Journal of Science Research and Reviews
PRINT ISSN: 1595-9074
E-ISSN: 1595-8329
DOI: https://doi.org/10.70882/josrar.2026.v3i4.250
Homepage: https://josrar.esrgngr.org
Original Research Article

Physics-Informed Neural Networks for Sonic Log Synthesis and Petrophysical Prediction: A Leave-One-Well-Out Uncertainty Assessment

*[1]Babalola, Oludare Olukayode, [2]Odunayo Christiana Ogunleye, [1]Racheal Foluke Oloruntola, [1]Damilola Rukaya Adedokun

[1]Department of Physics, Lead City University, Ibadan, Oyo State, Nigeria.
[2]Department of Geology, Obafemi Awolowo University, Ile-Ife. Osun State, Nigeria.
*Corresponding Author's email: darekay35@gmail.com

KEYWORDS
Physics-informed neural networks,
Synthetic sonic log,
Uncertainty,
Leave-one-well-out cross-validation.

ABSTRACT
Missing logs in petrophysical analysis of Niger Delta Wells (especially Sonic DT logs) is a common problem and most of the time a regression fit or a synthetic sonic log is sometimes used to solve for this missing logs. When a sonic log is synthesized rather than measured, the uncertainty associated with that synthesis may affect subsequent porosity and permeability predictions. This study builds a three-stage physics-informed neural network (PINN) pipeline for six TMB field wells. Stage 1 predicts sonic transit time from gamma ray, resistivity, spontaneous potential, and shale volume from TMB-05 and TMB-06, constrained by a Wyllie time-average residual. Stage 2 jointly predicts porosity and permeability from a wider log suite, tied together by a Timur-Coates residual. Stage 3 tests whether ensemble spread widens when its sonic input is synthetic rather than measured. Leave-one-well-out cross-validation gave Stage 1 R^2 between -1.47 and -0.45, a result random-split validation missed entirely. In Stage 2, one-fold initially failed because an input feature fell outside the training range; injecting random-magnitude synthetic noise into the DT input during training fixed it, recovering all folds. Porosity R^2 ranged from 0.67 to 0.88, permeability R^2 from 0.78 to 0.93. A controlled sweep of declared DT uncertainty, holding all other inputs constant, showed that ensemble spread narrowed rather than widened as declared uncertainty increased: the network relied more heavily on its more robust inputs and exhibited greater internal agreement—the opposite of the behaviour a downstream user would desire. Deep-ensemble spread captures cross-well extrapolation risk effectively, but it does not propagate a known upstream uncertainty downstream. Treating ensemble variance as a proxy for input-confidence tracking can therefore be misleading.

CITATION
Babalola, O. O., Ogunleye, O. C., Oloruntola, R. F., & Adedokun, R. D. (2026). Physics-Informed Neural Networks for Sonic Log Synthesis and Petrophysical Prediction: A Leave-One-Well-Out Uncertainty Assessment. *Journal of Science Research and Reviews*, 3(4), 285-298. https://doi.org/10.70882/josrar.2026.v3i4.250

INTRODUCTION

Sonic logs are essential for porosity estimation, seismic-to-well ties, and geomechanical analysis, yet they are frequently incomplete or unavailable because of tool failures, borehole conditions, or cost constraints at the wellsite (Yu et al., 2021). The conventional remedy—regression of available logs onto measured DT—has been explored with methods ranging from multilinear regression to gradient-boosted trees and was systematically benchmarked in an SPWLA-hosted contest (Yu et al., 2021). Although predictive accuracy has improved, most published studies have not examined how the uncertainty of a synthesized log propagates into subsequent petrophysical estimates. A synthetic DT carries a confidence interval that depends on the similarity between the target well and the training wells. If that interval is ignored and the synthetic curve is treated as equivalent to a measured log, porosity and permeability estimates inherit an unquantified error source.

Physics-informed neural networks (PINNs) offer one route to mitigate this problem. Raissi et al. (2019) formalized the embedding of known physical relationships into the loss function alongside a data-fitting term. The approach has been applied in petrophysics: Pothana and Ling (2025) demonstrated that a physics-integrated network predicting mineralogy and porosity from well logs generalized more successfully to blind wells than a conventional artificial neural network. Cross-well generalization remains the central practical challenge addressed in the present work.

A complementary gap concerns uncertainty quantification. Independently trained ensembles provide a computationally inexpensive estimate of predictive uncertainty without the full machinery of Bayesian neural networks (Lakshminarayanan et al., 2017). Whether such ensembles correctly reflect uncertainty that originates upstream—an uncertain synthetic log fed into a downstream model—has received less attention. The present study examines precisely this question.

A three-stage pipeline is constructed using six wells from the TMB field (Niger Delta Basin). Stage 1 synthesizes DT under a Wyllie time-average residual (Wyllie et al., 1956). Stage 2 predicts porosity and log-permeability under a Timur–Coates residual (Timur, 1968), accepting either measured or synthetic DT according to data quality. Stage 3 tests whether Stage-2 ensemble spread responds appropriately when its DT input changes from measured to synthetic. Three observations that emerged during pipeline construction are emphasized: (i) two data-integrity problems (a duplicated curve and a clipped curve) that would have corrupted the training set were identified by cross-well audit; (ii) leave-one-well-out (LOWO) validation reveals substantially poorer cross-well generalization than random-split validation for both stages; and (iii) an ensemble explicitly supplied with upstream uncertainty as an input feature still narrowed, rather than widened, its predictions—an outcome that is mechanistically explained and has direct practical consequences.

Geological Setting and Dataset

Study area

The TMB field lies within the Niger Delta Basin, a Tertiary-to-Recent deltaic succession of prograding sand and shale sequences deposited under fluvial, deltaic, and shallow-marine conditions (Owoyemi, 2005). Reservoir rocks are typically unconsolidated to poorly consolidated sandstones interbedded with shale. This lithology renders sonic-log-based porosity estimation both useful (sonic response is sensitive to porosity in clean sand) and occasionally inaccurate (shale intervals and poor borehole conditions depress the raw DT reading) (Ubuara et al., 2024). Six wells namely TMB-01, TMB-02, TMB-04, TMB-05, TMB-06, and TMB-07 were available as LAS files, each containing a conventional log suite of varying completeness.

Curve availability and data-quality control

Table 1 summarizes curve availability after a systematic data-quality audit. TMB-05 and TMB-06 retain the full suite required for Stage 1, including trustworthy measured DT. TMB-01 lacks RHOB and NPHI and has only sparse PHIE and SW. TMB-02 lacks RHOB, NPHI, and PHIE. The DT curves of TMB-04 and TMB-07 were deemed unreliable after cross-checking: TMB-04's DT is 99.3 % identical, sample-for-sample, to TMB-02's DT across their overlapping depth range, indicating duplication; TMB-02's own DT is clipped or saturated on 43.6 % of its readings (pinned at the curve minimum or maximum). Both curves were therefore excluded from the Stage-1 ground-truth set, leaving only TMB-05 and TMB-06 as trustworthy DT wells. A third, minor issue concerned TMB-06's PHIE: 28 % of values equalled a non-standard null sentinel (150.0) and were recoded to missing; a separate near-zero cluster (15 % of readings) was retained after confirmation that those points lie in high-VSH shale intervals.

No core-derived permeability exists for any of the six wells. A Timur–Coates pseudo-label, computed from PHIE and an irreducible water saturation estimated per well via the Buckles number, is therefore used. All permeability metrics reported herein should be interpreted with this limitation in mind.

Table 1: Curve availability by well after the data-quality audit

Well	Depth samples	GR	RHOB	NPHI	DT (trustworthy)	PHIE	SW/VSH	Resistivity
TMB-01	25,423	✓	✗	✗	✗	Sparse	✓	✓
TMB-02	25,901	✓	✗	✗	✗ (clipped)	✗	✗/✓	✓
TMB-04	25,900	✓	✓	✓	✗ (duplicated)	✓	✓	✓
TMB-05	23,064	✓	✓	✓	✓	✓	✓	✓
TMB-06	46,951	✓	✓	✓	✓	✓	✓	✓
TMB-07	41,155	✓	✓	✓	✗	✓	✓	✓

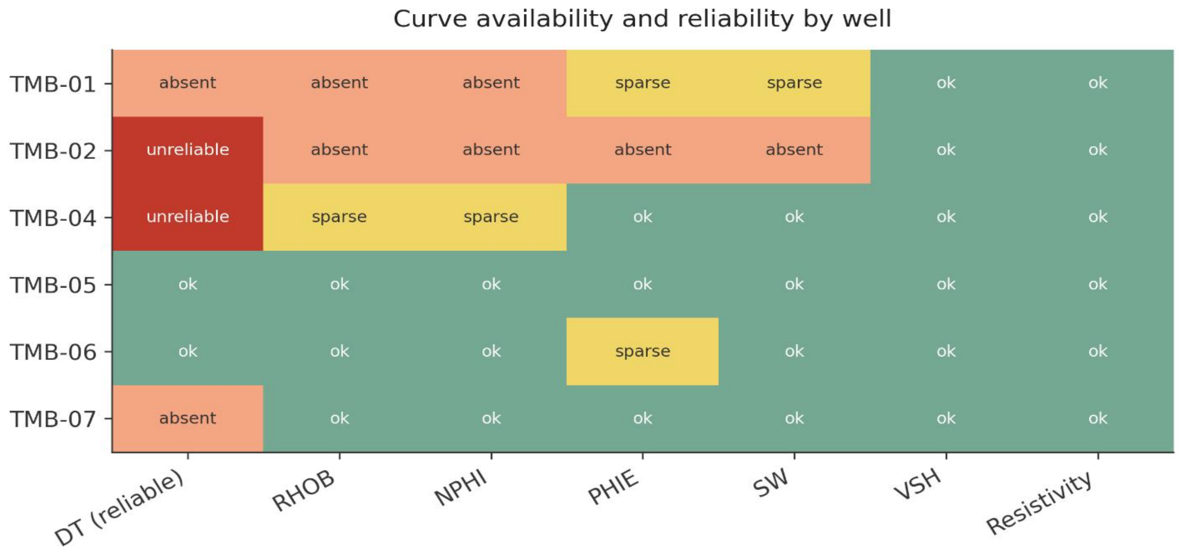


Figure 1: Curve availability by well following the data-quality audit

MATERIALS AND METHODS

Stratigraphic correlation of reservoir sands was performed across the six TMB wells using gamma-ray and resistivity logs (Figure 3). Four reservoir sands were correlated. The pipeline comprises three stages (Figure 2). Stage 1 predicts DT from GR, deep and shallow resistivity (log-transformed), spontaneous potential (SP), and shale volume (VSH), constrained by a Wyllie time-average residual. Stage 2 predicts porosity (PHIE) and log-permeability ($\log_{10}$ PERM) jointly from a wider suite that includes DT (measured where trustworthy, synthetic otherwise), constrained by a Timur–Coates residual. Stage 3 consists of diagnostics that test whether Stage-2 predictive uncertainty responds appropriately when DT changes from measured to synthetic.

Every network is a four-layer feed-forward architecture with 64 hidden units per layer, tanh activations, and dropout (rate 0.10–0.12) between layers, trained with the Adam optimizer (Kingma & Ba, 2014) at a learning rate of 2×10^{-3} for 200 epochs (Table 15). Ensembles comprise five independently initialized networks per stage; ensemble disagreement (standard deviation across member predictions) serves as the epistemic uncertainty estimate, following the deep-ensemble approach of Lakshminarayanan et al. (2017).

Stage 1: synthetic DT with a Wyllie residual

The physics constraint follows the Wyllie time-average equation (Wyllie et al., 1956):

$$\Delta t = \emptyset \times \Delta t_{\mathrm{fluid}} + (1-\emptyset) \times \Delta t_{\mathrm{matrix}} \quad (1)$$

where $\emptyset$ is porosity, and Δt_{fluid} and Δt_{matrix} are the transit times of the pore fluid and rock matrix respectively. Rather than fixing these at textbook constants, both were left as learnable scalar parameters, initialized near their conventional sandstone and mud-filtrate values (Δt_{matrix} of about 55.5 microseconds per foot, and Δt_{fluid} of about 189) and allowed to drift slightly during training due to the fact that the exact matrix mineralogy of the TMB sands is not independently constrained here. The physics loss is a masked mean-squared residual between the network's Δt prediction and the Wyllie-predicted Δt computed from that same sample's $\emptyset$ (available only for TMB-05 and TMB-06 among the ground-truth wells), weighted at five percent of the total loss relative to the data term. Training used only TMB-05 and TMB-06, the two wells with trustworthy measured DT as shown in figure 1 and every other well used a synthetic DT from the resulting five-member ensemble.

Stage 2: porosity and permeability with a Timur-Coates residual

The physics constraint here is the Timur equation (Timur, 1968):

$$k = \frac{a \times \emptyset^{m}}{S_{wirr}{}^{n}} \quad (2)$$

with a, m, and n treated as learnable parameters (log-space for a, to keep optimization stable), initialized at Timur's original sandstone coefficients (a of about 8581, m of 4.4, n of 2.0). The residual penalizes any mismatch between the network's own porosity output and its own log-permeability output, evaluated against the well's Buckles-derived S_{wirr}; a self-consistency constraint, since there is no core dataset to validate it. Inputs are GR, RHOB, NPHI, VSH, log-resistivity, DT, and, in the final version of the model, a DT-uncertainty feature described in the next section. Training and LOWO evaluation used the four wells with usable PHIE, RHOB, and NPHI coverage: TMB-04, TMB-05, TMB-06, and TMB-07 since they have almost the complete dataset.

Propagating Stage-1 uncertainty into Stage 2: three iterations

Getting Stage 2 to respond sensibly to Stage-1's uncertainty took three attempts, and all three are described here.The first version fed Stage 2 nothing about DT confidence at all; just the DT value itself, measured or synthetic, with no distinction between the two. The second version added a DT_std_in feature: the ensemble standard deviation from Stage 1, near-zero (0.05 microseconds per foot, a nominal floor) for the two measured-DT wells, and the genuine Stage-1 ensemble spread for the synthetic-DT wells. This improved three of the four LOWO folds noticeably, but the fourth, TMB-07, held out, collapsed to near-zero R2. The reason turned out to be simple once traced: TMB-07's true DT_std_in values (mean 2.36 microseconds per foot, up to 4.12) exceeded anything the training wells showed when TMB-07 was the one held out, so the network was extrapolating on that one feature into territory it had never seen ;a clean illustration of why LOWO, not a random split, is the validation scheme that actually catches this kind of thing - a random split would have mixed TMB-07's extreme values into the training set and hidden the problem entirely.

The third and final version fixes this directly by making a necessity for every training epoch to regenerates DT_in by adding fresh Gaussian noise of a randomly drawn magnitude (≥ 0.05 to 6.0 microseconds per foot), a range chosen putting into consideration the largest real uncertainty seen anywhere in the dataset; around each sample's clean DT value, with DT_std_in set to match the noise actually injected instead of a fixed DT_std_in tied to well identity. This declared uncertainty from well identity totally: every well, during training, sees the full range of possible confidence levels, so no held-out well's real uncertainty can fall outside what the network has encountered before. All Stage 2 results use this final, augmented-training version unless explicitly marked otherwise.

Cross-validation strategy

Every reported metric uses leave-one-well-out (LOWO) cross-validation: for each fold, one well is withheld entirely and the model trains on the rest, then predicts on the withheld well. This is a considerably more difficult in computation and give a more honest test than a random row-wise split, which allows a model to see some depths from every well during training and consequently overstates how well it would generalize to a genuinely new well. A random-split baseline is reported alongside LOWO throughout, specifically to make that gap visible rather than to serve as the headline number.

Uncertainty diagnostics

Three separate diagnostics probe whether Stage 2's ensemble spread reflects Stage-1's uncertainty. The first substitutes each measured-DT well's real DT with Stage-1's synthetic DT at the same depths, Monte Carlo sampling twelve draws from Stage-1's predictive distribution per sample, and compares the resulting Stage-2 ensemble spread to the spread obtained using the real, measured DT. The second is a controlled sweep: using real samples from the two measured-DT wells, DT_in and every other feature are held fixed at their true values while DT_std_in alone is swept manually from 0.05 to 6.0, isolating whether the network responds causally to declared uncertainty at all, independent of what Stage 1 actually estimated for that depth. The third is simply a cross-well comparison of Stage-2's own ensemble spread against each well's DT source, measured or synthetic.

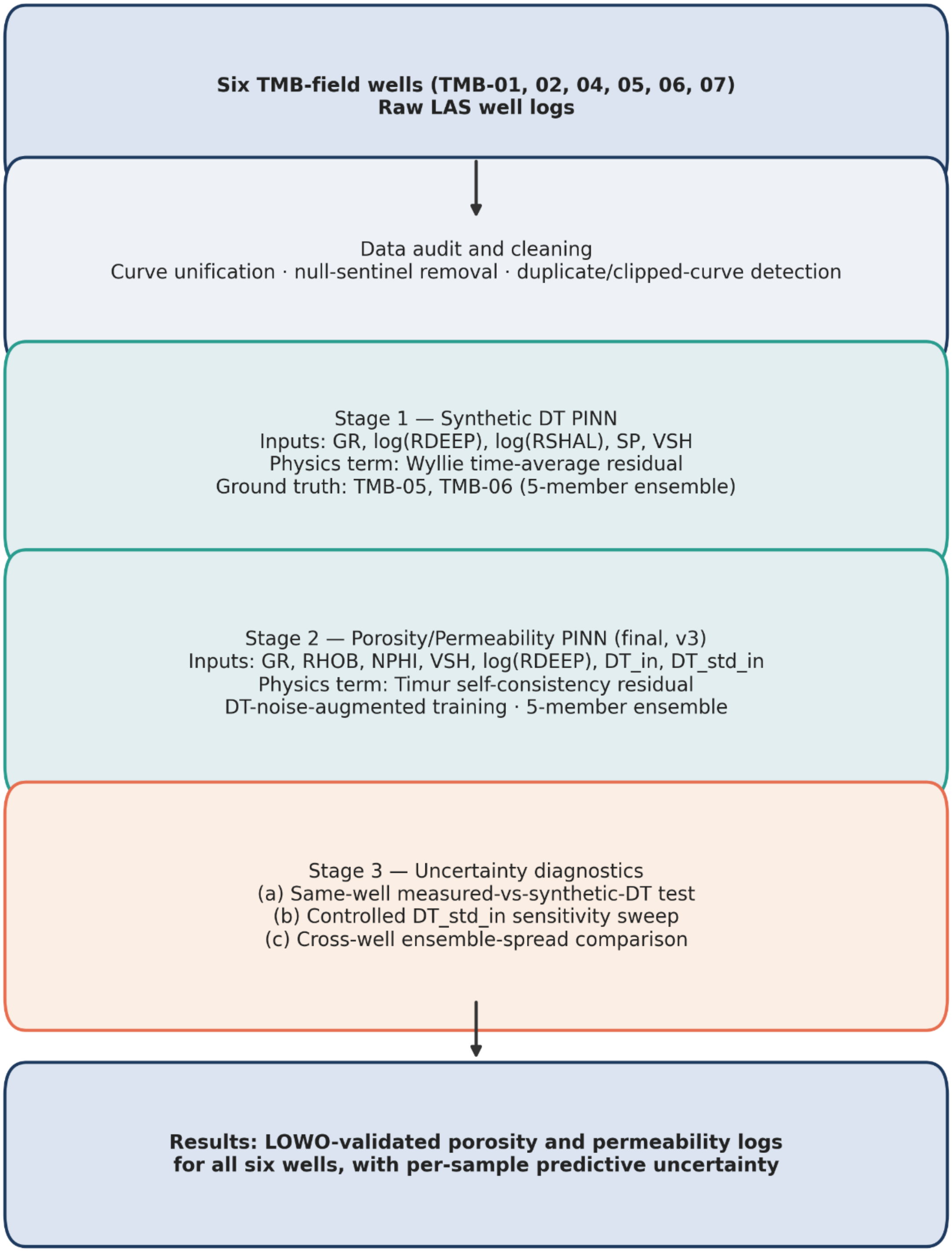


Figure 2: Methodology flowchart: synthetic DT generation (Stage 1), porosity-permeability prediction (Stage 2), and uncertainty diagnostics (Stage 3)

JOSRAR 3(4) JUL-AUG 2026 285-298

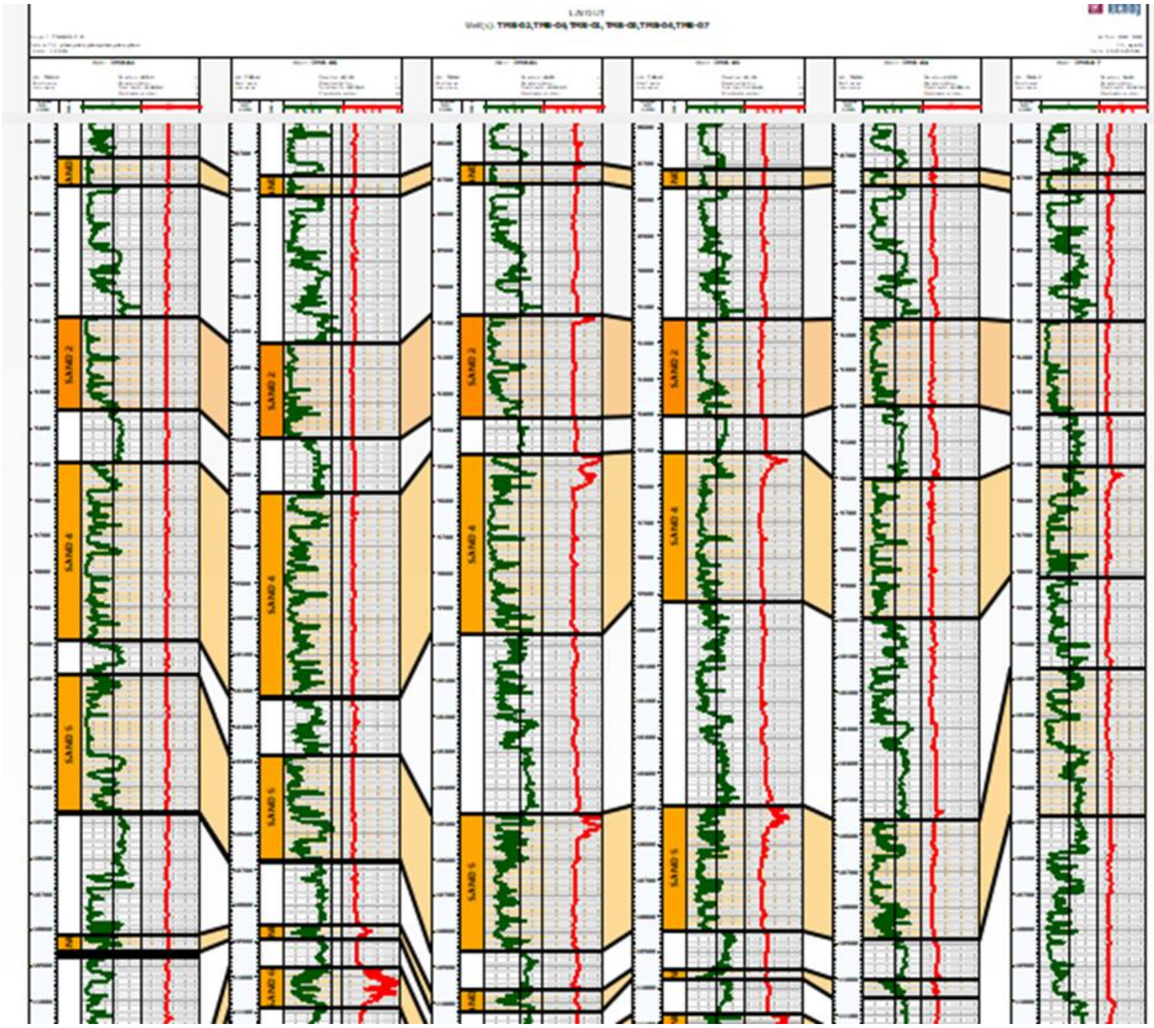


Figure 3: Stratigraphic correlation across the six TMB well using GR and resistivity log

RESULTS AND DISCUSSION

Stage 1: synthetic DT generalizes poorly across wells

Leave-one-well-out cross-validation between TMB-05 and TMB-06 produced MAE values of 7.52 μs/ft (TMB-05 held out) and 14.41 μs/ft (TMB-06 held out), against 5.06 μs/ft for a random 80/20 split pooling both wells (Table 2). R^2 under LOWO was negative in both folds (−1.47 and −0.45). A random split therefore overstated generalization performance by a substantial margin. With only GR, resistivity, SP, and VSH available as common features, the model lacks sufficient information to separate genuine cross-well geological variation from other sources of discrepancy. Two wells constitute a limited ground-truth base; the results demonstrate the LOWO-versus-random-split gap but do not support broad claims about typical cross-well DT variability across the basin.

Nevertheless, the ensemble's own uncertainty estimate behaved as expected for an epistemic proxy (Figure 4). Mean predictive standard deviation on the two Stage-1 training wells was 0.39–0.48 μs/ft; on wells never seen during training it rose four- to six-fold (1.70–2.36 μs/ft). Stage 1 therefore knows when it is extrapolating even though its absolute accuracy remains limited.

Table 2: Stage 1 leave-one-well-out cross-validation, TMB-05 and TMB-06

Held-out well	N	MAE (μs/ft)	RMSE (μs/ft)	R^2
TMB-05	4,232	7.52	9.43	-1.47
TMB-06	15,000*	14.41	18.30	-0.45
Random-split baseline	3,846	5.06	7.97	—

NB: *TMB-06 subsampled to 15,000 rows for training-time efficiency

Figure 4a: Stage 1 ensemble predictive uncertainty (mean std across five members), Stage-1 training wells vs. wells without trustworthy measured DT

Stage 2: a feature-extrapolation failure, diagnosed and fixed

The first version of Stage 2, with no DT-uncertainty information at all, gave LOWO porosity R^2 between 0.37 (TMB-04) and 0.75 (TMB-05), and permeability R^2 between 0.63 and 0.88. Adding DT_std_in as a fixed per-well feature (the second version) improved three of the four folds - TMB-04's porosity R^2 rose to 0.60, its permeability R^2 to 0.72 - but TMB-07 collapsed to a porosity R^2 of 0.008 and a permeability R^2 of 0.098, an order of magnitude worse than before the feature was added.

TMB-07's true DT_std_in range fell outside anything the training wells exhibited during that fold, and the network extrapolated badly on a single feature. The DT-noise-augmented third version resolves it (Table 3, Figure 4). Every fold now falls in a tight, sensible band: porosity R^2 of 0.68 to 0.88, permeability R^2 of 0.78 to 0.93. TMB-07 specifically moved from 0.008 to 0.671 (porosity) and from 0.098 to 0.924 (permeability) - not just recovered, but now one of the two strongest folds in the whole table. TMB-05 improved as well, from 0.82 to 0.88. TMB-04 and TMB-06 held roughly steady. The random-split baseline (porosity MAE 0.034, against a LOWO range of 0.028 to 0.052) again sits comfortably ahead of every LOWO fold, the same pattern as Stage 1 and the same caution about what a random split hides.

Table 3: Stage 2 leave-one-well-out cross-validation, final DT-noise-augmented model

Held-out well	n	PHIE MAE	PHIE R^2	$\log_{10}$(PERM) MAE	$\log_{10}$(PERM) R^2
TMB-04	2,021	0.047	0.716	0.616	0.817
TMB-05	4,111	0.028	0.879	0.503	0.785
TMB-06	15,000*	0.049	0.679	0.425	0.927
TMB-07	15,000*	0.052	0.671	0.296	0.924
Random-split baseline	7,226	0.034	—	0.315	—

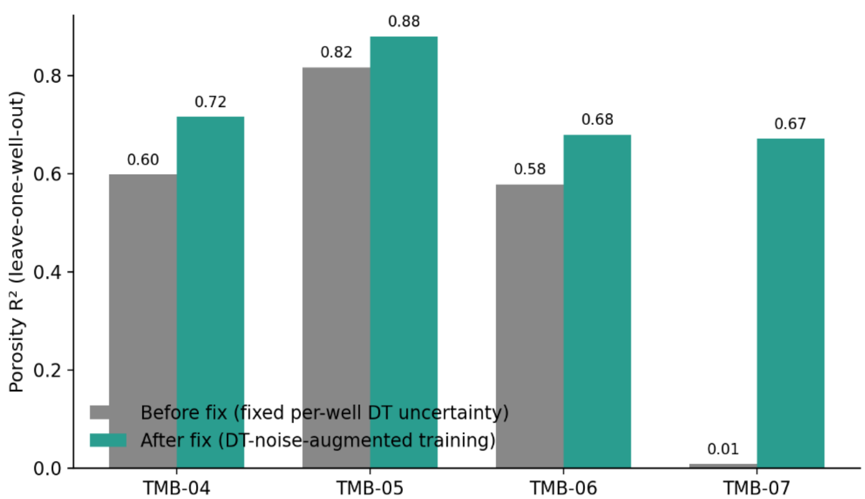


Figure 4b: Stage 2 leave-one-well-out porosity R^2 before and after the DT-noise-augmentation fix

Figure 5 shows out-of-fold predicted-versus-measured cross-plots for the final ensemble, pooled across all four ground-truth wells. Porosity predictions cluster tightly along the 1:1 line below about 0.25 and show more scatter above it, where TMB-06's higher-porosity intervals sit; permeability predictions - against the Timur-Coates pseudo-label, not a core measurement - track the 1:1 line reasonably well across four log10 orders of magnitude, which is more a statement about how strongly the physics residual constrains the two outputs together than about independent permeability accuracy.

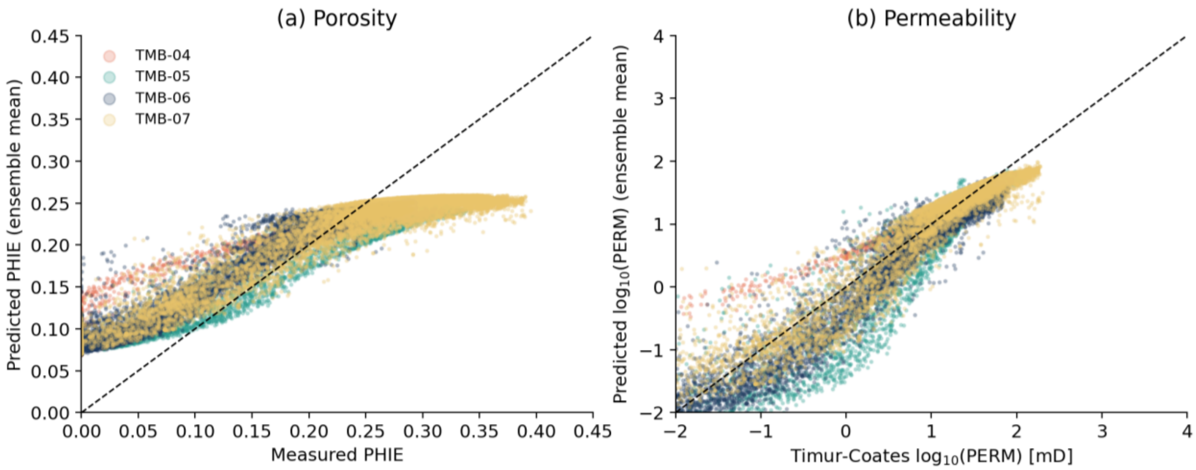


Figure 5: Measured vs. predicted porosity (a) and $\log_{10}$(permeability) (b), final ensemble, out-of-fold LOWO predictions across all four ground-truth wells

Timur parameters converged to values close to the coefficients Timur (1968) originally reported for sandstones across every LOWO fold: a between 7,067 and 7,915 (against Timur's 8,581), m between 4.56 and 4.59 (against 4.4), and n between 2.07 and 2.19 (against 2.0). That the network, free to move these parameters wherever the data pulled them, settled close to seventy-year-old core-calibrated values is a reasonable internal consistency check on the physics constraint, independent of anything to do with the uncertainty question.

Does Stage-2 uncertainty widen when its input is synthetic?

Same-well substitution of synthetic for measured DT at identical depths produced essentially no widening of Stage-2 ensemble spread (Table 4). Porosity standard-deviation ratios were 0.98 in both TMB-05 and TMB-06; permeability ratios were 0.93–0.95. Because these wells were Stage 1's training wells, their Stage-1 uncertainty is already small. A controlled sweep that held all features fixed while varying only DT_std_in from 0.05 to 6.0 μs/ft removed that objection. Ensemble spread narrowed systematically: porosity narrowing ratios reached 0.91–0.92 and permeability ratios 0.86–0.92 (Figure 6, Table 14). The network therefore responds causally to declared uncertainty, but by discounting the unreliable input and relying on the remaining consistent features, thereby increasing rather than decreasing agreement among ensemble members.

Cross-well comparison confirmed the absence of a simple monotonic relationship between input uncertainty and output spread: TMB-07, which carries the highest real DT uncertainty (mean 2.36 μs/ft), exhibited a lower permeability predictive spread than TMB-05, which has measured, high-confidence DT.

Table 4: Same-well diagnostic: Stage-2 ensemble predictive standard deviation, measured DT vs. Monte Carlo-propagated synthetic DT, at identical depths

Well	N	PHIE std (measured)	PHIE std (synthetic)	Ratio	$\log_{10}$(PERM) std (measured)	$\log_{10}$(PERM) std (synthetic)	Ratio
TMB-05	4,111	0.01376	0.01344	0.98	0.0953	0.0890	0.93
TMB-06	8,000*	0.01192	0.01170	0.98	0.0917	0.0875	0.95

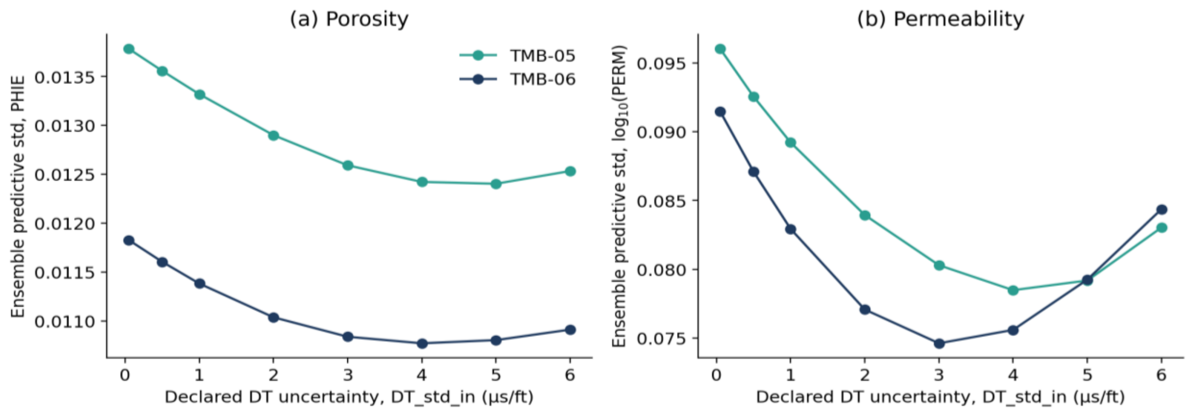


Figure 6: Controlled sensitivity sweep

Discussion

An ensemble that is explicitly informed, via an input feature, of the unreliability of its sonic log does not widen its predictions; it narrows them. This behaviour is rational at the level of a single prediction—discounting an unreliable input and relying on the remaining consistent features increases agreement among ensemble members—yet it is precisely the opposite of the behaviour required if ensemble spread is to serve as a downstream confidence interval on porosity or permeability. A practitioner who treated Stage-2 predictive variance as a confidence interval would therefore be most confident precisely where the input data were least reliable.

Two distinct forms of uncertainty are easily conflated. Cross-well epistemic uncertainty—whether the model recognizes that it has never seen a well of the present type—is captured effectively by deep ensembles, as evidenced by the four- to six-fold increase in Stage-1 predictive standard deviation on out-of-training wells. Sequential uncertainty propagation—whether a downstream model correctly reflects an upstream prediction it has been told is unreliable—is a different mechanism and is not learned by training an ensemble on cross-well variation alone.

Closing the sequential-propagation gap would require a loss function that explicitly penalizes over-confidence when DT_std_in is high, for example an input-dependent negative-log-likelihood objective. That experiment has not been performed here and is left as future work. Two further limitations are acknowledged: permeability metrics are validated only against a Timur–Coates pseudo-label, not core measurements; and Stage 1 rests on only two ground-truth wells, sufficient to demonstrate the LOWO gap but insufficient for broad generalization claims about basin-wide DT variability.

CONCLUSION

A three-stage physics-informed pipeline was constructed for six TMB-field wells. Data-quality problems (duplicated and clipped DT curves) were identified and removed before training. Physics residuals (Wyllie for sonic synthesis, Timur–Coates for porosity–permeability) kept both stages consistent with established rock-physics relationships; the Timur parameters recovered by the network converged close to the original 1968 sandstone coefficients. Leave-one-well-out validation demonstrated that cross-well generalization is substantially harder than random-split validation suggests, at both stages of the pipeline. A feature-extrapolation failure in Stage 2 was diagnosed and corrected by training-time noise augmentation.

The central finding is negative with respect to a common assumption: deep-ensemble spread, under the architecture examined here, does not widen in response to a known, explicitly supplied upstream uncertainty; it narrows, because the network rationally discounts an unreliable input in favour of reliable ones. Consequently, Stage-2 ensemble variance should not be interpreted as a proxy for the reliability of the DT input. Practitioners deploying sequential PINN workflows that rely on ensemble spread to flag upstream uncertainity are advised to test that assumption directly rather than to take it for granted.

REFERENCES

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv. https://doi.org/10.48550/arXiv.1412.6980

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. Advances in Neural Information Processing Systems, 30, 6402–6413.

Owoyemi, A. O. (2005). Sequence stratigraphy of Niger Delta, Delta field, offshore Nigeria (Doctoral dissertation, Texas A&M University).

Pothana, P., & Ling, K. (2025). Physics-integrated neural networks for improved mineral volumes and porosity estimation from geophysical well logs. Energy

Geoscience, 6(2), 100410. https://doi.org/10.1016/j.engeos.2025.100410

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Timur, A. (1968). An investigation of permeability, porosity, and residual water saturation relationships for sandstone reservoirs. The Log Analyst, 9(4), 8–17.

Ubuara, D. O., Olayinka, Y. A., Emujakporue, G. O., & Soronnadi-Ononiwu, G. C. (2024). Evaluation of formation susceptibility and sand production potential in an offshore field, Niger Delta Basin, Nigeria. Energy Geoscience, 5(1), 100213. https://doi.org/10.1016/j.engeos.2023.100213

Wyllie, M. R. J., Gregory, A. R., & Gardner, L. W. (1956). Elastic wave velocities in heterogeneous and porous media. Geophysics, 21(1), 41–70. https://doi.org/10.1190/1.1438217

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., & Izadi, H. (2021). Synthetic sonic log generation with machine learning: A contest summary from five methods. Petrophysics, 62(4), 393–406. https://doi.org/10.30632/PJV62N4-2021a4

APPENDIX

Well	Zones	Flag Name	Top	Bottom	Reference unit	Gross	Net	Net to Gross	Av_Shale Volume	Av_Porosity	Av_Water Saturation	contact (TVDSS)	Fluid type	contact type
TMB-04	SAND 6	RES	11144.21	11225.73	FT	81.511	57.14	0.701	0.258	0.249	0.546			
TMB-04	SAND 6	PAY	11144.21	11225.73	FT	81.511	32.864	0.403	0.292	0.257	0.316	11066.30	Oil	OWC
TMB-05	SAND 4	RES	9502.292	9999.79	FT	497.498	458.29	0.921	0.121	0.279	0.944			
TMB-05	SAND 4	PAY	9502.292	9999.79	FT	497.498	38.5	0.077	0.335	0.307	0.45	9537.61	Oil	OWC
TMB-05	SAND 5	RES	10490.33	10841.67	FT	351.341	323.169	0.92	0.126	0.249	0.904			
TMB-05	SAND 5	PAY	10490.33	10841.67	FT	351.341	51.5	0.147	0.176	0.232	0.39	10553.30	Oil	OWC
TMB-05	SAND 6	RES	11074.96	11205.94	FT	130.985	109.5	0.836	0.187	0.23	0.212			
TMB-05	SAND 6	PAY	11074.96	11205.94	FT	130.985	109.5	0.836	0.187	0.23	0.212	11192.80	Oil	OWC
TMB-06	SAND 7	RES	10857	10896.35	FT	39.35	15.667	0.398	0.33	0.286	0.299			
TMB-06	SAND 7	PAY	10857	10896.35	FT	39.35	15.667	0.398	0.33	0.286	0.299	10878.00	Oil	OWC
TMB-06	SAND 6	RES	10972.24	11088.59	FT	116.358	80	0.688	0.291	0.258	0.111			
TMB-06	SAND 6	PAY	10972.24	11088.59	FT	116.358	80	0.688	0.291	0.258	0.111	11068.70	Oil	ODT

Figure A1: Petrophysical Summary of the TMB fields

Table A1: Well inventory and depth coverage for the six TMB-field wells used in this study

Well	Depth top (ft)	Depth base (ft)	Sample interval (ft)	N samples
TMB-01	308.0	13019.0	0.5	25423
TMB-02	46.0	12996.0	0.5	25901
TMB-04	51.2	13000.7	0.5	25900
TMB-05	10.0	11541.5	0.5	23064
TMB-06	3849.5	11674.5	0.5	46951
TMB-07	2800.0	13088.5	0.5	41155

Table A2: Curve availability and reliability by well, established during the pre-modelling data audit

Well	DT (sonic)	RHOB	NPHI	PHIE	SW	VSH	Resistivity
TMB-01	absent	absent	absent	sparse (3.9%)	sparse (10.1%)	present	1 curve (LLD)
TMB-02	present, unreliable (43.6% clipped)	absent	absent	absent	absent	present	2 curves (LLS, RES)
TMB-04	present, unreliable (99.3% duplicate of TMB-02)	sparse (7.8%)	sparse (7.8%)	present	present	present	3 curves (LLS, RES, MSFL)
TMB-05	present, reliable	present	present	present	present	present	2 curves (LLD, LLS)
TMB-06	present, reliable	present	present	present, partly corrupted (28% null sentinel, cleaned)	present	present	3 curves (LLD, LLS, RT)
TMB-07	absent	present	present	present	present	present	3 curves (HDRS, HMRS, RES)

Table A3: Data integrity issues identified during the audit, with the resolution applied before model training

Issue	Description	Resolution
TMB-04 DTU curve duplicated from TMB-02	99.3% of TMB-04's DTU values are sample-for-sample identical to TMB-02's DT curve	Excluded from Stage-1 ground truth; treated as absent
TMB-02 DT curve clipped/saturated	43.6% of values pinned at exactly the curve's min (92.317) or max (151.3237)	Excluded from Stage-1 ground truth; treated as unreliable
TMB-06 PHIE non-standard null sentinel	28% of values equal exactly 150.0, distinct from the file-declared NULL of -999.25	Recoded to missing (NaN)
TMB-06 PHIE near-zero cluster	15% of values equal exactly 0.0; confirmed genuine (mean VSH ~0.97, i.e. shale) via cross-check against VSH and RHOB, not padding	Retained as valid zero-porosity shale
TMB-05 NPHI implausible negative values	0.9% of values below -1, physically impossible for a neutron porosity fraction	Recoded to missing (NaN)
Declared file NULL sentinel	-999.25, applied uniformly across all six wells	Recoded to missing (NaN) across all curves

Table A4: Stage-1 (synthetic DT) leave-one-well-out cross-validation metrics, with the random-split baseline for comparison

Held-out well	N samples	MAE (μs/ft)	RMSE (μs/ft)	R^2
TMB-05	4232	7.5173	9.4255	-1.4655
TMB-06	15000	14.4051	18.2989	-0.4527
RANDOM_SPLIT_BASELINE	3846	5.0599	7.9671	

Table A5: Stage-1 ensemble predictive uncertainty by well, contrasting wells inside vs. outside the DT training set

Well	N samples	Mean predictive std (μs/ft)	Median predictive std (μs/ft)	In Stage-1 training set
TMB-06	46948	0.3883	0.324	True
TMB-05	4232	0.4755	0.3997	True
TMB-04	10413	1.6984	1.864	False
TMB-01	23894	2.0801	2.4398	False
TMB-02	22101	2.2603	2.2886	False
TMB-07	40617	2.3634	2.6577	False

Table A6: Stage-2 leave-one-well-out metrics, v2 (fixed per-well DT_std_in feature, pre-fix)

Held-out well	N samples	PHIE MAE	PHIE R^2	log10(PERM) MAE	log10(PERM) R^2
TMB-04	2021	0.0481	0.7012	0.5515	0.8435
TMB-05	4111	0.0342	0.8205	0.4848	0.8148
TMB-06	15000	0.0421	0.7729	0.3816	0.9309
TMB-07	15000	0.0942	0.0081	1.3503	0.0976
RANDOM_SPLIT_BASELINE	7226	0.0325		0.2909	

Table A7: Stage-2 leave-one-well-out metrics, v3 (final, DT-noise-augmented training)

Held-out well	N samples	PHIE MAE	PHIE R^2	log10(PERM) MAE	log10(PERM) R^2
TMB-04	2021	0.047	0.7162	0.6159	0.8167
TMB-05	4111	0.0277	0.8795	0.5032	0.7847
TMB-06	15000	0.0495	0.6794	0.4251	0.9268
TMB-07	15000	0.0525	0.6715	0.2956	0.9239
RANDOM_SPLIT_BASELINE	7226	0.0335		0.3151	

Table A8: Direct comparison of Stage-2 v2 vs. v3 LOWO R^2 by well, showing the TMB-07 extrapolation fix

Held-out well	PHIE R^2 (v2)	PHIE R^2 (v3, final)	log10(PERM) R^2 (v2)	log10(PERM) R^2 (v3, final)
TMB-04	0.7012	0.7162	0.8435	0.8167
TMB-05	0.8205	0.8795	0.8148	0.7847
TMB-06	0.7729	0.6794	0.9309	0.9268
TMB-07	0.0081	0.6715	0.0976	0.9239

Table A9: Learned Timur-Coates physics parameters across LOWO folds, compared to the literature default

Held-out well	a	m	n
TMB-04	7853	4.4	1.97
TMB-05	7826	4.4	1.98
TMB-06	7915	4.4	2.02
TMB-07	7748	4.4	2.01
Literature default (Timur, 1968)	8581	4.4	2.0

Table A10: Cross-well Stage-2 ensemble predictive spread, final (v3) ensemble, by DT source

Well	DT source	N samples	Mean PHIE ensemble std	Mean log10(PERM) ensemble std
TMB-04	synthetic	2021	0.0112	0.0999
TMB-05	measured	4111	0.0138	0.0953
TMB-06	measured	15000	0.0119	0.091
TMB-07	synthetic	15000	0.0118	0.0649

Table A11: Stage-3 same-well measured-vs-synthetic-DT diagnostic, pre-fix (v1/v2) ensemble.

Well	N samples	PHIE std (measured DT)	PHIE std (synthetic DT)	log10(PERM) std (measured DT)	log10(PERM) std (synthetic DT)	PHIE widening ratio	log10(PERM) widening ratio
TMB-05	4111	0.0121	0.0115	0.0729	0.0704	0.9496	0.9653
TMB-06	8000	0.0108	0.0106	0.0818	0.0805	0.9782	0.9834

Table A12: ***Stage-3 same-well measured-vs-synthetic-DT diagnostic, re-run against the final (v3) ensemble***

Well	N samples	PHIE std (measured DT)	PHIE std (synthetic DT)	log10(PERM) std (measured DT)	log10(PERM) std (synthetic DT)	PHIE widening ratio	log10(PERM) widening ratio
TMB-05	4111	0.0138	0.0134	0.0953	0.089	0.9765	0.9337
TMB-06	8000	0.0119	0.0117	0.0917	0.0875	0.9822	0.9544

Table A13: Controlled DT_std_in sensitivity sweep: full results across the declared-uncertainty range

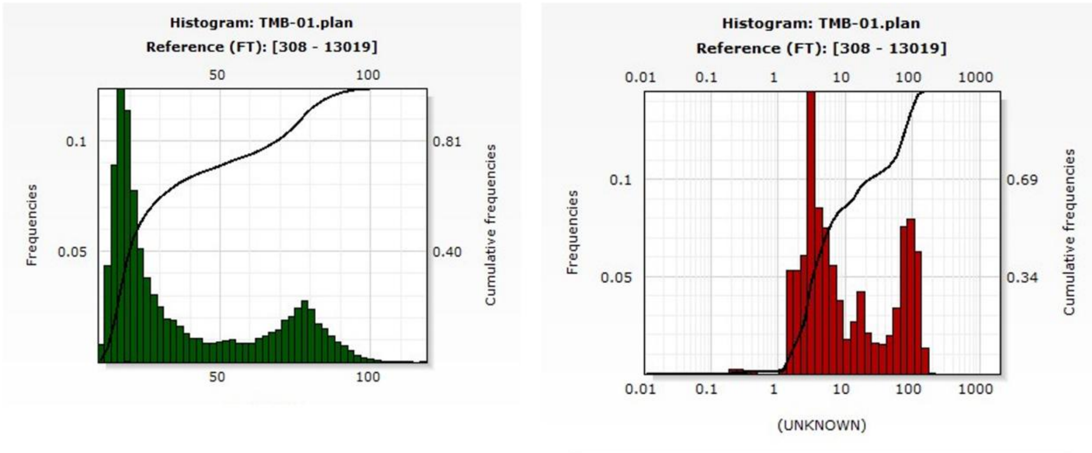
Well	Declared DT_std_in (µs/ft)	PHIE predicted mean	PHIE ensemble std	log10(PERM) predicted mean	log10(PERM) ensemble std
TMB-05	0.05	0.1757	0.0138	0.1416	0.096
TMB-05	0.5	0.176	0.0136	0.144	0.0926
TMB-05	1.0	0.1763	0.0133	0.1467	0.0892
TMB-05	2.0	0.1769	0.0129	0.1519	0.0839
TMB-05	3.0	0.1775	0.0126	0.1566	0.0803
TMB-05	4.0	0.1779	0.0124	0.1604	0.0785
TMB-05	5.0	0.1781	0.0124	0.1632	0.0792
TMB-05	6.0	0.1782	0.0125	0.1646	0.083
TMB-06	0.05	0.1711	0.0118	-0.2079	0.0915
TMB-06	0.5	0.1714	0.0116	-0.2082	0.0871
TMB-06	1.0	0.1718	0.0114	-0.2081	0.0829
TMB-06	2.0	0.1725	0.011	-0.207	0.0771
TMB-06	3.0	0.1731	0.0108	-0.2049	0.0746
TMB-06	4.0	0.1734	0.0108	-0.2019	0.0756
TMB-06	5.0	0.1734	0.0108	-0.1981	0.0792
TMB-06	6.0	0.1732	0.0109	-0.1938	0.0844

Table A14: Sensitivity-sweep widening ratio summary (declared uncertainty 0.05 to 6.0 µs/ft)

Well	PHIE std @ 0.05	PHIE std @ 6.0	PHIE ratio (6.0/0.05)	log10(PERM) std @ 0.05	log10(PERM) std @ 6.0	log10(PERM) ratio (6.0/0.05)
TMB-05	0.0138	0.0125	0.9093	0.096	0.083	0.8646
TMB-06	0.0118	0.0109	0.9226	0.0915	0.0844	0.9223

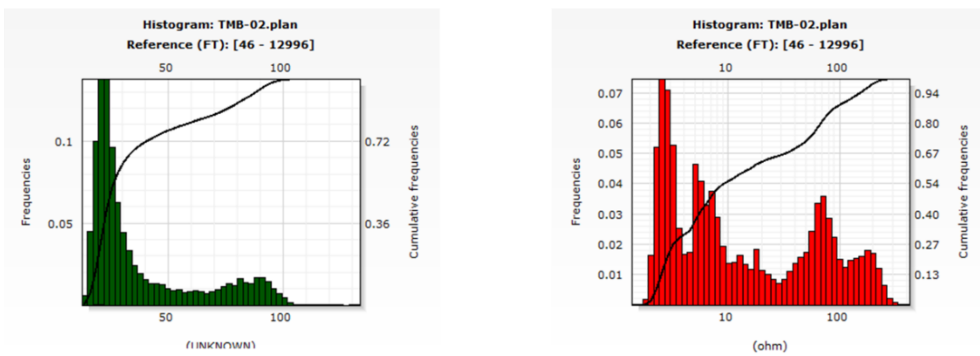
Table A15: Model architecture and training hyperparameters

Parameter	Value
Network type	Fully connected feed-forward PINN
Hidden layers	4
Hidden width	64
Activation	Tanh
Dropout rate	0.10 (Stage 1), 0.12 (Stage 2)
Optimizer	Adam
Learning rate	2×10^{-3}
Epochs	200
Ensemble size	5
Stage-1 physics term	Wyllie time-average residual (learnable DT_matrix, DT_fluid)
Stage-2 physics term	Timur self-consistency residual (learnable a, m, n)
Stage-2 DT noise range (v3)	Uniform(0.05, 6.0) µs/ft, redrawn every epoch
Row cap per well	15,000



DATA QC FOR TMB-01

Figure A2: Data QC for TMB-01



DATA QC FOR TMB-02

Figure A2: Data QC for TMB-02