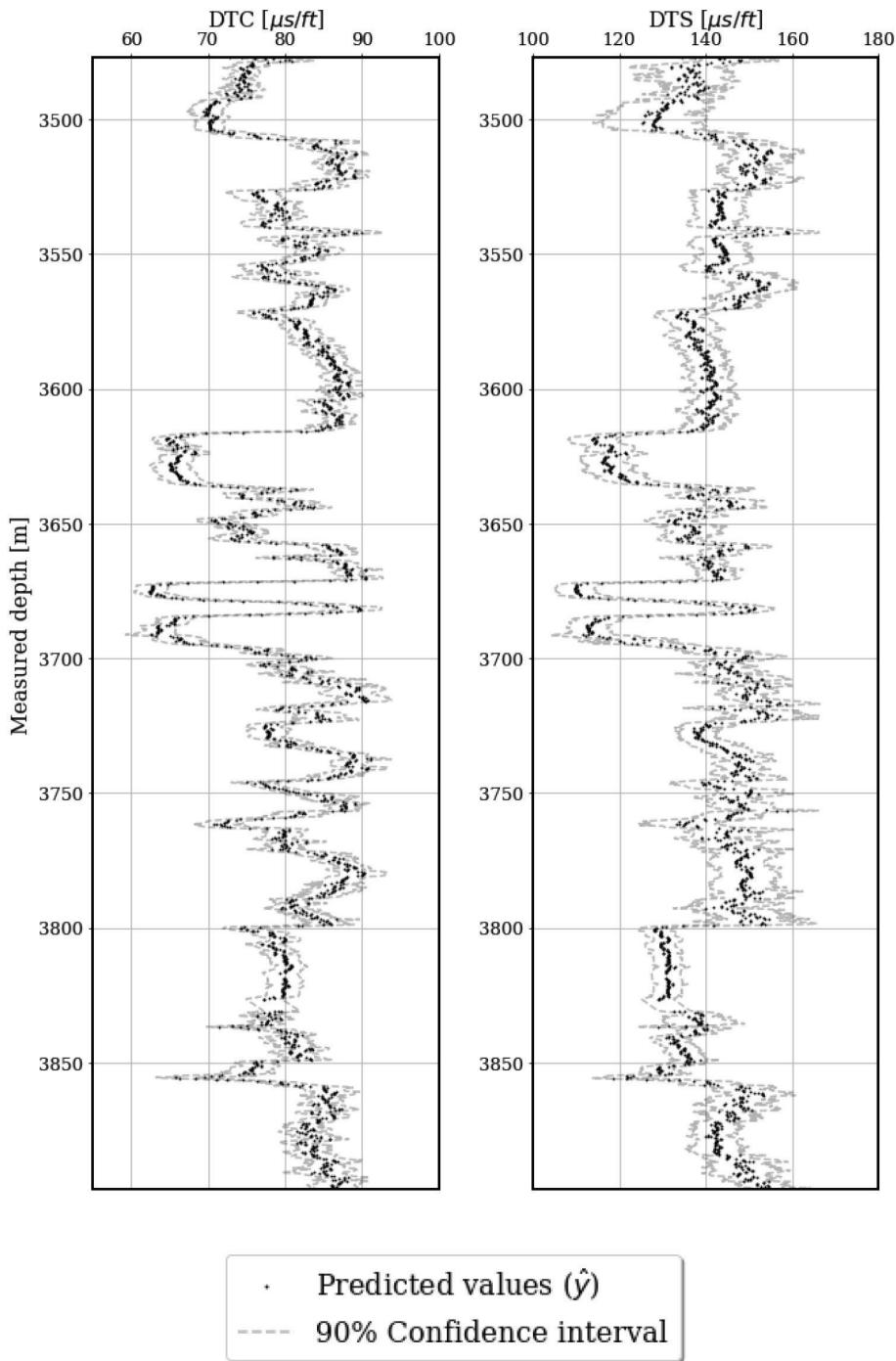


Fig. 11—Prediction example over testing data of DTC and DTS values and the 90% confidence interval for a correlated well with no sonic-logging information.

Sonic Well-Log Imputation Through Machine-Learning-Based Uncertainty Models

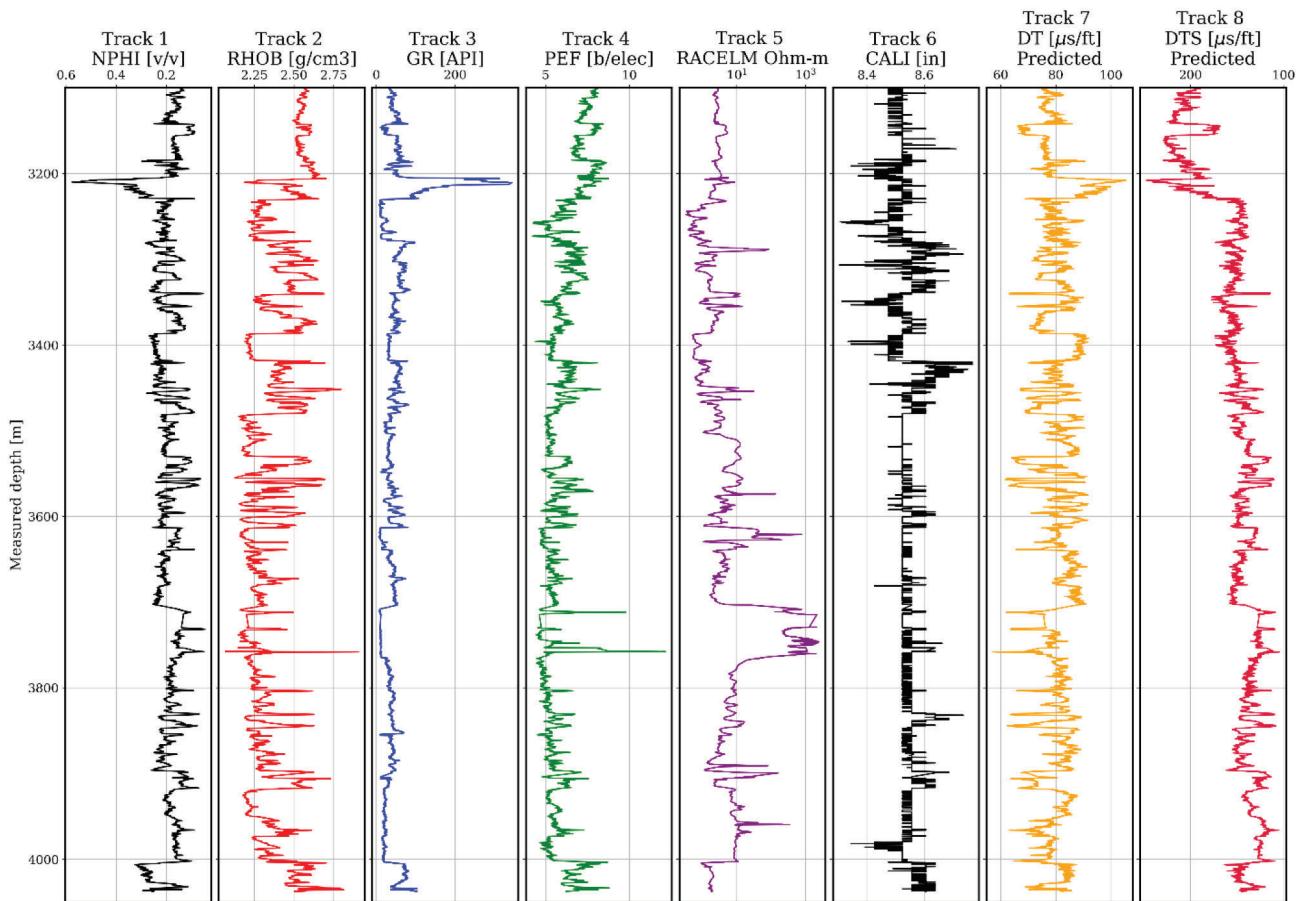


Fig. 12—Complete set of well logs from testing data used to predict DTC and DTS values in a well with no sonic-logging information.

CONCLUSIONS

We present and demonstrate a novel workflow for gradient boosting, decision-tree-based imputation of sonic, DTS, and DTC logs with accurate and precise predictions and uncertainty quantification. We provide a novel approach to quantifying and exploring the uncertainty in the imputation of missing well-log values.

Our proposed workflow uses the uncertainty model goodness metric to calculate robust uncertainty models and enables geoscientists and engineers to impute missing sonic well-logging information for improved geomechanical descriptions with uncertainty when sonic-logging measurements are unavailable.

ACKNOWLEDGMENTS

The authors thank the industry partners of the DIRECT consortium at The University of Texas at Austin for supporting this work. We gratefully recognize the two anonymous reviewers for their comments, which greatly improved this paper.

REPRODUCIBILITY

Workflow is publicly available on the author's GitHub repositories (github.com/emaldonadocruz and github.com/GeostatsGuy). We also provide an Uncertainty Tuning (UTuning) library in Python for hyperparameter tuning of machine-learning models (pypi.org/project/UTuning and github.com/emaldonadocruz/UTuning).

NOMENCLATURE

Abbreviations

| | |
|------|--|
| CALI | = caliper log, in. |
| DTC | = compressional slowness log, $\mu\text{s}/\text{ft}$ |
| DTS | = shear slowness log, $\mu\text{s}/\text{ft}$ |
| GBDT | = gradient boosting decision tree |
| GR | = gamma ray log, API |
| MAE | = mean absolute error, $\mu\text{s}/\text{ft}$ |
| MSE | = mean squared error, $\left[\frac{\mu\text{s}}{\text{ft}}\right]^2$ |
| NPHI | = neutron porosity log, v/v |
| PEF | = photoelectric factor log, b/elec |
| RHOB | = bulk density log, g/cm^3 |
| RT | = true resistivity log, $\Omega \cdot \text{m}$ |

Symbols

| | |
|----------------|--|
| F | = cumulative probability |
| h | = weak learner |
| M | = model |
| p | = probability interval |
| \mathbf{u} | = testing data point |
| X_{\max} | = maximum value of the predictor feature |
| X_{\min} | = minimum value of the predictor feature |
| X' | = normalized predictor feature value |
| y | = withheld testing values |
| ϵ | = learning rate |
| ξ | = indicator function |
| $\bar{\xi}(p)$ | = average indicator function |

REFERENCES

- Abdelaal, A., Elkataaty, S., and Abdulraheem, A., 2021, Pore Pressure Estimation While Drilling Using Machine Learning, Paper ARMA-IGS-21-115 presented at the ARMA/DGS/SEG 2nd International Geomechanics Symposium, held virtually, 1–4 November. DOI: 10.30632/PJV62N4-2021a4.
- Akinnikawe, O., Lyne, S., and Roberts, J., 2018, Synthetic Well Log Generation Using Machine Learning Techniques, Paper URTEC-2877021 presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, USA, 23–25 July. DOI: 10.15530/URTEC-2018-2877021.
- Alarifi, S.A., and Miskimins, J., 2021, A New Approach to Estimating Ultimate Recovery for Multistage Hydraulically Fractured Horizontal Wells by Utilizing Completion Parameters Using Machine Learning, Paper SPE-204470, *SPE Production & Operations*, **36**(3), 468–483. DOI: 10.2118/204470-PA.
- AlBahrani, H., and Morita, N., 2021, Risk-Controlled Wellbore Stability Criterion Based on a Machine-Learning-Assisted Finite-Element Model, Paper SPE-204101, *SPE Drilling & Completion*, **37**(1), 38–66. DOI: 10.2118/204101-PA.
- Al-Bulushi, N., Araujo, M., Kraaijveld, M., and Jing, X.D., 2007, Predicting Water Saturation Using Artificial Neural Networks (ANNs), Paper W, *Transactions, SPWLA Middle East Regional Symposium*, Abu Dhabi, UAE, 15–19 April.
- Alexeyev, A., Ostadhassan, M., Mohammed, R.A., Bubach, B., Khatibi, S., Li, C., and Kong, L., 2017, Well Log Based Geomechanical and Petrophysical Analysis of the Bakken Formation, Paper ARMA-2017-0942 presented at the 51st US Rock Mechanics/Geomechanics Symposium, San Francisco, California, USA, 25–28 June.
- Anifowose, F.A., Labadin, J., and Abdulraheem, A., 2017, Ensemble Machine Learning: An Untapped Modeling Paradigm for Petroleum Reservoir Characterization, *Journal of Petroleum Science and Engineering*, **151**(1), 480–487. DOI: 10.1016/j.petrol.2017.01.024.
- Asquith, G.B., Krygowski, D., Henderson, S., and Hurley, N., 2004, *Basic Well Log Analysis*, second edition, AAPG Methods in Exploration Series, **16**, American Association of Petroleum Geologists. ISBN: 9781629810492.
- Bateman, R.M., 2012, *Openhole Log Analysis and Formation Evaluation*, second edition, Society of Petroleum Engineers. ISBN: 978-1-61399-880-9.
- Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C., and Popp, J., 2013, Sample Size Planning for Classification Models, *Analytica Chimica Acta*, **760**(1), 25–33. DOI: 10.1016/j.aca.2012.11.007.
- Beleites, C., and Salzer, R., 2008, Assessing and Improving the Stability of Chemometric Models in Small Sample Size Situations, *Analytical and Bioanalytical Chemistry*, **390**(5), 1261–1271. DOI: 10.1007/s00216-007-1818-6.
- Blundell, C., Cornebise, J., Kavukcuoglu, K., and Wierstra, D., 2015, Weight Uncertainty in Neural Network, *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, **37**(1), 1613–1622. DOI: 10.5555/3045118.3045290. URL: <http://proceedings.mlr.press/v37/blundell15.pdf>. Accessed January 12, 2023.
- Chipman, H.A., George, E.I., and McCulloch, R.E., 2010, BART: Bayesian Additive Regression Trees, *The Annals of Applied Statistics*, **4**(1), 266–298. DOI: 10.1214/09-AOAS285.
- Deng, T., Xu, C., Jobe, D., and Xu, R., 2019, A Comparative Study of Three Supervised Machine-Learning Algorithms for Classifying Carbonate Vuggy Facies in the Kansas Arbuckle Formation, *Petrophysics*, **60**(6), 838–853. DOI: 10.30632/PJV60N6-2019a8.
- Depeweg, S., Hernández-Lobato, J.-M., Doshi-Velez, F., and Udluft, S., 2018, Decomposition of Uncertainty in Bayesian Deep Learning for Efficient and Risk-Sensitive Learning, *Proceedings of the 35th International Conference on Machine Learning*, Stockholm Sweden, **80**(1), 1184–1193. URL: <https://proceedings.mlr.press/v80/depeweg18a.html>. Accessed January 12, 2023.
- Deutsch, C.V., 1997, Direct Assessment of Local Accuracy and Precision, in Baafi, E.Y., and Schofield, N.A., editors, *Geostatistics Wollongong '96*, **1**, 115–125, Kluwer Academic Publishers, The Netherlands. URL: <https://ccg-server>.

- engineering.ualberta.ca/CCG%20Publications/Other/CVD%20Papers/01-Peer%20Reviewed/1997-1996/assess-local-accr-prec96.pdf. Accessed January 12, 2023.
- Dietterich, T.G., 2000, Ensemble Methods in Machine Learning, in Kittler, J., and Roli, F., editors, *MCS 2000: Multiple Classifier Systems, Lecture Notes in Computer Science*, **1857**, 1–15, Springer-Verlag, Berlin, Heidelberg. DOI: 10.1007/3-540-45014-9_1.
- Dong, S., Zeng, L., Du, X., He, J., and Sun, F., 2022, Lithofacies Identification in Carbonate Reservoirs by Multiple Kernel Fisher Discriminant Analysis Using Conventional Well Logs: A Case Study in an Oilfield, Zagros Basin, Iraq, *Journal of Petroleum Science and Engineering*, **210**, 110081. DOI: 10.1016/j.petrol.2021.110081.
- Du, Y., Tan, W., Jiang, C., Lu, D., and Li, D., 2008, An Effective Hash-Based Method for Generating Synthetic Well Log, *Proceedings, 2008 Third International Conference on Pervasive Computing and Applications*, **2**(1), 1017–1020. DOI: 10.1109/ICPCA.2008.4783672.
- Duan, T., Avati, A., Ding, D.Y., Thai, K.K., Basu, S., Ng, A.Y., and Schuler, A., 2020, NGBoost: Natural Gradient Boosting for Probabilistic Prediction, *Proceedings of the 37th International Conference on Machine Learning*, held virtually, 13–18 July. URL: <http://proceedings.mlr.press/v119/duan20a/duan20a.pdf>. Accessed January 12, 2023.
- Elkhatatny, S.M., Zeeshan, T., Mahmoud, M., Abdulazeez, A., and Mohamed, I.M., 2016, Application of Artificial Intelligent Techniques to Determine Sonic Time From Well Logs, Paper ARMA-2016-755 presented at the 50th US Rock Mechanics/Geomechanics Symposium, Houston, Texas, USA, 26–29 June.
- Equinor, 2018, Volve Data Village Dataset, URL: <https://data.equinor.com/dataset/Volve>. Accessed January 12, 2023.
- Eshkalak, M.O., Mohaghegh, S.D., and Esmaili, S., 2013, Synthetic, Geomechanical Logs for Marcellus Shale, Paper SPE-163690 presented at the SPE Digital Energy Conference, The Woodlands, Texas, USA, 5–7 March. DOI: 10.2118/163690-MS.
- Eskandari, H., Rezaee, M.R., and Mohammadnia, M., 2004, Application of Multiple Regression and Artificial Neural Network Techniques to Predict Shear Wave Velocity From Wireline Log Data for a Carbonate Reservoir South-West Iran, *CSEG Recorder*, **29**(7), 1–18. URL: <https://csegrecorder.com/articles/view/application-of-multiple-regression-and-artificial-neural-network-techniques>. Accessed January 12, 2023.
- Esmaeili, B., Rahimpour-Bonab, H., Kadkhodaie, A., Ahmadi, A., and Hosseinzadeh, S., 2022, Developing a Saturation-Height Function for Reservoir Rock Types and Comparing the Results With the Well Log-Derived Water Saturation, A Case Study From the Fahliyan Formation, Dorood Oilfield, Southwest of Iran, *Journal of Petroleum Science and Engineering*, **212**, 110268. DOI: 10.1016/j.petrol.2022.110268.
- Friedman, J.H., 2001, Greedy Function Approximation: A Gradient Boosting Machine, *The Annals of Statistics*, **29**(5), 1189–1232. DOI: 10.1214/aos/1013203451.
- Gal, Y., and Ghahramani, Z., 2016, Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, *Proceedings of the 33rd International Conference on Machine Learning*, **48**, 1050–1059. DOI: 10.5555/3045390.3045502.
- Hartigan, J.A., 1998, The Maximum Likelihood Prior, *The Annals of Statistics*, **26**(6), 2083–2103. DOI: 10.1214/aos/1024691462.
- Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.R., 2012, Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors, arXiv, arXiv:1207.0580. DOI: 10.48550/arXiv.1207.0580.
- Hutter, F., Hoos, H., and Leyton-Brown, K., 2014, An Efficient Approach for Assessing Hyperparameter Importance, *Proceedings of the 31st International Conference on Machine Learning*, Beijing, China, **32**(1), 754–762. DOI: 10.5555/3044805.3044891.
- Jamali, J., Javaherian, A., Wang, Y., and Ameri, M.J., 2022, The Behavior of Elastic Moduli With Fluid Content in the VTI Media, *Journal of Petroleum Science and Engineering*, **208**(A), 109308. DOI: 10.1016/j.petrol.2021.109308.
- Jia, X., Yang, J., Liu, R., Wang, X., Cotofana, S.D., and Zhao, W., 2021, Efficient Computation Reduction in Bayesian Neural Networks Through Feature Decomposition and Memorization, *IEEE Transactions on Neural Networks and Learning Systems*, **32**(4), 1703–1712. DOI: 10.1109/TNNLS.2020.2987760.
- Kazak, A., Simonov, K., and Kulikov, V., 2021, Machine-Learning-Assisted Segmentation of Focused Ion Beam-Scanning Electron Microscopy Images With Artifacts for Improved Void-Space Characterization of Tight Reservoir Rocks, Paper SPE-205347, *SPE Journal*, **26**(4), 1739–1758. DOI: 10.2118/205347-PA.
- Kim, S., Kim, K.H., Min, B., Lim, J., and Lee, K., 2020, Generation of Synthetic Density Log Data Using Deep Learning Algorithm at the Golden Field in Alberta, Canada, *Geofluids*, **2020**, Article ID 5387183, 1–26. DOI: 10.1155/2020/5387183.
- Kwon, Y., Won, J.-H., Kim, B.J., and Paik, M.C., 2020, Uncertainty Quantification Using Bayesian Neural Networks in Classification: Application to Biomedical Image Segmentation, *Computational Statistics & Data Analysis*, **142**, 106816. DOI: 10.1016/j.csda.2019.106816.
- Lakshminarayanan, B., Pritzel, A., and Blundell, C., 2017, Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles, *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6405–6416. DOI: 10.5555/3295222.3295387.
- Ledesma, A., Gens, A., and Alonso, E.E., 1996, Estimation of Parameters in Geotechnical Backanalysis—I. Maximum Likelihood Approach, *Computers and Geotechnics*, **18**(1), 1–27. DOI: 10.1016/0266-352X(95)00021-2.

- Li, H., and Misra, S., 2017, Prediction of Subsurface NMR T2 Distribution From Formation-Mineral Composition Using Variational Autoencoder, *SEG Technical Program Expanded Abstracts 2017*, 3350–3354. DOI: 10.1190/segam2017-17798488.1.
- Li, Y., Shen, J., Cai, W., and Zhou, X., 2022, Fractured Formation Evaluation by Seismic Attenuation Derived From Array Acoustic Log Waves Based on Modified Spectral Ratio Method and an Extended Biot's Poroelastic Model, *Journal of Petroleum Science and Engineering*, **209**, 109838. DOI: 10.1016/j.petrol.2021.109838.
- Liang, L., Lei, T., Donald, A., and Blyth, M., 2021, Physics-Driven Machine-Learning-Based Borehole Sonic Interpretation in the Presence of Casing and Drillpipe, Paper SPE-201542, *SPE Reservoir Evaluation & Engineering*, **24**(2), 310–324. DOI: 10.2118/201542-PA.
- Linero, A.R., 2017, A Review of Tree-Based Bayesian Methods, *Communications for Statistical Applications and Methods*, **24**(6), 543–559. DOI: 10.29220/CSAM.2017.24.6.543.
- Liu, H.-H., Zhang, J., Liang, F., Temizel, C., Basri, M.A., and Mesdour, R., 2021, Incorporation of Physics Into Machine Learning for Production Prediction From Unconventional Reservoirs: A Brief Review of the Gray-Box Approach, Paper SPE-205520, *SPE Reservoir Evaluation & Engineering*, **24**(4), 847–858. DOI: 10.2118/205520-PA.
- Long, W., Chai, D., and Aminzadeh, F., 2016, Pseudo Density Log Generation Using Artificial Neural Network, Paper SPE-180439 presented at the SPE Western Regional Meeting, Anchorage, Alaska, USA, 23–26 May. DOI: 10.2118/180439-MS.
- Maldonado-Cruz, E., and Pyrcz, M.J., 2021, Tuning Machine Learning Dropout for Subsurface Uncertainty Model Accuracy, *Journal of Petroleum Science and Engineering*, **205**, 108975. DOI: 10.1016/j.petrol.2021.108975.
- Maldonado-Cruz, E., and Pyrcz, M.J., 2022, Fast Evaluation of Pressure and Saturation Predictions With a Deep Learning Surrogate Flow Model, *Journal of Petroleum Science and Engineering*, **212**, 110244. DOI: 10.1016/j.petrol.2022.110244.
- Maleki, S., Moradzadeh, A., Riabi, R.G., Gholami, R., and Sadeghzadeh, F., 2014, Prediction of Shear Wave Velocity Using Empirical Correlations and Artificial Intelligence Methods, *NRIAG Journal of Astronomy and Geophysics*, **3**(1), 70–81. DOI: 10.1016/j.nrjag.2014.05.001.
- Malinin, A., Prokhorenkova, L., and Ustimenko, A., 2021, Uncertainty in Gradient Boosting via Ensembles, arXiv, arXiv:2006.10562v4. DOI: 10.48550/arxiv.2006.10562.
- Misra, S., Li, H., and He, J., 2021, Chapter 5 – Robust Geomechanical Characterization by Analyzing the Performance of Shallow-Learning Regression Methods Using Unsupervised Clustering Methods, in Misra, S., Li, H., and He, J., editors, *Machine Learning for Subsurface Characterization*, 129–155, Gulf Professional Publishing. DOI: 10.1016/B978-0-12-817736-5.00005-3.
- Neff, P., Steineder, D., Stummer, B., and Clemens, T., 2021, Estimation of Initial Hydrocarbon Saturation Applying Machine Learning Under Petrophysical Uncertainty, Paper SPE-203384, *SPE Reservoir Evaluation & Engineering*, **24**(2), 325–340. DOI: 10.2118/203384-PA.
- Olkogba, T.A., and Feng, Y., 2021, Practical Machine-Learning Applications in Well-Drilling Operations, Paper SPE-205480, *SPE Drilling & Completion*, **36**(4), 849–867. DOI: 10.2118/205480-PA.
- Peng, H., and Fan, Y., 2017, Feature Selection by Optimizing a Lower Bound of Conditional Mutual Information, *Information Sciences*, **418–419**(1), 652–667. DOI: 10.1016/j.ins.2017.08.036.
- Peng, H., Long, F., and Ding, C., 2005, Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**(8), 1226–1238. DOI: 10.1109/TPAMI.2005.159.
- Rolon, L.F., Mohaghegh, S.D., Ameri, S., Gaskari, R., and McDaniel, B., 2005, Developing Synthetic Well Logs for the Upper Devonian Units in Pennsylvania, Paper SPE-98013 presented at the SPE Eastern Regional Meeting, Morgantown, West Virginia, USA, 14–16 September. DOI: 10.2118/98013-MS.
- Savitzky, A., and Golay, M.J.E., 1964, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, *Analytical Chemistry*, **36**(8), 1627–1639. DOI: 10.1021/ac60214a047.
- Schölkopf, B., Williamson, R.C., Smola, A., Shawe-Taylor, J., and Platt, J., 1999, Support Vector Method for Novelty Detection, *Proceedings of the 12th International Conference on Neural Information Processing Systems*, 582–588, MIT Press, Cambridge, MA, USA. DOI: 10.5555/3009657.3009740.
- Singh, D., and Singh, B., 2020, Investigating the Impact of Data Normalization on Classification Performance, *Applied Soft Computing*, **97**(B), 105524. DOI: 10.1016/j.asoc.2019.105524.
- Souder, W.W., 2002, Using Sonic Logs to Predict Fluid Type, *Petrophysics*, **43**(5), 412–419.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., 2014, Dropout: A Simple Way to Prevent Neural Networks From Overfitting, *Journal of Machine Learning Research*, **15**(56), 1929–1958. URL: <http://jmlr.org/papers/v15/srivastava14a.html>. Accessed January 12, 2023.
- Tan, M., Peng, X., Cao, H., Wang, S., and Yuan, Y., 2015, Estimation of Shear Wave Velocity From Wireline Logs in Gas-Bearing Shale, *Journal of Petroleum Science and Engineering*, **133**, 352–366. DOI: 10.1016/j.petrol.2015.05.020.
- Tariq, Z., Elkhatatny, S., Mahmoud, M., and Abdulraheem, A., 2016, A New Artificial Intelligence Based Empirical Correlation to Predict Sonic Travel Time, Paper IPTC-19005 presented at the International Petroleum Technology Conference, Bangkok, Thailand, 14–16 November. DOI: 10.2523/IPTC-19005-MS.

- Tatsipie, N.R.K., and Sheng, J.J., 2021, Generating Pseudo Well Logs for a Part of the Upper Bakken Using Recurrent Neural Networks, *Journal of Petroleum Science and Engineering*, **200**, 108253. DOI: 10.1016/j.petrol.2020.108253.
- Vanegas, G., Nejedlik, J., Neff, P., and Clemens, T., 2021, Conditioning Model Ensembles to Various Observed Data (Field and Regional Level) by Applying Machine-Learning-Augmented Workflows to a Mature Field With 70 Years of Production History, Paper SPE-205188, *SPE Reservoir Evaluation & Engineering*, **24**(4), 809–826. DOI: 10.2118/205188-PA.
- Xu, Y., and Goodacre, R., 2018, On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning, *Journal of Analysis and Testing*, **2**(3), 249–262. DOI: 10.1007/s41664-018-0068-2.
- Ye, Y., Tang, S., Xi, Z., Jiang, D., and Duan, Y., 2022, A New Method to Predict Brittleness Index for Shale Gas Reservoirs: Insights From Well Logging Data, *Journal of Petroleum Science and Engineering*, **208**(B), 109431. DOI: 10.1016/j.petrol.2021.109431.
- Yin, J., and Li, N., 2022, Ensemble Learning Models With a Bayesian Optimization Algorithm for Mineral Prospectivity Mapping, *Ore Geology Reviews*, **145**, 104916. DOI: 10.1016/j.oregeorev.2022.104916.
- Yu, Y., Xu, C., Misra, S., Li, W., Ashby, M., Pan, W., Deng, T., Jo, H., Santos, J.E., Fu, L., Wang, C., Kalbekov, A., Suarez, V., Kusumah, E.P., Aviandito, M., Pamadya, Y., and Izadi, H., 2021, Synthetic Sonic Log Generation With Machine Learning: A Contest Summary From Five Methods, *Petrophysics*, **62**(4), 393–406. DOI: 10.30632/PJV62N4-2021a4.
- Zhang, J., Lang, J., and Standifird, W., 2009, Stress, Porosity, and Failure-Dependent Compressional and Shear Velocity Ratio and its Application to Wellbore Stability, *Journal of Petroleum Science and Engineering*, **69**(3–4), 193–202. DOI: 10.1016/j.petrol.2009.08.012.
- Zhang, W., Wu, C., Zhong, H., Li, Y., and Wang, L., 2021, Prediction of Undrained Shear Strength Using Extreme Gradient Boosting and Random Forest Based on Bayesian Optimization, *Geoscience Frontiers*, **12**(1), 469–477. DOI: 10.1016/j.gsf.2020.03.007.

ABOUT THE AUTHORS

Eduardo Maldonado-Cruz is a PhD candidate in petroleum engineering at The University of Texas at Austin in the Hildebrand Department of Petroleum and Geosystems Engineering. Before that, he worked as a lecturer at the National Autonomous University of Mexico (UNAM) and as a reservoir engineer for a private oil company in Mexico, where he developed multiple development projects in tight gas. He received his bachelor's and master's degrees from the National Autonomous University of Mexico (UNAM). Eduardo's specializations include reservoir engineering, numerical simulation, machine learning, and uncertainty modeling.

John T. Foster is an associate professor in the Hildebrand Department of Petroleum and Geosystems Engineering, the Department of Aerospace Engineering and Engineering Mechanics, and a core faculty member at the Oden Institute for Computational Engineering and Sciences at The University of Texas at Austin. Before joining UT-Austin, John was a faculty member in mechanical engineering at UTSA and was a senior member of the technical staff at Sandia National Laboratories, where he worked for 7 years. John received his BS and MS degrees in mechanical engineering from Texas Tech University and PhD from Purdue University. John is a registered Professional Engineer in the State of Texas and the cofounder and CTO of daytum. During his career in research, John has been involved in many projects ranging from full-scale projectile penetration field tests to laboratory experiments using Kolsky bars to modeling and simulation efforts using some of the world's largest computers. John's research interests are in experimental and computational mechanics and multiscale modeling with applications to geomechanics, impact mechanics, fracture mechanics, and anomalous transport processes.

Michael J. Pyrcz is an expert in geostatistical reservoir modeling. He has 15 years of experience consulting and teaching and has published over 30 peer-reviewed papers and one textbook on the topic. At Chevron, Michael was a company-wide expert responsible for (1) conducting and managing reservoir modeling and related R&D, (2) conducting and reviewing reservoir modeling projects, and (3) training and establishing standards for reservoir modeling. Michael joined The University of Texas at Austin as an associate professor in the Hildebrand Department of Petroleum and Geosystems Engineering and the Jackson School of Geosciences.