

Article

Leakage-Safe Probe-Assisted Contact Angle Prediction Using Nonnegative Surface-Energy Summaries and Physics-Residual Learning

Yuying Xia ¹, Wenbin Liu ^{2,*}, Mingyang Shen ¹, Rui Xing ¹ and Xuyang Gao ¹

¹ College of Computer Science and Technology, Jilin University, Changchun 130012, China; xiayy2123@mails.jlu.edu.cn (Y.X.); shenmy2123@mails.jlu.edu.cn (M.S.); xingrui2123@mails.jlu.edu.cn (R.X.); gaoxy2123@mails.jlu.edu.cn (X.G.)

² College of Software, Jilin University, Changchun 130012, China

* Correspondence: liuwenbin@jlu.edu.cn

Featured Application

The method supports target-clean screening of coatings, polymer films, membranes and other functional surfaces when non-target probe-liquid measurements are available.

Abstract

Contact-angle prediction from literature data is vulnerable to target leakage when solid surface-free-energy descriptors are reconstructed using the liquid, which is later treated as the target. We developed a target-masked workflow that removes the target liquid before fitting nonnegative Owens-Wendt-Rabel-Kaelble components by nonnegative least squares and uses that physical prediction to anchor residual learning. A row-level revision audit re-extracted or excluded mismatched legacy sources before all models were retrained. Development used nested source-group cross-validation; the fixed cross-source external confirmation set was excluded from selection. The revised residual model achieved mean absolute errors of 15.5 degrees in nested validation and 13.2 degrees on that confirmation set. Surface-cluster bootstrap supported improvement over physics, whereas source-cluster uncertainty remained substantial. Diagnostics showed source dependence, sparse roughness, and limited strict unseen-liquid support. The method is therefore positioned as an auditable, risk-aware and reproducible materials-screening tool for surfaces with at least two non-target probes, with explicit out-of-distribution risk and refusal conditions rather than universal transfer claims.

Keywords: contact angle; wettability; surface free energy; Owens-Wendt-Rabel-Kaelble; nonnegative least squares; target masking; residual learning; source-disjoint validation; conformal prediction



Academic Editor: Andrea Atrei

Received: 13 July 2026

Revised: 27 July 2026

Accepted: 29 July 2026

Published: 4 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Contact angle links interfacial thermodynamics with the observable wetting state of coatings, polymer films, membranes, biological interfaces and microfluidic materials. Young's relation, surface-energy decompositions, and the Wenzel and Cassie-Baxter descriptions remain the main physical starting points, but real literature measurements vary with roughness, heterogeneity, liquid history and protocol [1–10].

Recent machine-learning studies demonstrate strong prediction inside controlled domains. Peta used artificial neural networks to connect electrical-discharge-machining

variables and surface topography to wettability, reporting R^2 values above 0.9 [11]. Fu et al. combined micro/nanoscale morphology and molecular-scale chemistry descriptors with extreme gradient boosting (XGBoost) and SHapley Additive exPlanations (SHAP) for laser-chemical superwetting surfaces [12]. Sena et al. used experimental and molecular descriptors to predict water contact angle across a compact polymer collection [13]. These studies show what becomes possible when process, morphology and chemistry are measured consistently.

Other work addresses different prediction targets. Shaw et al. learned to measure a hydrophobic angle from droplet contours generated under Young-Laplace physics [14]. Reservoir studies use mineralogy, pressure, temperature, salinity and gas identity to predict advancing or receding angles within shale or carbonate domains [15–17]. Their high within-domain accuracy is not directly comparable with publication-disjoint transfer across unrelated functional surfaces because the available features, targets and grouping units differ.

Physics-based and physics-informed alternatives remain important. A recent perturbed-chain statistical associating fluid theory (PCP-SAFT) formulation estimated advancing and receding angles for smooth solids and pure or mixed liquids [18], while physics-informed autoencoders constrained droplet-profile analysis by the Young-Laplace equation [19]. Empirical physics features have also improved ensemble wettability models in reservoir tasks [20]. These approaches motivate the use of an interpretable physical anchor when heterogeneous training data are too small for an unconstrained learner.

A separate problem arises when solid surface-free-energy descriptors are reconstructed from probe contact angles. If the liquid later treated as the prediction target contributed to that reconstruction, the descriptor contains target information. A source-disjoint split, defined here as a split in which no publication source appears in more than one fold, does not remove this within-surface leakage. Surface-disjoint means that no source-conditioned surface instance, defined by source, material, treatment and reported surface state, crosses a split.

We therefore study target-masked probe-assisted prediction. Every target-liquid measurement is removed before fitting nonnegative surface-energy components. The resulting physical angle and fit diagnostics anchor compact residual models, and all model selection uses source-group out-of-fold predictions. The revision adds row-level source auditing, source-learning curves, new-material and new-liquid strata, strict leave-one-liquid-out diagnostics, practical error tolerances, and explicit refusal rules (Figure 1).

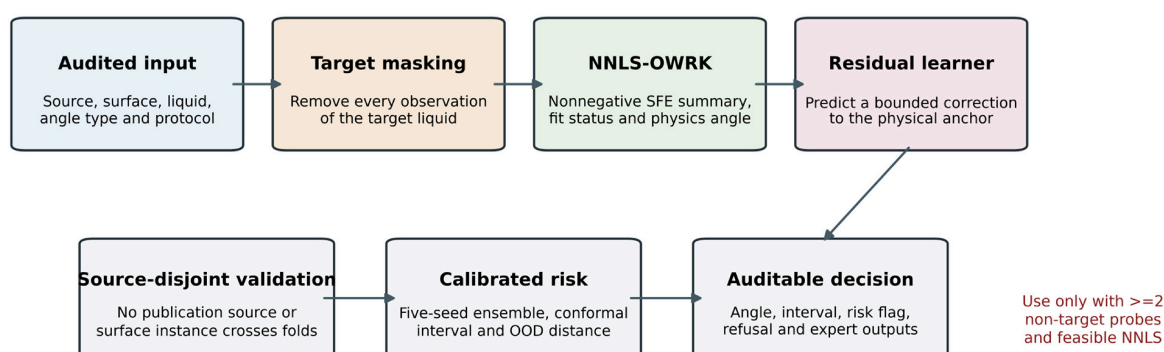


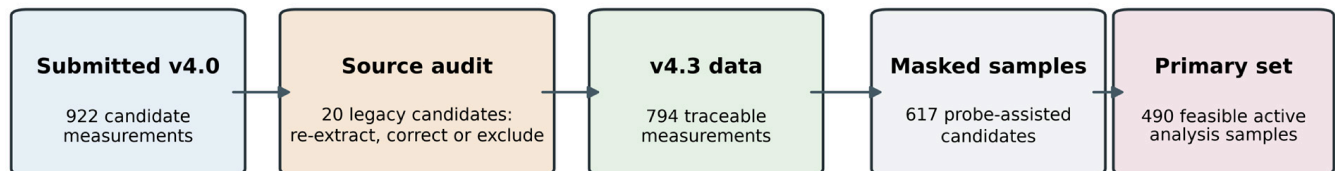
Figure 1. Leakage-safe probe-assisted contact-angle prediction and validation framework.

2. Materials and Methods

2.1. Data Construction, Provenance Audit and Prediction Task

The submitted v4.0 schema was an internally curated collection, not an established external database. It comprised five normalized long tables for sources, surfaces, liquids,

measurements and frozen splits. During revision, 20 legacy source candidates were checked against DOI metadata and original table or supplement locations. Nine were re-extracted row by row, three were retained after direct verification, two were retained after metadata correction, and six were excluded. Duplicate DOI identities were consolidated into one source group (Figure 2). The resulting v4.3 collection (Table 1) contains 794 measurements, 35 sources, 338 source-conditioned surface instances and 11 liquids.



Primary eligibility

target liquid removed | at least two unique non-target liquids | interior or boundary NNLS fit

Censored values, unsupported liquids, duplicate DOI identities and unverified source rows are not converted into point-valued targets.

Figure 2. Revision v4.3 provenance audit and primary eligibility flow.

Table 1. Revision v4.3 dataset and active probe-assisted cohorts.

Cohort	Samples	Sources	Surfaces
Complete v4.3 measurements	794	35	338
Primary eligible probe-assisted set	490	20	156
Development: training	204	6	62
Development: validation	19	2	6
Internal test	50	4	14
Legacy external	23	2	10
Fixed cross-source external confirmation	194	6	64

Quality control preserved the reported contact-angle type, measurement method, material and treatment state, uncertainty and replicate count when available. Liquid aliases and units were normalized; censored values were not converted to invented point targets; secondary records and untraceable rows were excluded. Raw publisher files are kept in the internal audit package, while the public reproducibility package records source locators and licensing constraints.

For each target, every measurement of the target liquid on the same surface is removed before any surface-energy reconstruction or feature calculation. Primary eligibility requires at least two unique non-target probe liquids and an interior or boundary nonnegative least-squares fit. Single-probe, singular, censored and insufficient-probe cases are retained only in audit files.

2.2. Target-Masked Nonnegative OWRK Reconstruction

For a liquid with total surface tension γ_L and dispersive and polar components γ_L^d and γ_L^p , the Owens-Wendt-Rabel-Kaelble (OWRK) relation is

$$\frac{\gamma_L(1 + \cos(\theta))}{2} = \sqrt{\gamma_S^d \gamma_L^d} + \sqrt{\gamma_S^p \gamma_L^p} \quad (1)$$

Here θ is the measured contact angle in degrees, and γ_s^d and γ_s^p are the solid dispersive and polar surface-free-energy (SFE) components in mJ m^{-2} . After removing the target liquid, nonnegative least squares (NNLS) estimates the square roots of the two solid components:

$$\min_{a,b \geq 0} \|X(a,b)^T - y\|_2^2, \quad a = \sqrt{\gamma_s^d}, \quad b = \sqrt{\gamma_s^p} \quad (2)$$

The target-masked fit returns both SFE components, the physical target-angle prediction, residual norm, probe counts and one of four statuses: interior fit, boundary fit, singular fit or insufficient probes. Squaring a negative unconstrained coefficient is prohibited because it would turn a physically invalid sign into an apparently valid energy component.

2.3. Physics-Residual Models and Implementation

The primary XGBoost residual tree [21] predicts the observed-minus-physical correction from surface, target liquid, reported experimental conditions and the target-masked physics summary. Random forest [22] is retained as a tree sensitivity model, and a compact neural expert [23] predicts a bounded correction in cosine space:

$$\cos(\hat{\theta}) = \text{clip}(\cos(\theta_{\text{NNLS}}) + \Delta_{\cos}, -1, 1), \quad \Delta_{\cos} = \delta_{\max} \tanh(z) \quad (3)$$

The neural objective (Table 2) combines angle Huber loss, cosine mean-squared error, residual shrinkage and heteroscedastic Gaussian negative log likelihood:

$$L = 1.00 L_{\text{Huber}} + 0.50 L_{\cos} + \lambda_r L_{\text{residual}} + 0.05 L_{\text{NLL}} \quad (4)$$

Table 2. Locked model and training parameters.

Component	Locked Setting
Surface/condition encoder	64-32; SiLU; dropout 0.10; 8-dimensional categorical embeddings
Target-liquid encoder	32-16; SiLU
Physics-summary encoder	32-16; SiLU
Residual and variance heads	64-32-1; bounded cosine residual; heteroscedastic log variance
Optimizer	AdamW; learning rate 0.001; weight decay 0.0001; batch size 64
Training control	Search: maximum 400 epochs and patience 50; final: maximum 1000 epochs and patience 100; gradient norm clipped at 5
Seeds	7, 19, 42, 67 and 99
Tree expert	Official XGBoost; 400 estimators; depth 3; learning rate 0.03; subsample 0.85; column subsample 0.85; minimum child weight 3
Random-forest expert	500 trees; square-root feature sampling; minimum leaf size 2
Loss weights	1.00 Huber (delta 5 degrees); 0.50 cosine MSE; 0.05 residual penalty; 0.05 Gaussian NLL
Selected neural controls	Maximum cosine correction 0.35; source-balanced sampling True

2.4. Source-Disjoint Validation and Fixed Cross-Source External Confirmation Set

Development uses nested source-group cross-validation (CV) with four outer folds and three inner folds. Preprocessing, hyperparameter selection and nonnegative fusion weights are fitted only within the corresponding training sources. The internal test, legacy external, and fixed cross-source external confirmation cohorts are not read by the selection code. The last cohort influenced early project architecture and is therefore called a fixed cross-source external confirmation set. Its results are descriptive confirmation, not the primary statistical evidence.

We report mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R^2), median absolute error (MedAE) and 90th-percentile absolute error

(P90AE). Paired 10,000-resample bootstraps resample complete surfaces; a source-cluster analysis tests sensitivity to the smaller number of independent publications.

2.5. Uncertainty, OOD Risk and Refusal

Five fixed seeds estimate ensemble variability. A 90% split-conformal interval [24,25] and an out-of-distribution (OOD) k-nearest-neighbor distance are calibrated from development out-of-fold (OOF) predictions only. The model requires at least two unique non-target probes and a feasible NNLS state. An OOD distance above the OOF 95th percentile or interval width above the OOF 90th percentile raises risk or refusal. Highly rough, porous, anisotropic or chemically heterogeneous surfaces are marked unvalidated because no defensible numeric cutoff can be inferred from the data.

2.6. Reviewer-Driven Robustness Audits

Additional analyses (Supplementary Figures S1 and S2 and Tables S1–S16) varied training-source count from two to seven, stratified OOF and fixed cross-source external confirmation set errors by material-family and target-liquid familiarity and performed strict leave-one-liquid-out diagnostics for water, diiodomethane, ethylene glycol and formamide. In the strict diagnostic, the held liquid is removed as both target and probe from all training samples. Fractions of predictions within 10 and 20 degrees quantify screening utility.

3. Results

3.1. Source Audit and Data Coverage

The row-level audit reduced the active primary set from the submitted 587-sample analysis to 490 traceable samples. This was not a cosmetic removal: several sources required replacement values and corrected angle types, and six legacy sources were excluded. Quantitative roughness was available for only 13 of 156 primary surfaces, representing 32 samples from 2 sources. Humidity, droplet volume, porosity and anisotropy were also incompletely reported. These variables cannot support a source-independent roughness cutoff.

The held-source learning diagnostic did not show a stable monotonic scaling law. Median residual XGBoost MAE changed from 16.8 degrees with two training sources to 13.9 degrees with seven, but the seven-source subsets ranged from 7.5 to 76.8 degrees. Composition effects therefore dominate any simple sample-count interpretation, and eight development sources provide limited source-level statistical power.

3.2. Nested Source-Group Validation

On 223 development samples (Table 3; Figure 3), NNLS physics produced 18.3 degrees MAE. The pre-specified physics-residual XGBoost model produced 15.5 degrees MAE and 0.43 R^2 . The model ranking and uncertainty are reported without using the fixed cross-source external confirmation set.

Table 3. Nested source-group cross-validation on the v4.3 development pool.

Model	<i>n</i>	MAE	RMSE	R^2	MedAE	P90AE
NNLS physics	223	18.3	26.3	0.39	12.6	43.3
Direct XGBoost	223	17.5	27.8	0.31	10.3	43.4
Neural residual	223	17.2	27.0	0.35	11.2	34.9
Residual RF	223	13.7	22.6	0.55	7.8	31.4
Residual XGBoost	223	15.5	25.4	0.43	9.1	35.2
OOF fusion	223	14.6	24.7	0.45	8.0	32.4

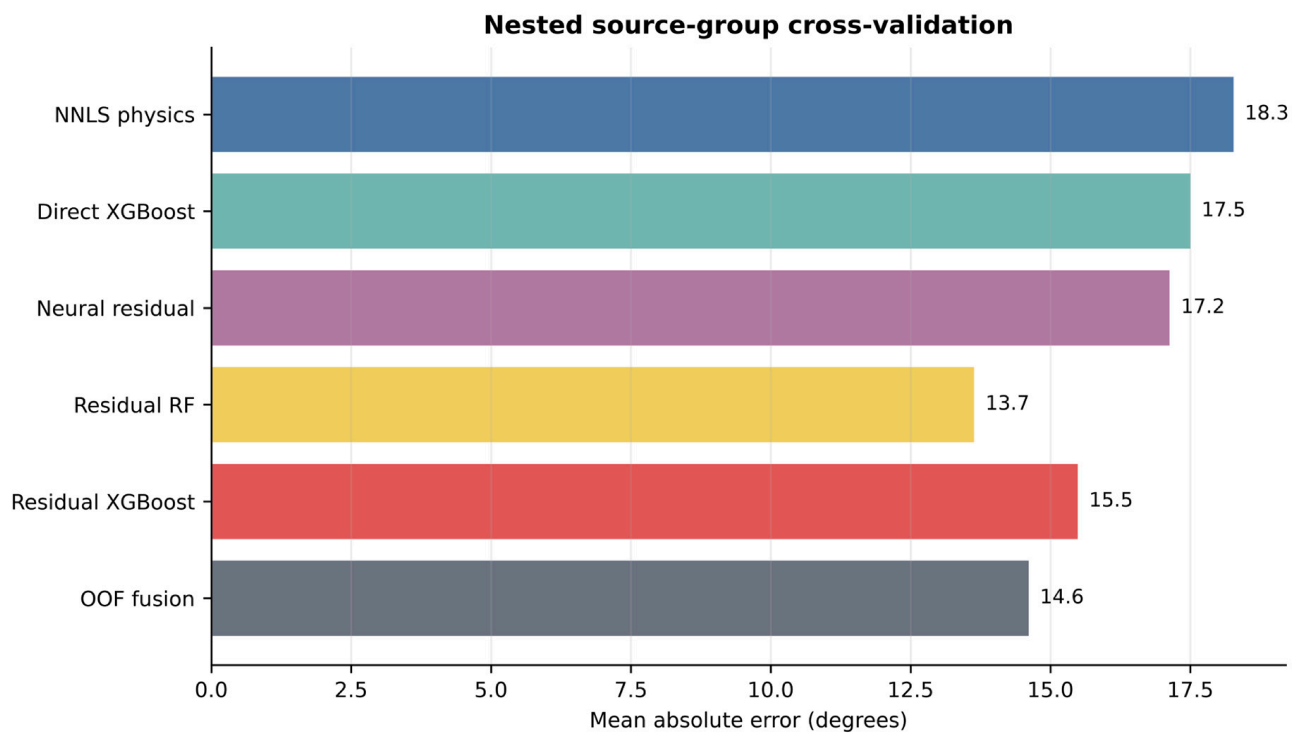


Figure 3. Mean absolute error under nested source-group cross-validation. Lower is better.

3.3. Fixed Cross-Source External Confirmation Set

The fixed cross-source external confirmation set (Table 4; Figure 4) contains 194 eligible samples. Physics-residual XGBoost achieved 13.2 degrees MAE. Because this cohort influenced early architecture decisions, it is not used to claim blind generalization and did not trigger model reselection.

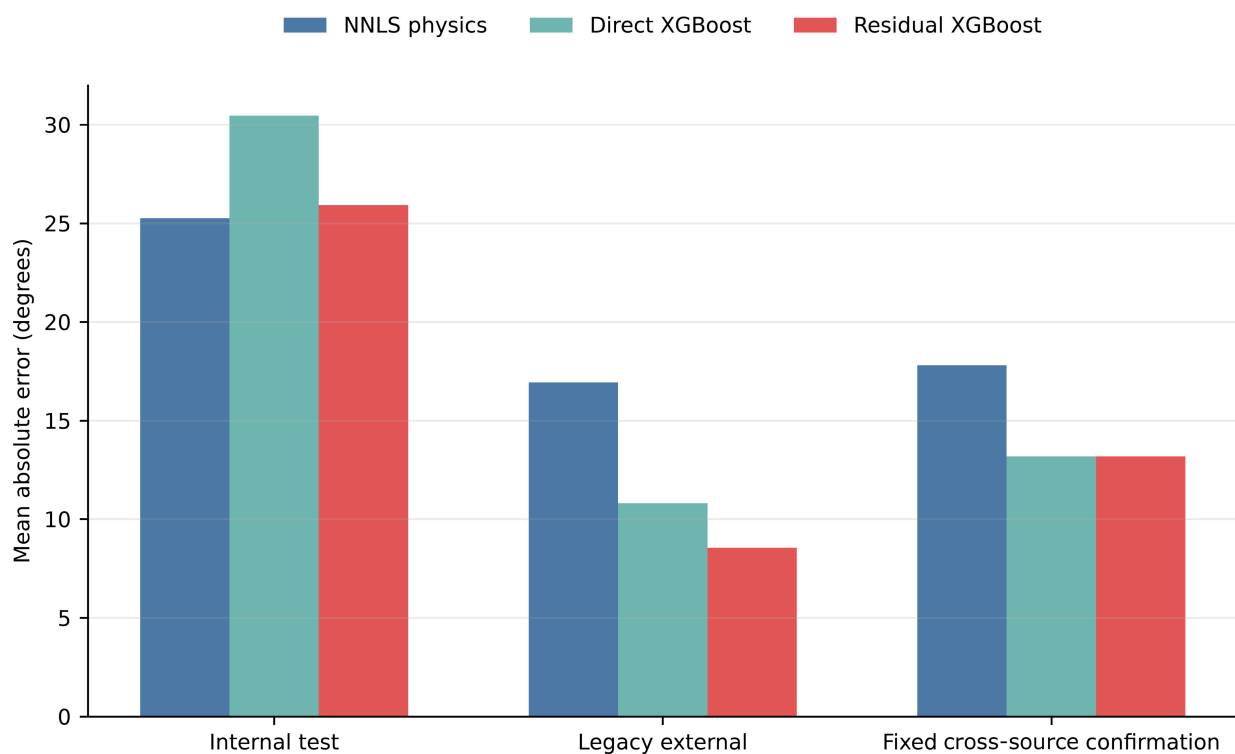


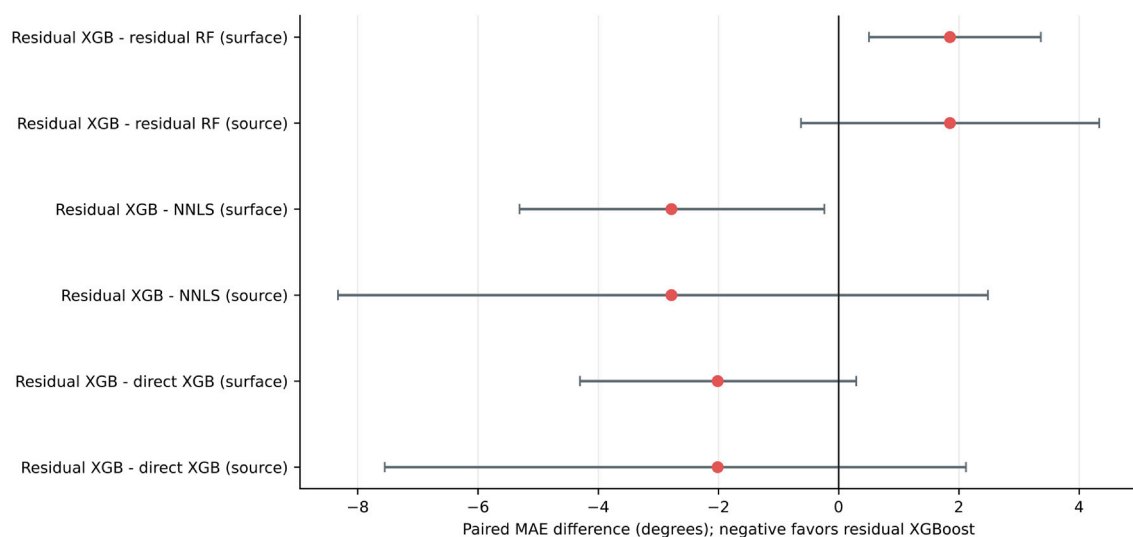
Figure 4. Model errors across fixed cross-source evaluation cohorts.

Table 4. Performance on the fixed cross-source confirmation set.

Model	<i>n</i>	MAE	RMSE	R ²	MedAE	P90AE
NNLS physics	194	17.8	25.3	0.60	12.4	38.6
Direct XGBoost	194	13.2	19.2	0.77	8.2	30.9
Neural residual	194	17.3	23.3	0.66	13.3	35.0
Residual RF	194	12.1	18.8	0.78	6.7	29.9
Residual XGBoost	194	13.2	19.2	0.77	8.3	29.7
OOF fusion	194	12.7	19.0	0.78	8.4	31.2

3.4. Bootstrap Evidence and Model Choice

Residual XGBoost minus NNLS (Figure 5) had a nested MAE difference of -2.8 degrees. The surface-cluster 95% interval was $[-5.3, -0.2]$, while the source-cluster interval was $[-8.3, 2.5]$. Any interval crossing zero is interpreted as directional evidence rather than statistical superiority. Residual RF obtained a lower nested MAE than residual XGBoost (13.7 versus 15.5 degrees); the paired XGBoost-minus-RF interval was $[0.5, 3.4]$ by surface and $[-0.6, 4.3]$ by source. Residual XGBoost remains the designated primary analysis because that model family was frozen in the revision protocol before the v4.3 audited rerun. RF is reported transparently as a stronger surface-level sensitivity result rather than being promoted after inspection.

**Figure 5.** Surface- and source-cluster paired bootstrap intervals for the primary residual model.

3.5. Practical Accuracy, Uncertainty and Selective Use

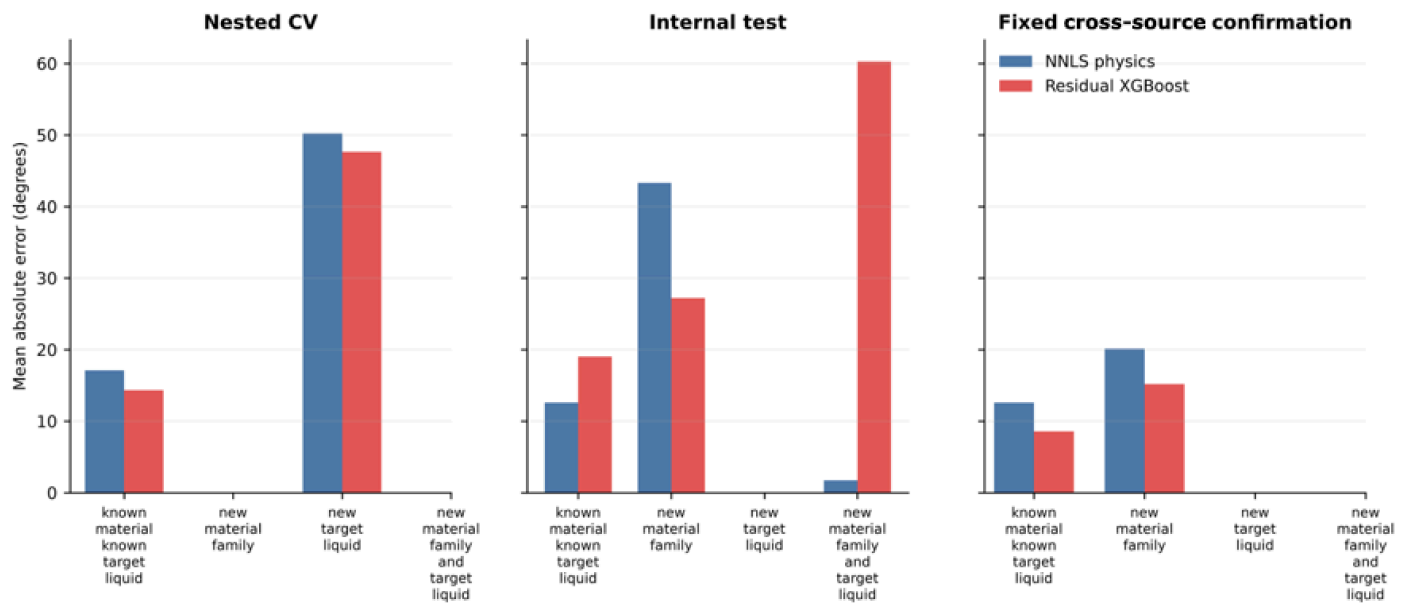
In nested validation, 52.5% of residual predictions were within 10 degrees and 79.8% were within 20 degrees. The corresponding fixed cross-source external confirmation proportions were 56.7% and 78.9%. The nested 90% conformal interval covered 88.3% of observations with a 78.2-degree mean width. The refusal rule retained 79.8% of samples and reduced fusion MAE from 14.6 to 11.9 degrees. These results support candidate screening and prioritization, not replacement of high-precision goniometric measurement. Calibrated intervals and refusal flags should accompany every point estimate.

3.6. New-Material and New-Liquid Boundaries

The internal dual-unknown stratum (Table 5; Figure 6) contained only 4 samples. NNLS MAE was 1.7 degrees, whereas residual XGBoost MAE rose to 60.3 degrees. This small but severe failure case is an explicit no-transfer warning for simultaneously unseen material families and target liquids.

Table 5. Physics-residual XGBoost by material-family and target-liquid novelty.

Evaluation	Training Familiarity	<i>n</i>	MAE	R ²
nested_source_cv	known material known target liquid	215	14.3	0.47
nested_source_cv	new target liquid	8	47.6	−0.24
internal test	known material known target liquid	24	19.0	−0.33
internal test	new material family	22	27.2	0.58
internal test	new material family and target liquid	4	60.3	−1454.78
fixed cross-source external confirmation	known material known target liquid	59	8.6	0.74
fixed cross-source external confirmation	new material family	135	15.2	0.77

**Figure 6.** Error stratified by material-family and target-liquid familiarity.

Strict leave-one-liquid-out evaluation was feasible only when enough other liquids remained for training and target-masked NNLS. Evaluated diagnostics were diiodomethane: 33.1 degrees MAE ($n = 65$); ethylene glycol: 23.6 degrees MAE ($n = 54$); formamide: 7.8 degrees MAE ($n = 28$). Other requested diagnostics were water: infeasible_after_strict_target_and_probe_exclusion. Small strata and infeasible liquids are reported as boundaries, not imputed evidence of transfer. In particular, deterioration for a held-out liquid is treated as a no-transfer warning rather than averaged away.

4. Discussion

The principal result is conditional. A target-masked physical summary can provide a useful anchor for a residual tree across heterogeneous sources, but the evidence does not support universal superiority. The source-cluster analysis is intentionally conservative because eight development sources provide much less independent information than the sample count suggests.

The compact neural expert achieved 17.2 degrees nested source-CV MAE, received a consensus fusion weight of 0.080, and did not provide a standalone performance advantage over the selected residual tree. With a small, heterogeneous development pool, neural capacity can fit sample-level patterns without learning a stable source-transfer correction. This result argues for compact physical summaries and grouped validation before increasing network complexity.

The residual target was chosen rather than direct XGBoost because nested source validation evaluates whether the learner corrects a physically interpretable baseline. Residual

XGBoost was pre-specified before the source-audited rerun. Residual RF subsequently produced lower fold-pooled MAE, but its source-cluster advantage remained uncertain; we report it as a sensitivity result instead of changing the primary model after inspection. A direct tree can also perform well in a fixed cohort, but selecting it after inspecting that cohort would turn confirmation into tuning. The residual formulation exposes when a learned correction overwhelms or contradicts the physical estimate.

Reported R^2 values above 0.9 in controlled manufacturing, polymer or reservoir studies are not contradicted by the lower cross-source values here. Those studies often use consistent process variables, direct morphology/chemistry, one material domain or random within-domain splits. The present task deliberately removes the target liquid and transfers across publication protocols with many missing descriptors. The resulting errors quantify a harder and less homogeneous problem.

4.1. Application Perspective

A supported use case is screening a functional surface for a target liquid after at least two other probe liquids have been measured. The physical estimate, learned correction, interval, OOD score and refusal flag can be inspected together to prioritize coating or membrane candidates. The model should not be used as a substitute for final contact-angle measurement or to certify a surface outside the audited material and protocol envelope.

4.2. Limitations and No-Use Conditions

Quantitative roughness, porosity, atomic asperity scale, anisotropy, surface chemistry, humidity and droplet volume are too sparsely and unevenly reported to estimate universal cutoffs. Roughness is confounded with source identity, so a numeric threshold would be misleading. Highly rough, porous, anisotropic, chemically heterogeneous, dynamic-wetting, novel two-dimensional and bio-inspired surfaces remain unvalidated unless they are represented by new independent sources. The hard input conditions, OOD P95 flag and interval-width P90 refusal rule are evidence-based operational bounds, not guarantees.

5. Conclusions

- Row-level provenance auditing replaced or removed unsupported legacy records and produced a v4.3 analysis set of 490 target-masked probe-assisted samples from 20 sources.
- Physics-residual XGBoost changed nested source-CV MAE from 18.3 to 15.5 degrees; source-cluster uncertainty determines whether this is reported as directional or statistically supported.
- The compact neural expert achieved 17.2 degrees nested MAE with consensus weight 0.080; it did not provide a standalone performance advantage over the selected residual tree.
- The model is a screening tool requiring at least two non-target probes and feasible NNLS; sparse roughness/condition metadata and only eight development sources preclude universal claims for novel materials or liquids.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app16157759/s1>. Figure S1: Held-source error versus the number of training sources. Lines show medians and ribbons show interquartile ranges; Figure S2: Strict leave-one-liquid-out diagnostics for liquids with sufficient remaining training data; Table S1: Revision v4.3 legacy-source decisions; Table S2: Contact-angle-type distribution after source audit; Table S3: Primary-analysis target-liquid distribution; Table S4: Primary-analysis material-family distribution; Table S5: Experimental-condition and uncertainty metadata completeness; Table S6: Quantitative roughness coverage by analysis cohort; Table S7: Selected v4.3 configuration; Table S8: Outer-fold

selected configurations; Table S9: Complete v4.3 core model metrics; Table S10: Surface- and source-cluster paired bootstrap comparisons; Table S11: Source-learning curve summary; Table S12: Strict leave-one-liquid-out results and infeasibility states; Table S13: Material-family and target-liquid novelty strata; Table S14: Fractions within practical absolute-error thresholds; Table S15: Recent contact-angle and wettability prediction studies; Table S16: Locked run manifest.

Author Contributions: Conceptualization, W.L. and M.S.; methodology, M.S. and W.L.; software, M.S.; validation, Y.X., R.X. and X.G.; formal analysis, M.S. and Y.X.; investigation, Y.X., R.X. and X.G.; resources, W.L.; data curation, Y.X., R.X., X.G. and M.S.; writing—original draft preparation, M.S. and Y.X.; writing—review and editing, W.L., M.S., Y.X., R.X. and X.G.; visualization, M.S.; supervision, W.L.; project administration, W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 92567204 and 62472194, the Fundamental Research Funds for the Central Universities, and the Jilin Science and Technology Research Project under Grant No. 20260203130SF.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The v4.3 source-audited data, frozen source- and surface-disjoint splits, training code, predictions, statistical outputs and run manifests are openly available in GitHub software release v4.3.0 (<https://github.com/jiaowoshendaxia/contact-angle-pcrnn-dataset-v3.4/releases/tag/v4.3.0> (accessed on 28 July 2026)) and Zenodo version record <https://doi.org/10.5281/zenodo.21545090> [26]. The all-versions concept DOI is <https://doi.org/10.5281/zenodo.20382890>. Publisher files that cannot be redistributed are represented by DOI, source locator and extraction audit only.

Acknowledgments: Generative AI tools were used for language editing, code review and document-formatting assistance. All source extraction, numerical results, scientific interpretations and final text were checked and approved by the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Young, T. An Essay on the Cohesion of Fluids. *Philos. Trans. R. Soc. Lond.* **1805**, *95*, 65–87. [CrossRef]
2. Dupre, A. *Theorie Mecanique de la Chaleur*; Gauthier-Villars: Paris, France, 1869.
3. Owens, D.K.; Wendt, R.C. Estimation of the Surface Free Energy of Polymers. *J. Appl. Polym. Sci.* **1969**, *13*, 1741–1747. [CrossRef]
4. Kaelble, D.H. Dispersion-Polar Surface Tension Properties of Organic Solids. *J. Adhes.* **1970**, *2*, 66–81. [CrossRef]
5. Fowkes, F.M. Attractive Forces at Interfaces. *Ind. Eng. Chem.* **1964**, *56*, 40–52. [CrossRef]
6. Good, R.J.; Girifalco, L.A. A Theory for Estimation of Surface and Interfacial Energies. *J. Phys. Chem.* **1960**, *64*, 561–565. [CrossRef]
7. Neumann, A.W.; Good, R.J.; Hope, C.J.; Sejpal, M. An Equation-of-State Approach to Determine Surface Tensions of Low-Energy Solids from Contact Angles. *J. Colloid Interface Sci.* **1974**, *49*, 291–304. [CrossRef]
8. van Oss, C.J.; Chaudhury, M.K.; Good, R.J. Interfacial Lifshitz-van der Waals and Polar Interactions in Macroscopic Systems. *Chem. Rev.* **1988**, *88*, 927–941. [CrossRef]
9. Wenzel, R.N. Resistance of Solid Surfaces to Wetting by Water. *Ind. Eng. Chem.* **1936**, *28*, 988–994. [CrossRef]
10. Cassie, A.B.D.; Baxter, S. Wettability of Porous Surfaces. *Trans. Faraday Soc.* **1944**, *40*, 546–551. [CrossRef]
11. Peta, K. Surface Wettability Modeling and Predicting via Artificial Neural Networks. *Materials* **2025**, *18*, 191. [CrossRef] [PubMed]
12. Fu, J.; Li, Y.; Song, X.; Liu, C.; Zhao, R.; Wang, H.; Wang, Q. Wettability Prediction and Feature Importance Analysis for Laser-Chemical Induced Superwetting Surface Based on Machine Learning. *J. Manuf. Process.* **2025**, *152*, 57–67. [CrossRef]
13. Sena, J.; Johannissen, L.O.; Blaker, J.J.; Hay, S. A Machine Learning Model for the Prediction of Water Contact Angles on Solid Polymers. *J. Phys. Chem. B* **2025**, *129*, 2739–2745. [CrossRef] [PubMed]
14. Shaw, D.G.; Liang, R.; Zheng, T.; Qi, J.; Berry, J.D. Accurate and Robust Static Hydrophobic Contact Angle Measurements Using Machine Learning. *Langmuir* **2024**, *40*, 16757–16770. [CrossRef] [PubMed]
15. Pan, B.; Song, T.; Yue, M.; Chen, S.; Zhang, L.; Edlmann, K.; Neil, C.W.; Zhu, W.; Iglauer, S. Machine Learning-Based Shale Wettability Prediction: Implications for H₂, CH₄ and CO₂ Geo-Storage. *Int. J. Hydrogen Energy* **2024**, *56*, 1384–1390. [CrossRef]

16. Mwakipunda, G.C.; Ibrahim, A.-W.; Kouassi, A.K.F.; Nafouanti, M.B.; Yu, L. Improving Wettability Estimation in Carbonate Formation Using Machine Learning Algorithms. *Int. J. Hydrogen Energy* **2025**, *111*, 781–797. [[CrossRef](#)]
17. Song, T.; Zhu, W.; Pan, B.; Song, H.; Chen, Z.; Yue, M. Development of Ensemble Learning Techniques and Sequential Model-Based Optimization for Enhancing the Generalizability of Shale Wettability Predictions. *Mar. Pet. Geol.* **2024**, *168*, 107010. [[CrossRef](#)]
18. Roosta, A.; Zendejboudi, S.; Rezaei, N. Estimating Advancing and Receding Contact Angles for Pure and Mixed Liquids on Smooth Solid Surfaces Using the PCP-SAFT Equation of State. *Phys. Chem. Chem. Phys.* **2025**, *27*, 6031–6042. [[CrossRef](#)] [[PubMed](#)]
19. Pan, C.; Qiao, C.; Feng, S.; Li, Y.; Zheng, Y.; Ye, H. Robust Measurement of Surface Tension and Contact Angle Using Physics-Informed Autoencoders. *Colloids Surf. A* **2026**, *730*, 138972. [[CrossRef](#)]
20. Otchere, D.A.; Mohammed, M.A.A.; Ganat, T.O.A.; Gholami, R.; Merican, Z.M.A. A Novel Empirical and Deep Ensemble Super Learning Approach in Predicting Reservoir Wettability via Well Logs. *Appl. Sci.* **2022**, *12*, 2942. [[CrossRef](#)]
21. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of KDD*; ACM: New York, NY, USA, 2016; pp. 785–794.
22. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
23. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
24. Vovk, V.; Gammerman, A.; Shafer, G. *Algorithmic Learning in a Random World*; Springer: New York, NY, USA, 2005.
25. Lei, J.; G'Sell, M.; Rinaldo, A.; Tibshirani, R.J.; Wasserman, L. Distribution-Free Predictive Inference for Regression. *J. Am. Stat. Assoc.* **2018**, *113*, 1094–1111. [[CrossRef](#)]
26. Xia, Y.; Liu, W.; Shen, M.; Xing, R.; Gao, X. *Leakage-Safe Probe-Assisted Contact Angle Prediction Reproducibility Package, Source-Audited Revision v4.3*, Version v4.3.0; Zenodo: Geneva, Switzerland, 2026. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.