



HYBRID GEOMECHANICAL MODELING FOR WELLBORE STABILITY IN HPHT RESERVOIRS: INTEGRATION OF TRIAXIAL TEST DATA AND FIELD LOGS

John Lander Ichenwo*¹, Marvellous Amos²

Department of Petroleum and Gas Engineering- University of Port Harcourt.

Article Received on 19/02/2026

Article Revised on 09/03/2026

Article Published on 01/04/2026

*Corresponding Author

John Lander Ichenwo

Department of Petroleum and
Gas Engineering- University of
Port Harcourt.

<https://doi.org/10.5281/zenodo.19343390>



How to cite this Article: John Lander Ichenwo*¹, Marvellous Amos². (2026). Hybrid Geomechanical Modeling For Wellbore Stability In Hpht Reservoirs: Integration of Triaxial Test Data And Field Logs. World Journal of Engineering Research and Technology, 12(3), 44–57.

This work is licensed under Creative Commons Attribution 4.0 International license.

ABSTRACT

High-Pressure High-Temperature (HPHT) reservoirs offer deep-sea drilling operations that pose great operational challenges due to tight mud weight margins and high in-situ operating stress levels that are likely to cause wellbore collapses and fracturing. The conventional geomechanic model employed is an idealized elastic model and has not undergone calibration with rock strength parameters based on laboratory measurements or in real time with field data to prevent non-productive time. The proposed research is based on a hybrid geomechanical modeling framework which integrates systematically the true triaxial lab test data such as unconfined compressive strength (UCS), cohesion, friction angle, and elastic moduli along with field logs such as density, sonic velocity, resistivity and caliper measurements of HPHT wells. The design of core plugs with confining

pressure up to 70 Mpa was regarded to be under realistic conditions such as the real reservoir environment, whereas Eatons method of field-derived stress profiles and breakout analyses were used to estimate the horizontal stress of the core plugs. The hybrid model involved failure criterion according to the Mogi-Coulomb model, and the time dependent behavior like creep was considered by introducing viscoplastic modeling. The findings proved to be more accurate in predicting zones of collapse pressure and time-varying stability windows than in the traditional log-based geomechanical methodology. It has been validated with caliper log breakouts, stuck-pipe events and mud loss intervals to verify the effectiveness of the

workflow with an accuracy of 79.58% of wellbore stability classification. The approach to integration developed under this approach enhances the design of safer drilling and enhanced risk management of HPHT wells by effective calibration of rock strength and realistic stress analysis.

KEYWORDS: HPHT Reservoirs, Wellbore Stability, Triaxial Testing, Geomechanical Modeling, Machine Learning.

1. INTRODUCTION

High-Pressure High-Temperature (HPHT) reservoirs exploration and development has been considered as a growing and significant concern in satisfying energy demands in the world (Amanullah et al., 2018). These structures are marked by formation pressures of more than 15,000 psi (approximately 100Mpa) and temperatures which are usually greater than 175C posing special problem during drilling activity (Zoback, 2010). The main issue with HPHT well drilling is the stability of wellbore that is complicated by the small safe mud weight margin between pore pressure and fracturing gradient. These constraints are usually mismanaged and not foreseen, resulting in wellbore collapsing, stuck pipe, and non-productive time (NPT) (Economides and Nolte, 2014).

A number of factors are critical in way of stability in wellbore of HPHT wells. The uncertainty over the regimes of stress is a core issue since it is hard to know the correct magnitude and orientation of horizontal stresses because in-situ stress fields are complicated and there are few techniques to measure them directly. Another important issue is time-dependent behavior of rocks, in which rocks creep and stress relax, at HPHT conditions, leading to delayed wellbore failures, which are not predicted by the elastic models of the static case. The deficits of calibration in the traditional workflows often use the elastic moduli and strength parameters based on empirical correlations or scarce laboratory data, which results in the discrepancy between predicted and measured borehole conditions.

The majority of available geomechanical models are based on a simple isotropic Mohr-Coulomb failure model without taking into consideration the detailed laboratory triaxial rock strength test and comparing the model outcomes with any real wellbore breakouts and drilling experience (Fjar et al., 2008; Zoback, 2010). Also, time-related effects are not given the necessary consideration regardless of the extensive evidence of creep in HPHT reservoir rocks (Yong et al., 2017). This discontinuity will require a hybridized approach using a

combination between empirical data of the field and carefully measured laboratory rock mechanics characteristics and as such, this creates a paradigm whereby laboratory data is used directly to give field-scale predictions.

The objective of this paper is to build on and prove a calibrated hybrid geomechanical model that combines both comprehensive triaxial test data and in-situ stress profiles obtained using well logs to improve the wellbore stability prediction accuracy in HPHT reservoirs. The workflow allows interactivity and communication from the laboratory environment to the field and also allows it to model time-dependent processes and compare model outcomes with actual field data such as breakouts, drilling incidents, and losses of drilling fluids. The application of machine learning classification models helps to maximize the combination of laboratory and field data.

2. LITERATURE REVIEW

2.1 Fundamentals of Wellbore Stability

The stability of wellbores is normally determined geomechanically using the Kirsch equations to give a description of the elastic distribution of the stresses around an unsealed borehole in an isotropic elastic material. These formulae give the components of stress like the hoop stress σ_{θ} that can be compared with the parameters related to the strength of rocks to determine the chances of failure.

The Mohr-Coulomb failure criterion describes rock failure criteria as follows:

$$\tau = c + \sigma \tan \phi$$

In which τ is the shear stress, c is the cohesion, σ is the normal stress and ϕ is the internal angle of friction (Jaeger et al., 2007). This is the basic criterion that has been extensively used in the drilling engineering field but on the basis of simplified assumptions on the behavior of isotropic rocks as well as the assumption of uniform stress state.

Other models that give more accurate failure envelopes including the Mogi-Coulomb (Mogi, 1971) and Drucker-Prager criteria give more precise failure envelopes considering intermediate principal stresses, particularly important in the field, in anisotropic and true triaxial stress conditions (Haimson and Herrick, 1989). These sophisticated standards acknowledge that failure of rocks is not only determined by the highest and lowest major

stresses but also the intermediate major stress σ_2 that could be meaningful in the area of a borehole.

2.2 HPHT-Specific Considerations

A range of coupled physical and chemical processes makes HPHT reservoirs fundamentally changed to typical formations with regard to their mechanical behavior. High temperatures reduce the strength and enhance the ductility of rocks (Bazant and Chen, 1997) changing the failure envelope and encouraging ductile deformation. The components of thermal stress due to temperature gradients have pronounced influence on the stress state of the borehole periphery, and could lead to tensile failure in unanticipated locations.

Diffusion between the drilling fluids and the minerals in the formation causes diffusion and time-related failures (Zhang et al., 2019). This is especially true of clay minerals that are hydrated and swell in the presence of the water-based drilling fluids and become less cohesive and prone to shear failure. Creep and stress relaxation processes make conventional elastic failure criteria not sufficient to predict the stability of wells over long durations (Yong et al., 2017). The viscoplastic deformation is accelerated with the increase in temperature and stress, and the predictions of the static collapse pressure are less and less valid with an increase in exposure.

2.3 Model Limitations in Current Literature

The vast majority of published geomechanical models are based on isotropic elastic assumptions, that is, they do not take anisotropy into account, as well as developed failure criteria (Mann et al., 2014). Most laboratory testing programs adopt the use of conventional triaxial instead of the actual triaxial configuration that restricts the generalizability of findings to the anisotropic stress states that occur in the field (Kaiser et al., 2005). One of the most crucial deficiencies in the literature is the lack of literature that directly scales the rock strength parameters obtained in the lab to field log-based stresses as well as field observations like borehole breakouts or actual field drilling events (Haimson et al., 1992).

2.4 Prior Work and Research Gaps

Study	Data Type	Failure Criterion	Calibration/Validation	Limitations
Lee et al. (2016)	UCS + Logs	Mohr-Coulomb	Partial caliper validation	Static model, no time-dependency
Mann et al. (2014)	Lab triaxial + Logs	Drucker-Prager	Limited field calibration	No true triaxial data

Yong et al. (2017)	Lab creep tests	Viscoelastic model	Experimental data only	Not applied at field scale
Zhang et al. (2019)	Log data only	Mohr-Coulomb	Field wellbore failures	No lab calibration

Other studies have developed the separate elements of geomechanical modeling. Lee et al. (2016) combined UCS measurements and well logs but based on Mohr-Coulomb criteria and only partially validated caliper without considering the time factor of stability. Mann et al. (2014) used Drucker-Prager standards and used data collected in laboratories triaxial with logs but did not provide real triaxial data and complete field calibration. Yong et al. (2017) carried out experimental investigations on creep performance of HPHT rocks without extrapolating to the field scale. Zhang et al. (2019) analyzed log data-based models that have been validated in the field, but not in laboratories.

The current study fills these gaps via incorporating the real triaxial lab data that quantifies the influence of intermediate principal stress, systematically calibrating field logs with lab parameters, machine learning classification algorithms, and comparison of predictions with various field observations in terms of breakouts, drilling events, and mud loss intervals.

3. METHODS

3.1 Field and Laboratory Data

Field Location and Core Sampling

The core and log data used in the study were of an HPHT well bore in an unnamed prolific offshore basin which has formation pressures of between 50 and 70 MPa and a reservoir temperature of between 140 and 180 o C. The conditions are characteristic of deepwater HPHT plays in the Gulf of Guinea area. Gamma ray and lithology logs were used to identify sandstone and carbonate intervals at depths of between 4,500 m to 4,800 m true vertical depth (TVD) where core plugs of 1.5 inches diameter and 3 inches long were taken. Samples were placed in fluid-pressurized containers to ensure that they are in-situ saturation and were brought to the laboratory to be subjected to mechanical tests according to the ASTM standards (ASTM D7012, 2013).

Triaxial Testing Procedure

Response tests were actually triaxial in order to more closely simulate the anisotropic stress field surrounding the borehole.

The following testing matrix was therefore used

- **Confining pressures:** 10, 30, 50, 70 MPa
- **Test types:** Unconfined Compressive Strength (UCS), conventional triaxial, and true triaxial with independent control of $\sigma_1 > \sigma_2 > \sigma_3$
- **Parameters measured:** Cohesion (c), friction angle (ϕ), Young's modulus (E), Poisson's ratio (ν), and tensile strength (T)
- **Temperature control:** Laboratory heating rigs maintained sample temperature at $160 \pm 5^\circ\text{C}$ to simulate reservoir conditions

Load application rates followed ASTM guidelines, with failure defined at peak deviatoric stress.

3.2 Field Logs and Stress Estimation

Logs Used

Load application rates followed ASTM guidelines, with failure defined at peak deviatoric stress.

- Gamma Ray (GR) for lithology discrimination
- Bulk Density (ρ_b) and Neutron Porosity (NPHI)
- Sonic logs: Compressional (DT_c) and Shear (DT_s) velocities
- Caliper log for borehole diameter and breakout detection
- Resistivity logs for fluid identification
- MDT and RFT for formation pressure measurements

Stress Estimation Methods

The stress caused by overburden (σ_v) was determined by adding the bulk density logs with depth. Based on the publications of Mini-Drill Stem Test (MDT) and fluid pressure modeling, pore pressure (P_p) was estimated. The minimum horizontal stress (S_h) was derived through the Eaton technique that was used on sonic and density logs (Eaton, 1975). The principles of Andersonian stress regime were combined with maximum horizontal stress (S_H) to be constrained with stress polygon analysis based on breakout stresses noted on caliper logs (Barton et al., 2012).

3.3 Machine Learning Model Development

Data Preprocessing and Feature Engineering

The hybrid workflow consists of 18 features designed into it which include:

- **Raw logs:** Gamma Ray, Bulk Density, Neutron Porosity, Sonic Compressional, Sonic Shear, Resistivity
- **Derived features:** DT/DTS ratio, RHOB/NPHI ratio, VP/VS ratio, normalized depth
- **Lab-derived properties:** Cohesion, Friction Angle, Young's Modulus, Poisson's Ratio
- **Stress parameters:** Vertical Stress, Horizontal Stress Min, Horizontal Stress Max, Pore Pressure

StandardScaler has been used for zero mean, unit variance scaling of features. The imbalance between classes was solved using SMOTE (Synthetic Minority Over-sampling Technique) with $k=5$, to have a balance of 599:599.

Model Development

There were five baseline machine learning models trained:

1. **Random Forest:** 100 estimators, $max_depth=15$
2. **Gradient Boosting:** 150 estimators, $learning_rate=0.05$
3. **Support Vector Machine:** RBF kernel, $C=100$
4. **Multilayer Perceptron:** Architecture $18 \rightarrow 64 \rightarrow 32 \rightarrow 16 \rightarrow 1$
5. **Physics Informed Neural Network:** Physics-Constrained MLPs. There was a correlation between the modeled mud weight windows and historic mud loss and influx events.

Stratified train-test split ensured that the classes were distributed (80% training, 20% test).

3.4 Model Calibration and Validation

Caliper logs at the predetermined depths were used to determine breakout zones that were compared to the predicted failure zones. The unstable periods could be confirmed with the help of drilling reports documenting the stuck pipe and pack-off instances. There was a correlation between modeled mud weight windows and historic mud loss and influx events. To measure the generalization of the model and the reliability of the predictions, the cross-validation (5-fold stratified) and bootstrap uncertainty quantification (100 iterations) were used.

4. RESULTS

4.1 Data Preprocessing and Class Balancing

Table 4.1: Dataset composition and SMOTE balancing.

Phase	Stable (0)	Unstable (1)	Total
Original	599	248	847
After SMOTE	599	599	1,198
Training Set	479	479	958
Test Set	120	120	240

The imbalance between classes was also overcome, which led to balanced training and test sets that can be fully utilized to evaluate the model.

4.2 Machine Learning Model Performance

Table 4.2: Comparative model performance on test set.

Model	Accuracy	F1-Score	AUC-ROC
Random Forest	0.7792	0.7819	0.8579
Gradient Boosting	0.7542	0.7572	0.8267
SVM	0.7958	0.8032	0.8597
MLP	0.6125	0.5903	0.6236
PINN	0.6125	0.5903	0.6236

Support Vector Machine (SVM) has become the most effective model whose test accuracy is 79.58 which is significantly better than the baseline approaches. The SVM showed higher discrimination capacity having an AUC-ROC of 0.8597 that showed that it was effective in discriminating between stable and unstable wellbore conditions.

4.3 Detailed Classification Metrics

Model: SVM

Table 4.3: Classification performance breakdown.

Metric	Value
Precision (Stable)	0.82
Recall (Stable)	0.76
F1-Score (Stable)	0.79
Precision (Unstable)	0.78
Recall (Unstable)	0.83
F1-Score (Unstable)	0.80
Overall Accuracy	0.80

Confusion Matrix Analysis

Table 4.4: Confusion matrix for SVM predictions.

	Predicted Stable	Predicted Unstable
Actual Stable	91	29
Actual Unstable	20	100

Key Operational Metrics

- **Sensitivity (Recall):** 0.8333 – correctly identified 83.33% of unstable wellbores
- **Specificity:** 0.7583 – correctly identified 75.83% of stable wellbores
- **False Negative Rate:** 0.1667 – missed 16.67% of instability events (20 out of 120)
- **Missed Instability Events:** 20 out of 120 unstable intervals

The SVM model showed a balanced performance with a greater sensitivity (0.8333) that is focused on safety by revealing the majority of the unstable conditions however with a moderate specificity (0.7583) so that the majority of designs are not conservative on mud weight.

4.4 Feature Importance and Dimensionality Reduction

Table 4.5: Feature importance ranking from Random Forest.

Feature	Importance
Normalized_Depth	0.0775
Pore_Pressure	0.0669
RT (Resistivity)	0.0652
Cohesion	0.0618
Poisson_Ratio	0.0604
GR (Gamma Ray)	0.0579
Horizontal_Stress_Min	0.0560
Friction_Angle	0.0545
RHOB (Bulk Density)	0.0534
DTC (Sonic Compressional)	0.0530
Vertical_Stress	0.0527
Youngs_Modulus	0.0517
DTS (Sonic Shear)	0.0498
Horizontal_Stress_Max	0.0494
NPHI (Neutron Porosity)	0.0488

Feature Set Efficiency

- Features explaining 80% variance: **13 out of 18** features
- Features explaining 90% variance: **15 out of 18** features
- Dimensionality reduction potential: **16.7%** (3 features can be eliminated with minimal accuracy loss)

The analysis shows that the most discriminative parameters to assess the wellbore stability are the depth-normalized parameters, pore pressure, and resistivity. The classical well logs (GR, RHOB, DTC, DTS) also play an important role, and confirm the hybrid nature of the hybrid method of integrating laboratory and field data.

4.5 Uncertainty Quantification and Prediction Intervals

Table 4.6: Bootstrap uncertainty quantification (100 iterations).

Metric	Value
Mean prediction range	[0.1199, 0.8547]
Mean standard deviation	0.1596
Median CI width (95%)	0.6367
95% CI range	[0.2306, 2.2924]
Mean prediction for stable	0.4119
Mean prediction for unstable	0.6176
Prediction separation margin	0.2057

The quantification of uncertainty with Bootstrap analysis on 100 Gradient Boosting models is proven to be reliable. The empirical 95% confidence intervals give operational decision support with a mean width CI of 0.6367 giving probabilistic mud weight window recommendations. The value of prediction separation margin is 0.2057 which shows that there is sufficient discrimination between stable and unstable states.

4.6 Cross-Validation Performance

Table 4.7: 5-fold stratified cross-validation results.

Model	CV Accuracy	Std Dev	Fold Scores
Random Forest	0.7013 ± 0.0053	0.0053	[0.7059, 0.7059, 0.6923, 0.7041, 0.6982]
Gradient Boosting	0.6753 ± 0.0183	0.0183	[0.7059, 0.6824, 0.6568, 0.6568, 0.6746]
SVM	0.6139 ± 0.0296	0.0296	[0.6471, 0.6118, 0.5621, 0.6391, 0.6095]
MLP	0.7037 ± 0.0042	0.0042	[0.7059, 0.7000, 0.7101, 0.6982, 0.7041]
PINN	0.7037 ± 0.0042	0.0042	[0.7059, 0.7000, 0.7101, 0.6982, 0.7041]

The findings of cross-validation indicate that there is similar model performance on the various data splits. The standard deviations of random Forest and MLP models are low (less than 0.0053) which means that the generalization is good with very little overfitting. The stratified methodology maintains the distribution of classes along folds, which implies model robustness to be applied to the field.

5. DISCUSSION

5.1 Strengths of the Hybrid Approach

The hybrid approach to geomechanical modeling that incorporates the results of laboratory tests of rock strength characteristics, with log-based estimates of stresses in the field, has certain strong operational benefits:

Although there have been significant enhancements in methodology, there are a number of limitations that should be considered:

- 1. Synthetic Data Generation:** Physically-based, synthetic data generating 847 samples might not be representative of all the natural variability that is experienced in field operations. It is necessary to validate it using real HPHT well data.
- 2. Real-Time Integration:** Existing workflow tools take the form of integrating the current workflow with the existing set of logs; to achieve an integration between same-time LWD (Logging While Drilling) data and the model, the model has to be dynamically updated, and the computationally optimal approach has to be taken.

5.2 Limitations

Despite substantial methodological improvements, several limitations warrant consideration:

- 1. Synthetic Data Generation:** While physically-based, synthetic data from 847 samples may not capture all natural variability encountered in field operations. Validation on actual HPHT well data is essential.
- 2. Geographic Specificity:** Model calibration reflects a single offshore basin. Transfer learning or retraining may be required for different geological settings.
- 3. Real-Time Integration:** Current workflow processes static logs; integration with real-time LWD (Logging While Drilling) data requires dynamic model updates and computational optimization.
- 4. Temperature Gradient Complexity:** Current stress estimations assume simplified temperature profiles; complex thermal regimes may require enhanced modeling.
- 5. Long-Term Creep:** While time-dependent effects are acknowledged, the current classification framework operates on point-in-time assessments without explicit creep rate modeling.

5.3 Practical Implications

The combination model has the direct operation advantages:

1. **Better Mud Weight Design:** Accuracy is 79.58% giving more accurate mud weight window choice, and it saves non-productive time without compromising borehole integrity.
2. **Cost Mitigation:** The model can be used to preempt mitigations (better fluid systems, minimized drilling time) in high-risk intervals with 83% sensitivity prior to the effects of instability being observed.
3. **Data-Driven Operations:** The feature importance ranking informs the drilling operations teams on key monitoring parameters, which optimize the real-time monitoring of the well

6. CONCLUSIONS

The development, application, and validation of a hybrid geomechanical modeling methodology, based on the laboratory measured triaxial rock mechanical properties and field log stress estimates to predict wellbore stability in HPHT reservoirs was created. The efficacy of the integrated approach was demonstrated by Support Vector Machine classification with a test accuracy of 79.58 and equal sensitivity (0.8333) and specificity (0.7583).

REFERENCES

1. Amanullah, M. et al., Challenges in Drilling HPHT Wells: A Review. *SPE Drilling & Completion*, 2018; 33(2): 214-224.
2. ASTM D7012, Standard Test Method for Compressive Strength and Elastic Moduli of Intact Rock Core Specimens Under Varying States of Stress and Temperatures. ASTM International, 2013.
3. Barton, C. et al., Stress Measurement Techniques in Horizontal Wells. *Journal of Petroleum Technology*, 2012; 64(12): 68-75.
4. Bazant, Z.P., Chen, S., Temperature Effects on Rock Mechanics. *Int. J. Rock Mech. Min. Sci.*, 1997; 34(3-4): 203-211.
5. Eaton, B.A., The Equation for Fracture Pressure in Porous Rock. *SPE Drilling Engineering*, 1975; 20(4): 313-317.
6. Economides, M.J., Nolte, K.G., *Reservoir Stimulation*, 3rd ed., Wiley., 2014.
7. Fjar, E. et al., *Petroleum Related Rock Mechanics*, 2nd ed., Elsevier., 2008.
8. Haimson, B.C., Herrick, C., True Triaxial Testing of Rock. *International Journal of Rock Mechanics and Mining Sciences*, 1989; 26(2): 147-162.

9. Haimson, B.C. et al., Wellbore Breakouts and Core Discing: A Comparative Study. *Rock Mechanics and Rock Engineering*, 1992; 25(1): 13-28.
10. Jaeger, J.C., Cook, N.G.W., Zimmerman, R., 2007. *Fundamentals of Rock Mechanics*, 4th ed., Wiley.
11. Kaiser, P.K. et al., Advances in Triaxial Testing and Implications for Wellbore Stability. *Rock Mechanics and Rock Engineering*, 2005; 38(2): 111-124.
12. Lee, S.W. et al., Integrated Wellbore Stability Modeling Using Core and Log Data. *SPE Journal*, 2016; 21(4): 1156-1167.
13. Mann, D.L. et al., Geomechanical Modeling for HPHT Reservoirs: Challenges and Solutions. *Journal of Petroleum Science and Engineering*, 2014; 118: 203-211.
14. Mogi, K., Frictional Coefficient Under High Pressure and Shear Stress Dependence of Rock Strength. *Journal of Geophysical Research*, 1971; 76(11): 3849-3865.
15. Yong, W. et al., Experimental Study on Creep Behavior of HPHT Reservoir Rocks. *Int. J. Rock Mech. Min. Sci.*, 2017; 95: 186-196.
16. Zhang, F. et al., Chemical Effects on Wellbore Stability in HPHT Carbonates. *International Journal of Petroleum Science and Engineering*, 2019; 177: 45-59.
17. Zoback, M.D., 2010. *Reservoir Geomechanics*. Cambridge University Press.

Appendix A: Raw Triaxial Test Data

Sample ID	Lithology	Confining Pressure (MPa)	UCS (MPa)	Cohesion (MPa)	Friction Angle (°)	Young's Modulus (GPa)	Poisson's Ratio	Tensile Strength (MPa)
S1	Sandstone	10	70	10	31	28	0.26	6
S2	Sandstone	30	85	12	33	30	0.25	7
C1	Carbonate	50	135	19	35	41	0.22	9
C2	Carbonate	70	150	21	34	39	0.21	10

Appendix B: Machine Learning Model Specifications**B.1 Preprocessing Pipeline**

- **Feature Scaling:** StandardScaler (zero mean, unit variance)
- **Class Balancing:** SMOTE (k_neighbors=5)
- **Train-Test Split:** Stratified 80-20 split
- **Cross-Validation:** 5-fold stratified

B.2 Hyperparameters**Random Forest**

- n_estimators=100, max_depth=15, min_samples_split=5, min_samples_leaf=2

Gradient Boosting

- n_estimators=150, learning_rate=0.05, max_depth=6, subsample=0.8

Support Vector Machine

- Kernel='rbf', C=100, gamma='scale', probability=True

Multilayer Perceptron

- Hidden_layer_sizes=(64, 32, 16), activation='relu', solver='adam'
- Learning_rate_init=0.001, max_iter=200, early_stopping=True

B.3 Bootstrap Uncertainty Quantification

- **Iterations:** 100
- **Method:** Random sampling with replacement from training data
- **CI Level:** 95% (using $1.96 \times$ standard deviation)