Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.36210

Package-Conditioned Active Learning for Thermally Conductive Insulators

Haowei Zheng

College of Computer and Information Science and College of Software, Southwest University, Chongqing, China
zzz23842609@163.com

Abstract. Thermally conductive electrical insulators are needed to remove heat from heterogeneous artificial-intelligence hardware, yet conventional screening often maximizes bulk conductivity without considering electrical confidence, thermoelastic compatibility, or package-level benefit. ThermoDielectric-AL is introduced as a package-conditioned, uncertainty-aware workflow combining 1,056 open dielectric records, 1,181 open elastic records, physics-derived Debye, Slack, and expansion descriptors, chemistry-cluster validation, and sequential acquisition. Leave-one-composition-cluster-out validation shows that random cross-validation inflates coefficient-of-determination values by 0.19–0.32. An extremely randomized-tree ensemble nevertheless retains useful out-of-domain ranking for band gap and a Slack thermal-conductivity proxy, while calibrated ensemble intervals provide a conservative electrical-feasibility gate. In retrospective search, Thompson sampling recovers 94.8% of the true top 1% after evaluating 25.4% of the candidate pool, compared with 21.9% for random selection. A one-dimensional package model indicates rapidly diminishing temperature benefit above about 30 W m-1 K-1. The combined evidence prioritizes Al2O3, BN, and Si3N4 as defensible insulating candidates and retains AlN as a higher-uncertainty prospect. All thermal quantities are calculated screening proxies rather than measurements, defining testable priorities for package-specific experiments.

Keywords: active learning, thermal management, electrical insulators, materials informatics, heterogeneous integration

1. Introduction

Advanced packages integrate logic, high-bandwidth memory, and accelerators through dense interconnects [1, 2], improving communication but creating hotspots, constrained dielectrics, and anisotropic heat paths. Temperature affects transport, leakage, timing, and lifetime [3]; even with architectural cooling [4], insulating layers must remove heat efficiently.

Artificial intelligence and open databases accelerate materials discovery [5, 6], but package insulators must jointly satisfy thermal, electrical, expansion, and processing requirements. Dielectric and elastic datasets provide complementary evidence [7, 8] through materials-informatics tools [9], yet overlap is sparse and random splits may test interpolation rather than chemical transfer [10].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.36210

Active learning selects evaluations using uncertainty and utility [11, 12] and has enabled closed-loop discovery [13], thermal-conductivity prediction [14], and polymer trade-off searches [15]. ThermoDielectric-AL connects sequential search to electrical confidence, expansion compatibility, and package benefit, asking which defensible candidate should be evaluated next rather than which crystal has the largest bulk-conductivity proxy.

2. Materials and methods

2.1. Data, provenance, and descriptors

The matminer dielectric_constant branch contained 1,056 Materials Project structures, Perdew–Burke–Ernzerhof (PBE) gaps, and electronic/ionic dielectric tensors [7, 16]; elastic_tensor_2015 contained 1,181 structures, Voigt–Reuss–Hill moduli, Poisson ratios, and anisotropies [8]. Identifiers and reduced formulas defined evidence tiers without equating polymorphs.

Descriptors summarized elemental fractions, atomic attributes, density, atomic volume, site count, space group, and lattice. Median imputation, scaling, principal-component analysis, and clustering were fitted within training folds. Clusters used k-means with 30 initializations and no target labels. Processing used pymatgen [17].

2.2. Physics-derived proxies

For density ρ, shear modulus G, and bulk modulus K, transverse and longitudinal velocities were calculated as

$$v_t = (G/\rho)^{1/2}, \qquad v_l = [(K + 4G/3)/\rho]^{1/2} \tag{1}$$

The mean acoustic velocity and Anderson Debye temperature [18] were

$$v_m = [(2v_t^{-3} + v_l^{-3})/3]^{-1/3} \tag{2}$$

$$\theta_D = (h/k_B)[3n/(4\pi V)]^{1/3} v_m \tag{3}$$

A Poisson-ratio Grüneisen proxy and the 300 K Slack conductivity proxy [19, 20] were

$$\gamma = 3(1+\nu)/[2(2-3\nu)] \tag{4}$$

$$\kappa_{Slack} = 3.1 \times 10^{-6} M_{av} \theta_D^3 \delta / (\gamma^2 n_p^{2/3} T) \tag{5}$$

Here M_{av} is mean atomic mass, δ the cube root of atomic volume, n_p the primitive-cell atom count, and $T = 300$ K. Numerically integrated Debye heat capacity gave

$$\alpha_L = \gamma C_V / (3KV) \tag{6}$$

A Cahill-like minimum-conductivity descriptor was retained [21, 22]. Because these equations omit anharmonicity, defects, porosity, texture, and interfaces, all thermal outputs are proxies.

DOI: 10.54254/2755-2721/2026.36210

2.3. Validation, uncertainty, and acquisition

Ridge regression, extremely randomized trees (ExtraTrees) [23, 24], and a neural network were compared in scikit-learn [25]. For primary leave-one-composition-cluster-out (LOCCO) validation, eight clusters from 18 principal components were held out sequentially; random five-fold validation diagnosed interpolation. ExtraTrees used 120 trees, minimum leaf size two, and 72% of features per split; the neural network used 64/32 hidden units, L_2 regularization, and early stopping.

Tree dispersion was calibrated against LOCCO errors [26, 27] to obtain a 90% band-gap lower confidence bound (LCB_{90}). For eight seeds, acquisition began with 60 labels; a 50-tree model was refitted at budgets of 60, 120, 180, 240, and 300 records. Random, greedy, upper-confidence-bound (UCB; $\mu + 1.5\sigma$), and Thompson-style ($\mu + z\sigma$) policies were compared. The fixed oracle weighted thermal and coefficient-of-thermal-expansion (CTE) scores 0.68/0.32; electrical predictions remained a separate gate.

2.4. Package-conditioned decision rule

Feasibility required $LCB_{90} \geq 1.25$ eV, Slack proxy ≥ 8 W m^{-1} K^{-1}, and expansion of 0–12 ppm K^{-1}. The thermal score was logistic in log conductivity, centred at 10 W m^{-1} K^{-1}; the CTE score was $\exp(-|\alpha_L - 2.6|/5)$; electrical confidence was logistic in LCB_{90}, centred at 2.5 eV. Utility weighted these scores 0.55/0.20/0.25; weights were fixed preferences, not fitted coefficients. Pareto ranking maximized thermal proxy and gap confidence while minimizing expansion mismatch. Actinides, Be, Cd, Hg, Pb, Tl, and As were excluded; this is not a complete safety assessment.

For power P, layer area A, thickness t, and conductivity k, package sensitivity was evaluated using

$$\Delta T_{layer} = Pt/(Ak) \tag{7}$$

Calculations used $P = 400$ W, $A = 400$ mm^2, $t = 40$ μm, $k = 1$ –300 W m^{-1} K^{-1}, and a 1.4 W m^{-1} K^{-1} SiO_2-like baseline; this is a bottleneck calculation, not a full package solver.

2.5. Statistics, reproducibility, and availability

All random operations used seed 20260712. Metrics were R^2, mean absolute error (MAE), root mean squared error (RMSE), and Spearman correlation; acquisition results report mean ± standard deviation across eight initial pools. Intervals and seed variability are descriptive, not significance tests. The reproducibility bundle includes source JSON and checksums, derived tables, descriptors, predictions, all seed traces, plotting code, a claim–evidence audit, a pinned environment, and a one-command CPU workflow. Reviewed data and code snapshots should receive a DOI and public link before submission.

3. Results and discussion

3.1. Workflow and cross-source evidence

ThermoDielectric-AL links open density-functional-theory (DFT) data, physics proxies, chemical holdouts, uncertainty gating, acquisition, and package response (Figure 1). It makes no deposition, thermoreflectance, breakdown, or device-measurement claim.

DOI: 10.54254/2755-2721/2026.36210

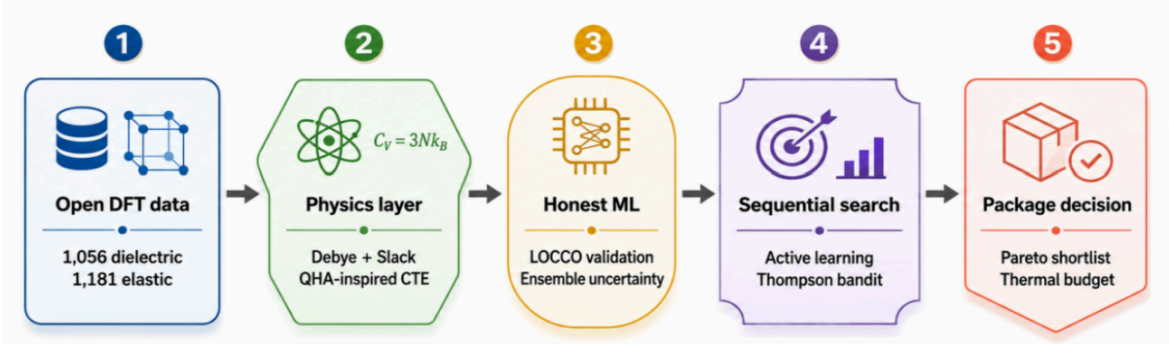

Figure 1. Workflow from open dielectric and elastic records through physics proxies, chemical holdouts, calibrated uncertainty, acquisition, and package decision; the dashed boundary marks future physical validation

Only 49 identifiers (4.6%/4.1% of the two sources) and 63 reduced formulas overlap (Table 1). The dielectric branch trains electrical surrogates; the elastic branch supplies thermal candidates. Identifier matches provide direct evidence, formula matches cautious transfer, and remaining electrical values are model-transferred.

Table 1. Open-data provenance and evidence roles

Component	Records	Observed or derived quantities	Role
Dielectric DFT	1,056	PBE gap; dielectric tensors	Electrical surrogate
Elastic DFT	1,181	Structure; elastic moduli; anisotropy	Debye, Slack, and expansion proxy pool
Exact identifier overlap	49	Same identifier in both datasets	Direct cross-source evidence
Reduced-formula overlap	63	Formula-level match	Tier-I evidence with polymorph caution

Figure 2 shows separated chemical manifolds and broad thermal variation: thousand-record datasets yield only tens of exact joint observations, requiring domain-shift validation.

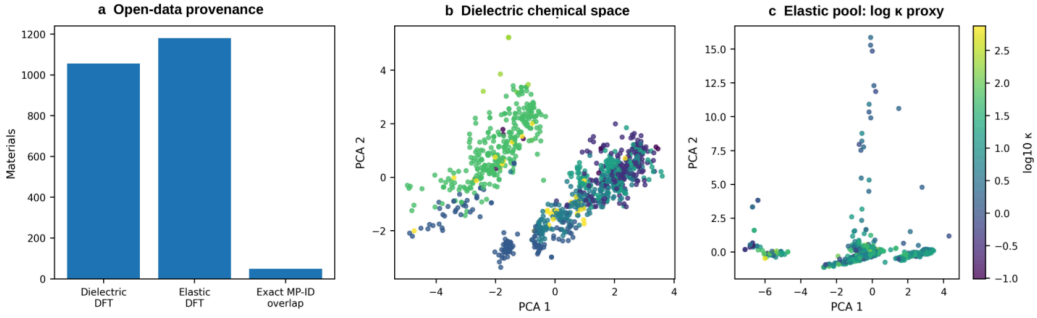

Figure 2. Data and chemical-space heterogeneity: (a) dataset overlap, (b) dielectric principal-component projection, and (c) elastic candidates coloured by log Slack proxy

DOI: 10.54254/2755-2721/2026.36210

3.2. Chemical holdouts and calibrated feasibility

Under LOCCO, ExtraTrees gave PBE-gap/Slack R^2 values of 0.485/0.454 and Spearman coefficients of 0.686/0.696 (Table 2), supporting ranking rather than metrology. Random validation inflated R^2 by 0.19–0.32; neural PBE-gap R^2 was 0.401 and did not exceed ExtraTrees. Ridge slightly exceeded ExtraTrees for dielectric R^2 (0.369/0.340) but lacked tree-wise uncertainty.

Table 2. Predictive performance under interpolation and chemical-domain shift

Target	Model / protocol	R^2	MAE	RMSE	Spearman ρ
PBE gap (eV)	ExtraTrees / LOCCO	0.485	0.856	1.151	0.686
PBE gap (eV)	ExtraTrees / random	0.720	0.570	0.848	0.849
$\log_{10}$ static dielectric	ExtraTrees / LOCCO	0.340	0.172	0.231	0.601
$\log_{10}$ static dielectric	ExtraTrees / random	0.658	0.106	0.166	0.857
$\log_{10}$ Slack κ proxy	ExtraTrees / LOCCO	0.454	0.303	0.412	0.696
$\log_{10}$ Slack κ proxy	ExtraTrees / random	0.644	0.236	0.332	0.817
PBE gap (eV)	Neural network / LOCCO	0.401	0.934	1.241	0.624

Parity plots show central compression and heteroscedastic errors (Figure 3). Multipliers 1.790/2.268 achieved 90%/95% PBE-gap coverage; dispersion–error Spearman 0.456 supported using LCB_{90}. BeO (114.8 W m^{-1} K^{-1}; 3.14 eV) was hazard-excluded; carbon (750.5 W m^{-1} K^{-1}) and B_6O failed at 0.29/0.93 eV. These are evidence, not measured-property, failures.

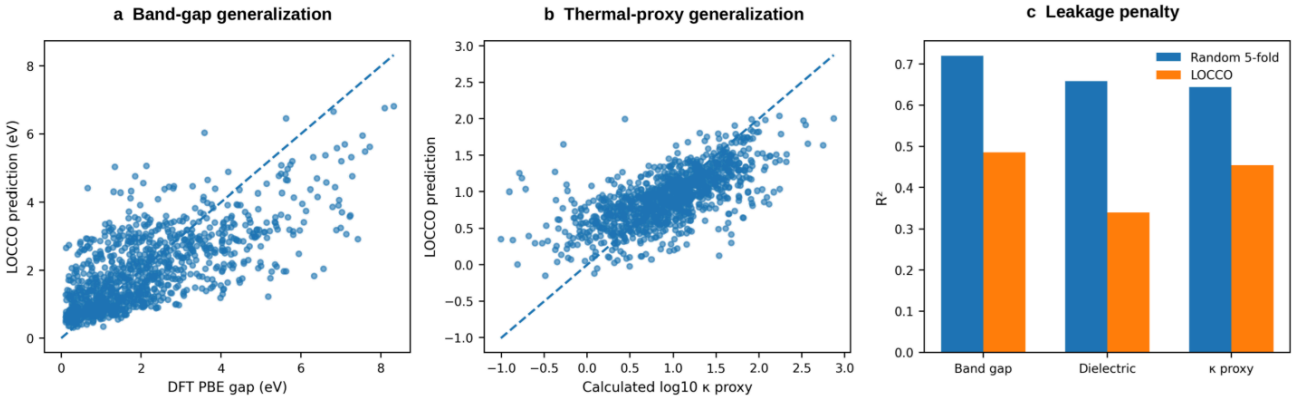

Figure 3. Out-of-domain validation: (a) PBE-gap parity, (b) Slack-proxy parity, and (c) R2 inflation of 0.19–0.32 under random splitting

3.3. Sequential-search efficiency

At 25.40% of the pool, Thompson, UCB, and greedy acquisition recovered 94.8%, 93.8%, and 92.7% of the top 1%, versus 21.9% randomly (Table 3). Best utility saturated earlier than recall (Figure 4), favouring a contextual bandit for this finite, one-step-feedback pool. A single-candidate search may stop near the utility plateau; discovering a robust family requires sustained recall.

Table 3. Retrospective sequential-search benchmark (mean ± standard deviation over eight seeds)

Budget	Pool evaluated	Acquisition	Top-1% recall	Best utility
120	10.16%	Random	5.2 ± 7.6%	0.883

DOI: 10.54254/2755-2721/2026.36210

Table 3. (continued)

120	10.16%	Greedy mean	61.5 ± 14.0%	0.949
120	10.16%	UCB	61.5 ± 12.5%	0.950
120	10.16%	Thompson bandit	43.8 ± 16.5%	0.942
300	25.40%	Random	21.9 ± 10.9%	0.941
300	25.40%	Greedy mean	92.7 ± 11.3%	0.951
300	25.40%	UCB	93.8 ± 8.6%	0.952
300	25.40%	Thompson bandit	94.8 ± 6.2%	0.952

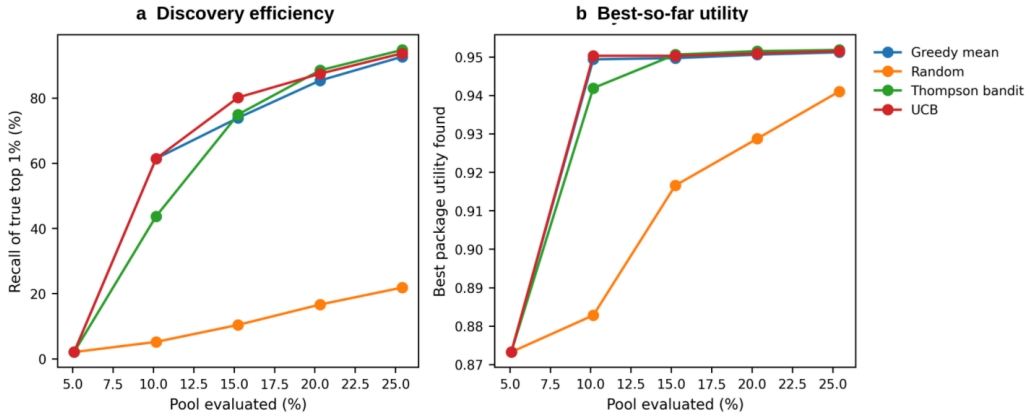

Figure 4. Retrospective acquisition across eight seeds. Model-guided policies outperform random selection and converge near 25.4%; best utility saturates before top-set recall

3.4. Package sensitivity and candidate shortlist

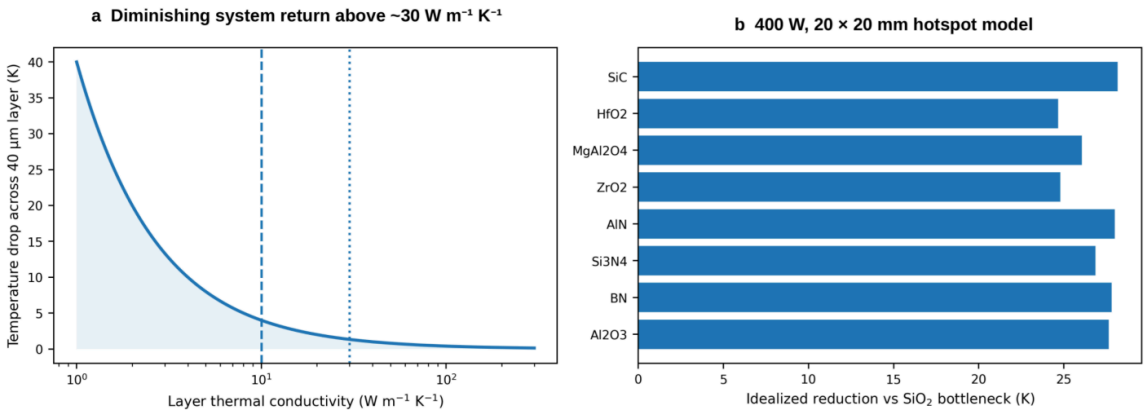

Figure 5. Conductivity diminishing returns: (a) calculated drop across the 40 µm layer and (b) reduction from a 1.4 W m-1 K-1 baseline; values are model outputs

The SiO_2-like layer contributes 28.57 K. At 10, 30, and 100 W m^{-1} K^{-1} it contributes 4.00, 1.33, and 0.40 K; 30 W m^{-1} K^{-1} captures 95.3% of the maximum reduction, while increasing to 100 gains only 0.93 K (Figure 5). Thus electrical confidence, expansion, anisotropy, interfaces, and processing dominate small increments. The model omits spreading, boundary resistance, bumps, underfill, vias, temperature dependence, and cooling conditions [28].

DOI: 10.54254/2755-2721/2026.36210

Al_2O_3 and BN share top utility (0.775; Table 4, Figure 6): their proxies/LCB_{90} values are 43.3 W m^{-1} K^{-1}/3.10 eV and 52.2/4.15; BN anisotropy of 217 makes orientation and polymorph control essential. Si_3N_4 offers close expansion matching and a 3.51 eV bound. AlN reaches 68.7 W m^{-1} K^{-1} but remains Tier II with no formula match and a 1.83 eV bound. ZrO_2 and HfO_2 retain gap confidence but have near-10 W m^{-1} K^{-1} proxies; $MgAl_2O_4$ is intermediate and prediction-dependent. SiC is retained only as a semiconducting heat-spreader benchmark. Unlike earlier screens [29-31], ranking reflects package sensitivity.

Table 4. Evidence-tiered shortlist for thermally conductive electrical insulation

Formula	Evidence	Slack κ proxy	CTE proxy	Gap LCB_{90}	Utility
Al_2O_3	Tier I	43.3	6.03	3.10	0.775
BN	Tier I	52.2	10.58	4.15	0.775
Si_3N_4	Tier I	23.3	5.92	3.51	0.740
AlN	Tier II	68.7	6.33	1.83	0.648
ZrO_2	Tier I	10.6	9.22	2.85	0.509
$MgAl_2O_4$	Tier II	15.9	9.07	2.21	0.496
HfO_2	Tier I	10.3	8.90	2.72	0.491
SiC	Benchmark	99.3	3.75	1.92	0.735

Slack κ is in W m^{-1} K^{-1}; coefficient-of-thermal-expansion (CTE) proxy is in ppm K^{-1}; gap LCB_{90} is in eV.

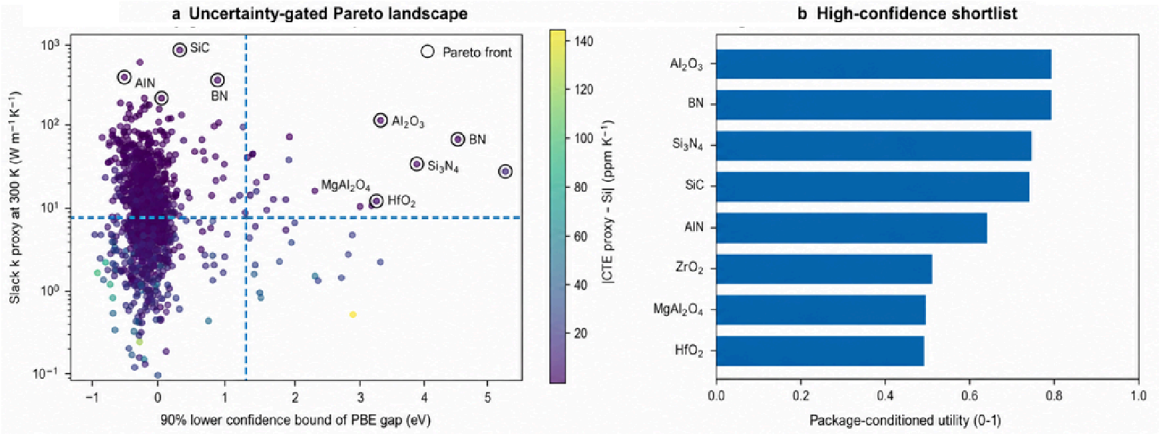

Figure 6. Uncertainty-gated shortlist: (a) Slack proxy versus gap LCB90, coloured by expansion mismatch, and (b) combined thermal, expansion, and electrical utility

3.5. Limitations and experimental route

This decision protocol does not provide measured properties. PBE gaps are not breakdown fields; formula evidence misses polymorph, texture, stoichiometric deviation, and deposition-induced phase; Slack theory omits film defects, porosity, grain boundaries, and interfaces. Expansion and package approximations add uncertainty, so Table 4 defines hypotheses, not datasheet values.

A closed-loop thin-film study should vary density, phase, and texture for Al_2O_3, BN, Si_3N_4, and AlN; measure conductivity by time- or frequency-domain thermoreflectance (TDTR/FDTR), resistivity, breakdown, dielectric response, stress, roughness, thickness, and boundary conductance;

DOI: 10.54254/2755-2721/2026.36210

then select composition–process pairs by uncertain package benefit for silicon-interposer or glass-carrier testing.

4. Conclusion

Package-conditioned active learning turns insulator screening into an auditable allocation problem. Chemical holdouts reveal random-split optimism, calibrated intervals enforce electrical confidence, and sequential acquisition recovers most high-utility candidates after evaluating roughly one quarter of the pool. Package sensitivity shifts emphasis from excessive conductivity to expansion, anisotropy, interfaces, and processing. Al_2O_3, BN, and Si_3N_4 are best supported; AlN remains promising but less certain. Higher-fidelity phonon, reliability, interface, and experimental evidence can update the workflow.

Acknowledgement

The author acknowledges the Materials Project, matminer, pymatgen, and dataset maintainers. Funding details must be completed before submission.

Author contribution

H.Z. conceived the study, performed the workflow and analysis, interpreted results, and wrote the manuscript. Additional contributions must be recorded before submission.

Competing interests

The author declares no competing interests; this statement requires review before submission.

References

[1] Waldrop, M. M. (2016). The chips are down for Moore's law. *Nature, 530*, 144–147.
[2] Sheikh, F., Nagisetty, R., Karnik, T., & Kehlet, D. (2021). 2.5D and 3D heterogeneous integration: Emerging applications. *IEEE Solid-State Circuits Magazine, 13*, 77–87.
[3] Pop, E. (2010). Energy dissipation and transport in nanoscale devices. *Nano Research, 3*, 147–169.
[4] Tuckerman, D. B., & Pease, R. F. W. (1981). High-performance heat sinking for VLSI. *IEEE Electron Device Letters, 2*, 126–129.
[5] Merchant, A., et al. (2023). Scaling deep learning for materials discovery. *Nature, 624*, 80–85.
[6] Jain, A., et al. (2013). The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials, 1*, 011002.
[7] Petousis, I., et al. (2017). High-throughput screening of inorganic compounds for novel dielectric and optical materials. *Scientific Data, 4*, 160134.
[8] de Jong, M., et al. (2015). Charting the complete elastic properties of inorganic crystalline compounds. *Scientific Data, 2*, 150009.
[9] Ward, L., et al. (2018). Matminer: An open source toolkit for materials data mining. *Computational Materials Science, 152*, 60–69.
[10] Dunn, A., Wang, Q., Ganose, A., Dopp, D., & Jain, A. (2020). Benchmarking materials property prediction methods: The Matbench test set and Automatminer reference algorithm. *npj Computational Materials, 6*, 138.
[11] Lookman, T., Balachandran, P. V., Xue, D., & Yuan, R. (2019). Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials, 5*, 21.
[12] Balachandran, P. V., Xue, D., Theiler, J., Hogden, J., & Lookman, T. (2016). Adaptive strategies for materials design using uncertainties. *Scientific Reports, 6*, 19660.
[13] Kusne, A. G., et al. (2020). On-the-fly closed-loop autonomous materials discovery via Bayesian active learning. *Nature Communications, 11*, 5966.

DOI: 10.54254/2755-2721/2026.36210

[14] Tranås, R., Løvvik, O. M., & Berland, K. (2023). Lattice thermal conductivity from first principles and active learning with Gaussian process regression. arXiv:2309.06786.
[15] Liu, Y., Xu, J., Zhang, R., Jiang, M., & Luo, T. (2026). Active learning-enabled multi-objective design of thermally conductive and mechanically compliant polymers. arXiv:2603.23494.
[16] Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical Review Letters, 77*, 3865–3868.
[17] Ong, S. P., et al. (2013). Python Materials Genomics (pymatgen): A robust, open-source Python library for materials analysis. *Computational Materials Science, 68*, 314–319.
[18] Anderson, O. L. (1963). A simplified method for calculating the Debye temperature from elastic constants. *Journal of Physics and Chemistry of Solids, 24*, 909–917.
[19] Toher, C., et al. (2014). High-throughput computational screening of thermal conductivity, Debye temperature, and Grüneisen parameter using a quasi-harmonic Debye model. *Physical Review B, 90*, 174107.
[20] Slack, G. A. (1973). Nonmetallic crystals with high thermal conductivity. *Journal of Physics and Chemistry of Solids, 34*, 321–335.
[21] Cahill, D. G., Watson, S. K., & Pohl, R. O. (1992). Lower limit to the thermal conductivity of disordered crystals. *Physical Review B, 46*, 6131–6140.
[22] Xia, Y., et al. (2023). A unified understanding of minimum lattice thermal conductivity. *Proceedings of the National Academy of Sciences USA, 120*, e2302541120.
[23] Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning, 63*, 3–42.
[24] Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.
[25] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.
[26] Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in Neural Information Processing Systems, 30*.
[27] Tran, K., et al. (2020). Methods for comparing uncertainty quantifications for material property predictions. *Machine Learning: Science and Technology, 1*, 025006.
[28] Zhu, K., Huang, D., Costero, L., & Atienza, D. (2025). 3D-ICE 4.0: Accurate and efficient thermal modelling for 2.5D/3D heterogeneous chiplet systems. arXiv:2512.05823.
[29] Gopakumar, A., Pal, K., & Wolverton, C. (2022). Identification of high-dielectric-constant compounds from statistical design. arXiv:2206.04750.
[30] Qu, J., Zageceta, D., Zhang, W., & Zhu, Q. (2019). High-dielectric ternary oxides from crystal-structure prediction and high-throughput screening. arXiv:1909.04195.
[31] Choudhary, K., et al. (2020). High-throughput density-functional perturbation theory and machine-learning predictions of infrared, piezoelectric and dielectric responses. *Scientific Data, 7*, 351.