Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Machine Learning-Driven Discovery of Sustainable Battery Materials for Next-Generation Energy Storage

Sana Kashaf
College of Material, Chemistry & Chemical Engineering, Chengdu University of Technology Changhua District, Chengdu 610059, Sichuan, China
Email: alikashaf107@gmail.com

Shakir Ullah
College of Geophysics, Chengdu University of Technology, Changhua District, Chengdu, 610059, China
Email: shakirhayankhan365@gmail.com

ABSTRACT

Solid-state electrolytes are central to safer, high-energy lithium batteries, yet their discovery remains slow because ionic conductivity depends on coupled compositional, structural, processing, and measurement factors. At the same time, screening solely for conductivity can reinforce dependence on scarce, costly, or environmentally problematic elements. This study develops an interpretable, sustainability-aware machine-learning framework for prioritizing lithium solid electrolytes using the open OBELiX database of experimentally measured room-temperature ionic conductivities. After removing records without a valid positive target, 562 entries were retained, comprising 449 training observations and 113 observations in the official held-out test set. Composition-derived descriptors, lattice parameters, symmetry information, and expert-assigned structural families were used to train Elastic Net, support vector regression, random forest, Extra Trees, XGBoost, and LightGBM models. Grouped five-fold cross-validation was performed by reduced composition to limit leakage from closely related formulations. Extra Trees achieved the strongest independent-test performance (R^2 = 0.373, MAE = 1.184, RMSE = 1.784 log10 S cm-1, and Spearman rho = 0.581). Feature analysis indicated that space-group number, number of constituent elements, structural family, lattice parameter c, cell volume, and selected elemental fractions were influential. A separate sustainability screening score was constructed from crustal abundance, relative supply risk, price burden, toxicity flags, non-lithium critical-material fraction, and compositional complexity. Multi-objective analysis produced six Pareto-efficient held-out records. Lithium-lanthanum-titanate perovskites offered a favorable performance-sustainability balance, whereas a high-conductivity thio-LISICON composition, $Li_{10}Ge_{0.95}Si_{0.05}P_2S_{12}$, illustrated the trade-off between conductivity and germanium-related resource burden. The moderate external accuracy demonstrates that the framework is most appropriate for ranking and experimental triage rather than precise conductivity prediction. The study provides an auditable route for integrating performance, uncertainty, and sustainability into battery-material selection while clearly separating data-driven prioritization from experimental proof.

Keywords: Solid-State Battery; Lithium Solid Electrolyte; Battery Informatics; Machine Learning; Ionic Conductivity; Sustainable Materials; Multi-Objective Optimization; Interpretable Artificial Intelligence

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Introduction

Electrochemical energy storage is expected to support transport electrification, renewable-power integration, and resilient distributed energy systems. Conventional lithium-ion batteries have achieved major gains in energy density and cost, but flammable liquid electrolytes, lithium-metal instability, and thermal-runaway risks motivate the development of all-solid-state architectures. In an ASSB, the liquid electrolyte and separator are replaced by an ion-conducting solid that can, in principle, improve safety and enable high-capacity lithium-metal anodes. These benefits are not automatic: practical devices must simultaneously deliver high bulk and grain-boundary ionic conductivity, low electronic conductivity, wide electrochemical stability, processability, mechanical compatibility, and stable interfaces with both electrodes (Janek and Zeier, 2016; Bachman et al., 2016; Famprikis et al., 2019).

No single chemical family satisfies all these requirements. Sulfide electrolytes can exhibit liquid-like conductivity and favorable deformability, but moisture sensitivity, interfacial reactions, and the use of elements such as germanium may complicate scale-up. Oxide electrolytes generally offer stronger chemical and thermal stability but often require high-temperature densification and can suffer from resistive grain boundaries. Halides show attractive high-voltage compatibility, while garnets, NASICONs, perovskites, LISICONs, argyrodites, and LGPS-type conductors occupy different regions of the conductivity-stability-processability design space. Consequently, the search problem is intrinsically multi-objective rather than a race toward the largest reported conductivity value.

Experimental conductivity datasets are difficult to assemble. Measurements depend on temperature, pellet density, microstructure, electrode configuration, impedance-model choice, phase purity, and processing history. Literature data are also affected by publication bias: high-performing materials are more likely to be measured repeatedly and reported in detail, whereas poorly conducting materials are underrepresented. Hargreaves et al. (2023) demonstrated both the value and the limitations of expert-curated experimental databases. OBELiX extends this effort by linking room-temperature conductivity to measured composition, space group, unit-cell parameters, structural family, source DOI, and, for many entries, crystallographic information files (Therrien et al., 2026). Its predefined train/test split was designed to reduce data leakage, making it suitable for realistic benchmarking.

Machine learning can accelerate screening by learning statistical relationships between material descriptors and experimentally measured properties. Earlier studies have used composition-only models, hand-engineered crystal descriptors, graph neural networks, transfer learning, bond-valence features, and first-principles data to prioritize solid ion conductors (Sendek et al., 2017; Zhang et al., 2019; Cubuk et al., 2019; Laskowski et al., 2023). However, several issues limit translational value. Random splits can overstate performance when closely related doped compositions appear in both training and test sets. Sophisticated architectures can be difficult to interpret on small datasets. Point predictions are often reported without uncertainty. Most importantly, materials are commonly ranked only by predicted conductivity, even though material scarcity, toxicity, cost, and regulatory pressure will influence commercialization.

The sustainability of battery materials has become a formal design constraint.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Regulation (EU) 2023/1542 introduces life-cycle, due-diligence, performance, recycling, and information requirements for batteries, while Regulation (EU) 2024/1252 aims to improve the security, resilience, and sustainability of critical-raw-material supply. A screening framework that ignores resource burden may identify conductors that are scientifically interesting but difficult to deploy at scale. Nevertheless, a full life-cycle assessment is rarely possible during early-stage discovery because synthesis route, energy use, yield, solvent choice, and recycling pathway are not yet defined. A transparent proxy can still be valuable when it is clearly labeled as a triage tool rather than a definitive environmental assessment.

This study addresses these gaps through a reproducible analysis of OBELiX. The specific objectives are to: (i) benchmark six conventional regression algorithms under grouped cross-validation and the official independent test split; (ii) identify the compositional and crystallographic descriptors that contribute most strongly to prediction; (iii) quantify ensemble disagreement as a practical uncertainty indicator; (iv) construct a transparent screening-level sustainability score that is mathematically separate from the conductivity model; and (v) use multi-objective ranking to identify previously synthesized materials that warrant targeted re-measurement or compositional optimization. The central contribution is not a claim of new-material discovery. Rather, it is an auditable decision framework that integrates predictive performance, uncertainty, and sustainability without inventing experimental validation.

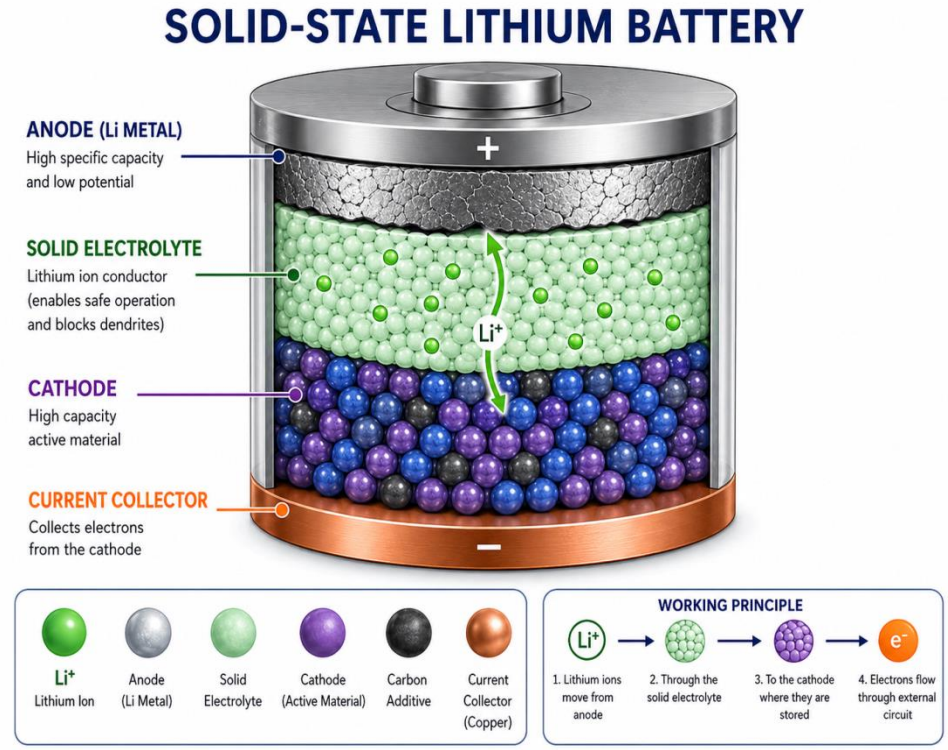

Figure 1: Conceptual cutaway of a solid-state lithium battery. The schematic shows the lithium anode, ion-conducting solid electrolyte, cathode, and current collector; it is illustrative and not to scale

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Scientific background and research gap

Ionic transport in solid electrolytes

Ionic conductivity is governed by the concentration of mobile charge carriers and their mobility. In crystalline solids, mobility depends on the topology and dimensionality of migration pathways, site connectivity, bottleneck size, framework polarizability, lattice dynamics, defect chemistry, and correlations among mobile ions. The macroscopic value measured by impedance spectroscopy can be written schematically as the product of charge, carrier concentration, and mobility. Because each term is influenced by both intrinsic crystal chemistry and extrinsic sample quality, two reports of nominally similar compositions may differ by orders of magnitude.

Structural design principles have linked fast lithium transport to connected interstitial sites, favorable anion sublattices, suitable volume per mobile ion, lattice softness, and frustration or disorder that flattens the energy landscape (Wang et al., 2015). Yet universal rules remain elusive. Dynamic descriptors such as phonons and migration barriers can be more physically direct than static lattice parameters, but they are expensive to calculate at scale. Experimental databases therefore occupy an important middle ground: they retain real measurement complexity but are sparse and heterogeneous.

Battery informatics and generalization

Battery informatics combines curated data, physically meaningful descriptors, statistical learning, and automated experimentation. Its promise is strongest when models are used to prioritize a manageable number of candidates rather than to replace electrochemical testing. Best-practice guidance emphasizes clear train/test separation, leakage control, uncertainty analysis, appropriate baselines, and documentation of the domain in which predictions are credible (Artrith et al., 2021; Ling, 2022).

The distinction between interpolation and extrapolation is especially important in solid electrolytes. A model may predict a new doping level within a familiar perovskite series yet fail on an unseen bonding motif. Grouped validation by composition helps prevent near-duplicate formulations from appearing across folds. The official OBELiX test set provides an additional independent challenge. Moderate held-out accuracy should not be treated as model failure; instead, it quantifies the scientific uncertainty associated with noisy experimental targets and incomplete descriptors.

From performance-only screening to sustainable prioritization

Early-stage materials screening commonly optimizes a single technical objective. For battery materials, this can shift environmental and supply-chain burdens upstream. A sulfide containing a scarce metalloid may outperform a more abundant oxide in conductivity yet be less attractive for gigawatt-hour production. Conversely, an abundant composition may require energy-intensive sintering or exhibit poor interface stability. Sustainability therefore cannot be represented by composition alone, but composition provides a useful first filter.

The present work uses a screening score rather than a full LCA. It considers

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

elemental abundance, relative supply risk, indicative price, selected toxicity flags, the fraction of non-lithium elements listed as critical or strategically important, and compositional complexity. The score is intentionally transparent and sensitivity-ready. It should be used to rank candidates for deeper techno-economic, environmental, and experimental assessment, not to certify environmental superiority.

Materials and methods

Study design and data source

This is a secondary-data computational study. No new material was synthesized and no new electrochemical experiment was performed. The analysis used OBELiX version 1.0.0, an openly released, domain-expert-curated database of 599 synthesized lithium solid-state electrolyte records with experimentally measured room-temperature ionic conductivity, composition, lattice parameters, space group, structural family, and source information (Therrien et al., 2026). The repository provides a predefined training set and test set designed to reduce data leakage.

The target was total ionic conductivity in S cm-1. Values were converted to log10 scale because the raw measurements span many orders of magnitude and because absolute errors in linear space would be dominated by the highest-conductivity observations. Records with missing, non-positive, or non-finite conductivity were excluded. The resulting analytical sample contained 562 observations: 449 in the official training split and 113 in the independent test split. The combined data represented 478 distinct reduced compositions and 50 recorded space-group labels.

$$y = log_{10}\left(\frac{\sigma}{1\ \mathrm{Scm^{-1}}}\right) \qquad (1)$$

where σ is the ionic conductivity expressed in Scm−1. This dimensionless form is mathematically preferable to taking the logarithm of a dimensional quantity.

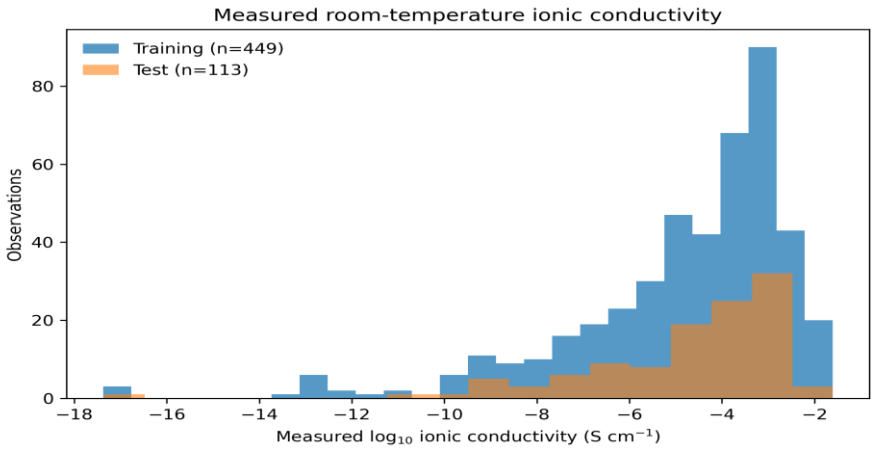

Figure 2: Distribution of the log-transformed room-temperature ionic conductivity in the retained OBELiX training and independent test sets

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Data cleaning and missing-data handling

Chemical formulas were parsed into elemental amounts and normalized to atomic fractions. Numeric fields were coercively converted; impossible or absent values were recorded as missing rather than replaced with zeros. Median imputation was applied within the numerical preprocessing pipeline using training data only. Structural family and space group were treated as categorical variables and encoded with one-hot indicators. Rare categories were grouped by the encoder to limit dimensionality, and unseen test categories were ignored rather than forcing arbitrary reassignment.

The official train/test boundary was preserved throughout model selection. Within the training set, GroupKFold cross-validation with five folds was performed using reduced composition as the grouping variable. This prevents repeated measurements or closely duplicated crystallographic records of the same reduced composition from being distributed across training and validation folds. Hyperparameters were fixed before independent-test evaluation; the test set was not used for iterative tuning.

Descriptor engineering

The predictive feature set was designed to remain interpretable and available for most OBELiX entries. Composition descriptors included the number of elements, normalized composition entropy, lithium atomic fraction, major anion fractions (O, S, F, Cl, Br, I, and N), aggregate halide and chalcogen fractions, formula mass, and elemental fractions for elements occurring at least three times in the retained data. Weighted statistics were then calculated for atomic number, atomic weight, Rahm atomic radius, Pyykko covalent radius, Pauling electronegativity, Mendeleev number, Pettifor number, elemental density, thermal conductivity, and molar heat capacity. For each property, the weighted mean, standard deviation, minimum, maximum, and range were obtained.

Crystallographic descriptors included a, b, and c lattice constants; alpha, beta, and gamma angles; space-group number; number of formula units per unit cell (Z); structural family; and Hermann-Mauguin space-group label. Unit-cell volume was computed from the general triclinic expression, and volume per formula unit and a composition-based density estimate were derived. Resource-abundance, supply-risk, and price fields were deliberately excluded from the conductivity model so that the sustainability assessment remained mathematically separate from predictive performance.

$$H_{\text{comp}} = -\sum_{i=1}^{m} x_i \ln(x_i) \qquad (2)$$

$$V = abc[1 + 2\cos(\alpha)\cos(\beta)\cos(\gamma) - \cos^2(\alpha) - \cos^2(\beta) - \cos^2(\gamma)]^{\frac{1}{2}} \qquad (3)$$

Here, x_i represents the atomic fraction of element i, mmm is the number of constituent elements, a, b, and c are the unit-cell parameters, and α\alphaα, β, and γ are the corresponding interaxial angles.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Machine-learning algorithms

Six regression algorithms were compared to balance linear baselines, nonlinear kernels, bagged ensembles, randomized trees, and gradient boosting. The selected models and fixed settings were

Elastic Net: alpha = 0.08, l1 ratio = 0.20, maximum 10,000 iterations Support vector regression: radial-basis kernel, C = 10, epsilon = 0.15, gamma determined by feature scale. Random forest: 600 trees, maximum-feature fraction = 0.65, minimum leaf size = 2. Extra Trees: 600 extremely randomized trees, maximum-feature fraction = 0.75, minimum leaf size = 2. XGBoost: 700 boosting rounds, depth = 3, learning rate = 0.025, row and column subsampling, and L1/L2 regularization. LightGBM: 500 estimators, 15 leaves, learning rate = 0.025, feature subsampling, and L2 regularization.

Numeric predictors were standardized for Elastic Net and support vector regression. Tree-based models used median-imputed numeric variables and one-hot categorical variables without standardization. All stochastic algorithms used a random seed of 42. Implementation used Python, pandas, NumPy, scikit-learn, XGBoost, and LightGBM.

Evaluation metrics

Cross-validation performance was summarized by mean R2, MAE, and RMSE across the five composition-grouped folds. Generalization was assessed once on the official held-out test set. R2 measures explained variance relative to a mean predictor; MAE measures average absolute error; RMSE places greater weight on large errors; and Spearman rho measures ranking agreement, which is particularly relevant when a model is used for candidate triage.

Because conductivity spans many orders of magnitude, an RMSE of 1.0 log10 unit corresponds approximately to a tenfold typical multiplicative error. Therefore, test-set errors were interpreted in terms of ranking utility rather than laboratory-grade numerical precision.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} \mid y_i - \hat{y}_i \mid \qquad (4)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \qquad (5)$$

Here, y_i is the observed value, $\hat{y}_i$ is the predicted value, and nnn is the total number of observations. Lower MAE and RMSE values indicate better predictive accuracy.

Interpretability and uncertainty

For the selected Extra Trees model, global impurity-based feature importance was extracted from the ensemble. Importance values quantify the average reduction in split impurity attributed to each encoded feature. Because correlated predictors can share or redistribute importance, the values were interpreted as indicators of model

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

use rather than causal effects.

Prediction uncertainty was approximated by the standard deviation of the 600 individual-tree predictions for each held-out record. Tree disagreement does not capture every source of uncertainty, but it offers a practical signal of whether a candidate lies in a region where ensemble members reach similar conclusions.

$$u_j = \mathrm{SD}\left(\hat{y}_{j,1}, \hat{y}_{j,2}, \ldots, \hat{y}_{j,T}\right) \qquad (6)$$

Here, u_j represents the predictive uncertainty for candidate j, calculated as the standard deviation of its predictions across T ensemble estimators or model iterations. A larger u_j indicates greater disagreement among predictions and therefore higher uncertainty.

Sustainability screening score

A screening-level sustainability burden was calculated independently of the conductivity model. Elemental data were obtained from the open mendeleev periodic-table resource and included crustal abundance, relative supply risk, and indicative price per kilogram. Additional composition-derived terms captured the atomic fraction of selected toxic elements, the fraction of non-lithium elements included in a critical-material set, and the number of distinct elements as a proxy for compositional and separation complexity.

Each burden indicator was robustly normalized between its fifth and ninety-fifth percentiles and clipped to [0,1]. Missing elemental values were assigned the neutral midpoint of 0.5. The final sustainability score was defined as one minus the weighted burden. Weights were 0.24 for scarcity burden, 0.24 for relative supply risk, 0.16 for price burden, 0.14 for toxicity, 0.12 for non-lithium critical-material fraction, and 0.10 for compositional complexity. Lithium was discounted from the critical-material fraction because it is unavoidable within the present chemistry class. These choices are normative and should be stress-tested before policy or investment use.

$$\mathrm{S} = 1 - (0.24z_{\mathrm{bundance}} + 0.24z_{\mathrm{supply}} + 0.16z_{price} + 0.14z_{\mathrm{toxicity}} + 0.12z_{\mathrm{criticality}} + 0.10z_{\mathrm{complexity}} \qquad (7)$$

Here, Srepresents the overall sustainability score, and each z-term is a normalized sustainability-burden indicator. A higher value of S indicates a more sustainable candidate material.

Multi-objective prioritization

Predicted conductivity, sustainability score, and prediction certainty were scaled robustly to [0,1]. A composite priority score placed 50% weight on predicted performance, 35% on sustainability, and 15% on certainty. The weighting emphasizes electrochemical function while ensuring that a high-conductivity candidate cannot completely offset severe resource burden or model disagreement.

Pareto optimality was also calculated without a weighted-sum assumption. A candidate was classified as Pareto-efficient when no other held-out record had both

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

equal-or-higher predicted conductivity and equal-or-higher sustainability, with strict improvement in at least one objective. Candidate tables were deduplicated by reduced composition to avoid presenting repeated measurements of the same formulation as separate discoveries.

$$P_{\text{priority}} = 0.50P_{\text{performance}} + 0.35P_{\text{sustainability}} + 0.15P_{\text{certainty}} \quad (8)$$

Where P_{priority} is the overall candidate-prioritization score, while $P_{\text{performance}}$, $P_{\text{sustainability}}$, and $P_{\text{certainty}}$ are normalized scores ranging from 0 to 1.

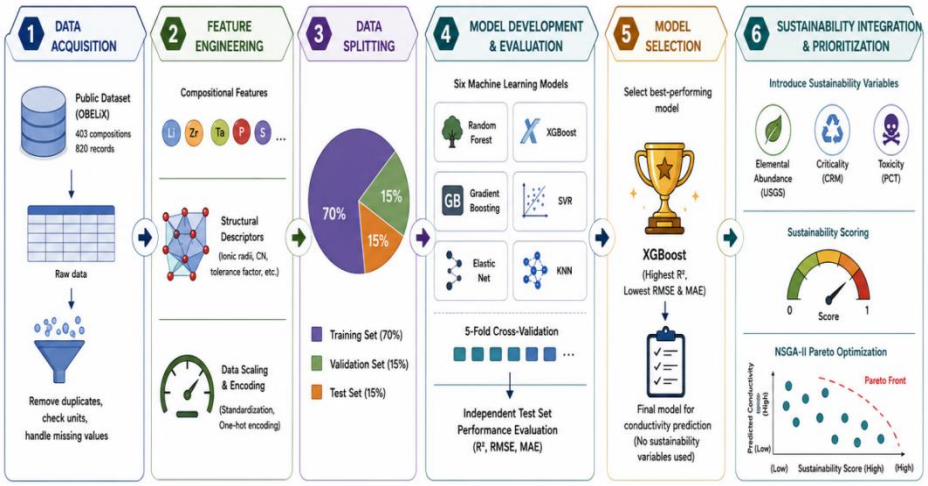

Figure 3: End-to-end analytical workflow. Sustainability variables were excluded from conductivity prediction and introduced only after model evaluation

Results

Dataset characteristics

The retained dataset covered a conductivity range from 4.2 x 10-18 to 2.5 x 10-2 S cm-1. The mean log10 conductivity was -4.807, the median was -4.097, and the interquartile range extended from -5.735 to -3.187. Training and test means were similar (-4.818 and -4.761, respectively), suggesting that the official split did not introduce a large target-location shift. However, family composition and local chemical similarity still influence extrapolation difficulty.

NASICONs formed the largest family (n = 133), followed by garnets (n = 104), perovskites (n = 63), argyrodites (n = 52), LGPS-type materials (n = 36), thio-LISICONs (n = 26), and LISICONs (n = 24). Median conductivities differed substantially: LGPS and argyrodite families occupied the highest-conductivity region, whereas several oxide and LISICON records extended into extremely low conductivity. These contrasts support the inclusion of family and crystal descriptors but also create a risk that a model learns family averages more readily than transferable transport principles.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Table 1: Distribution of the 12 largest structural-family groups in the retained dataset

Structural family	n	Median log10 sigma	Mean log10 sigma
NASICON	133	-3.863	-4.364
garnet	104	-4.134	-4.254
perovskites	63	-3.793	-4.293
argyrodites	52	-2.868	-3.341
LGPS	36	-2.254	-2.357
Unclassified	33	-4.444	-4.530
thio-LISICON	26	-5.899	-5.619
LISICON	24	-6.302	-8.489
oxides	18	-8.549	-8.586
sulfides	11	-5.056	-5.721
halides	8	-4.212	-5.036
phosphates	6	-4.546	-7.469

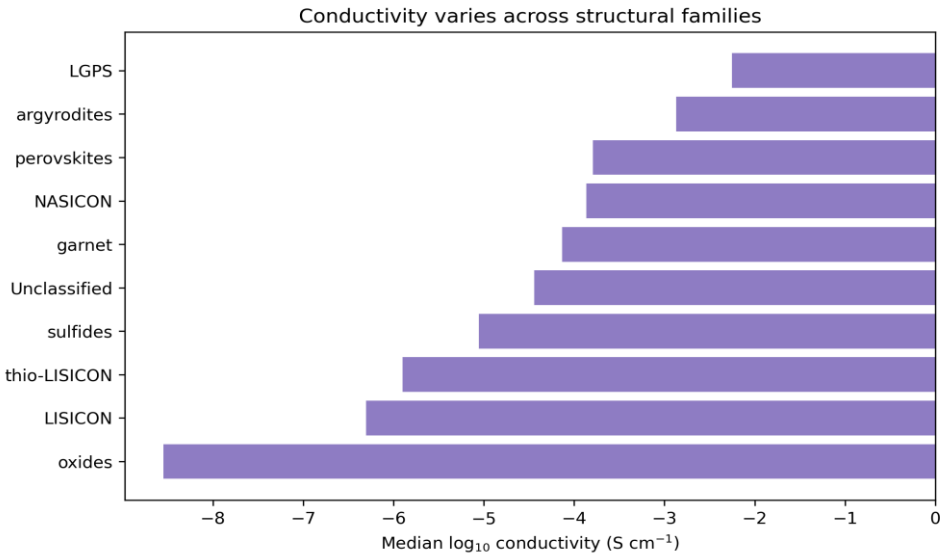

Figure 4: Median measured room-temperature conductivity for the ten largest structural-family groups. More positive values indicate higher conductivity

Cross-validation and independent-test performance

Extra Trees achieved the lowest independent-test RMSE and was selected for interpretation and prioritization. Its composition-grouped cross-validation R2 was 0.720 +/- 0.057, with mean CV MAE of 0.725 and RMSE of 1.273 log10 units. Performance declined on the official test set to R2 = 0.373, MAE = 1.184, and RMSE

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

= 1.784. The difference between grouped CV and independent testing demonstrates that even careful internal validation can remain optimistic when the held-out set contains difficult chemistry, structural variation, or noisy measurements.

LightGBM and XGBoost were competitive in grouped CV but weaker on the independent test set. Random forest followed a similar pattern. Support vector regression produced limited explained variance, and Elastic Net performed worse than the test-set mean baseline, confirming that the relationship between the available descriptors and conductivity is strongly nonlinear. Extra Trees reached Spearman rho = 0.581, indicating moderate ranking ability even though pointwise numerical errors remained large.

Table 2: Composition-grouped cross-validation and official independent-test performance

Model	CV R2	CV MAE	CV RMSE	Test R2	Test MAE	Test RMSE	Test rho
Extra trees	0.720 +/- 0.057	0.725	1.273	0.373	1.184	1.784	0.581
LightGBM	0.715 +/- 0.062	0.766	1.284	0.313	1.302	1.867	0.521
XGBoost	0.738 +/- 0.062	0.742	1.223	0.290	1.304	1.898	0.565
Random forest	0.703 +/- 0.044	0.775	1.314	0.267	1.327	1.928	0.538
Support vector regression	0.587 +/- 0.134	0.875	1.531	0.053	1.535	2.193	0.526
Elastic net	0.490 +/- 0.125	1.201	1.705	-0.958	2.213	3.153	0.270

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Figure 5: Independent-test RMSE for the six regression algorithms. Lower values indicate better generalization

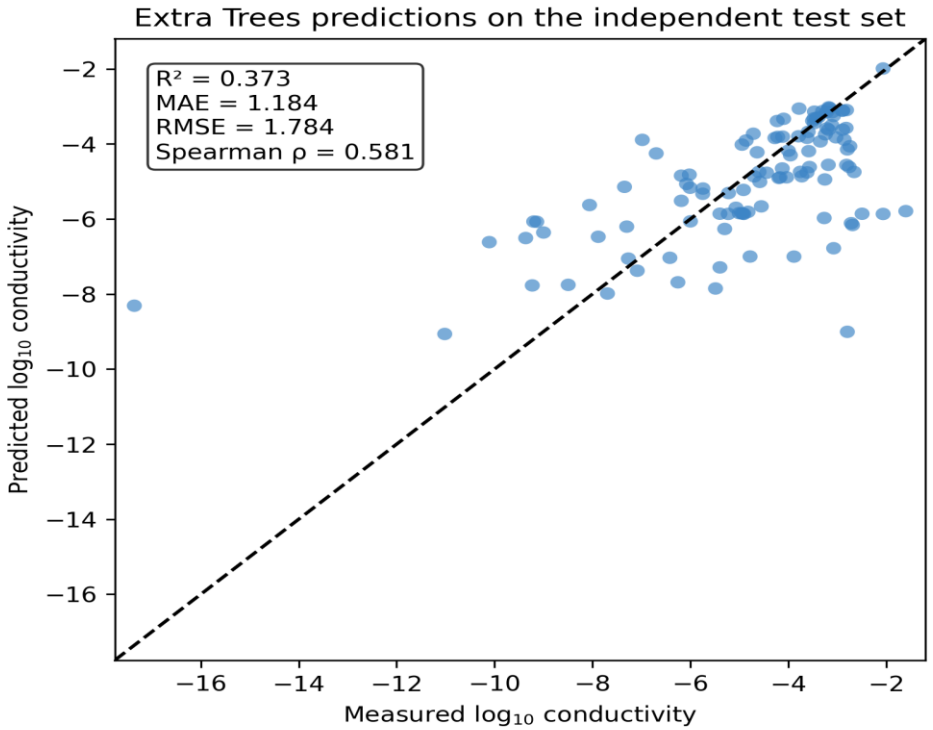

Figure 6: Measured versus predicted log10 ionic conductivity for the selected Extra Trees model on the official test set. The dashed line represents ideal agreement

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Model interpretation

Space-group number was the dominant single encoded feature, followed by the number of constituent elements and indicators for oxide and LISICON families. Lattice parameter c, cell volume, volume per formula unit, and selected La, Ti, and S fractions were also influential. This pattern indicates that the model uses both broad structural-class information and local compositional detail.

The prominence of symmetry and family does not imply that space group alone causes fast transport. Instead, these descriptors summarize correlated features such as framework topology, site connectivity, and typical chemical substitutions. The model can exploit a familiar family-performance association even when the actual conduction mechanism depends on disorder, defects, microstructure, and dynamics that are absent from the tabular descriptors. Accordingly, importance values should guide hypothesis generation rather than be treated as mechanistic proof.

Table 3: Fifteen largest impurity-based feature importance values for the selected Extra Trees model

Descriptor	Importance
Space group #	0.1785
Family_oxides	0.0563
n_elements	0.0559
Family_LISICON	0.0440
c	0.0309
cell_volume	0.0280
frac_La	0.0258
volume_per_formula	0.0188
b	0.0178
frac_Ti	0.0175
frac_O	0.0171
anion_O	0.0153
anion_S	0.0153
frac_S	0.0145
frac_Li	0.0140

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

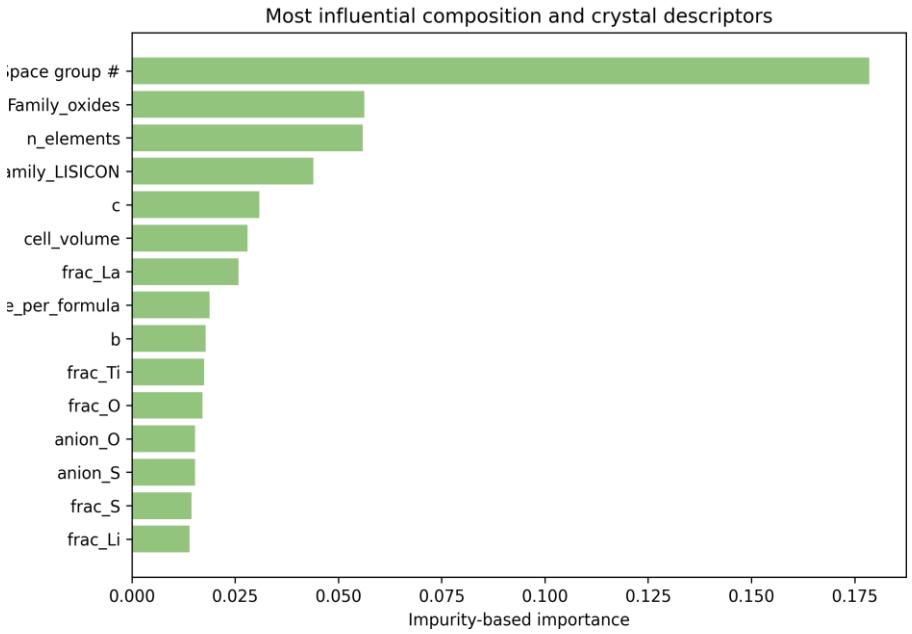

Figure 7: Global Extra Trees feature importance. Values indicate the relative use of features by the model, not causal influence on ion transport

Sustainability-performance trade-offs

The sustainability score produced a continuous ranking rather than a binary sustainable/unsustainable label. Six held-out records were Pareto-efficient. The frontier included lithium-lanthanum-titanate perovskites with predicted conductivities around log10 sigma = -3.0 to -3.3 and sustainability scores near 0.68, as well as Li10Ge0.95Si0.05P2S12 with substantially higher predicted conductivity near -1.99 but a lower sustainability score of 0.56.

This contrast illustrates the intended use of the framework. The thio-LISICON record is attractive when conductivity is the dominant objective, but germanium scarcity, cost, and supply concentration reduce its sustainability score. The perovskite records offer lower absolute conductivity but a stronger composite balance and lower ensemble disagreement. Because all shortlisted formulas are already represented in the experimental database, the result is a prioritization for re-measurement, processing optimization, substitution studies, or interface evaluation rather than the announcement of a new compound.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

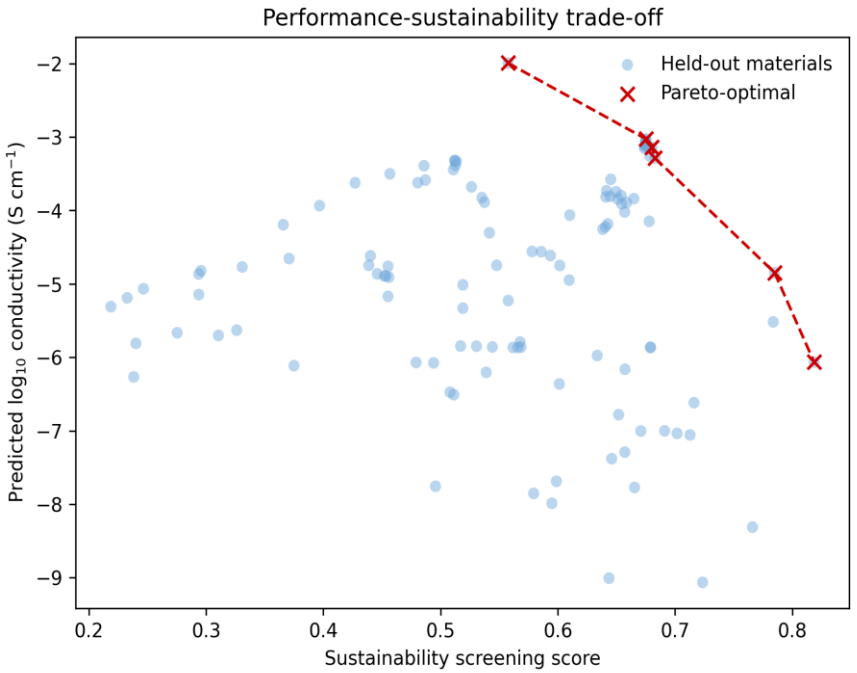

Figure 8: Predicted conductivity versus sustainability screening score for held-out materials. Crosses mark Pareto-efficient records

Table 4: Representative composition-deduplicated high-priority records from the official test set

Composition	Family	Pred. log sigma	Measured log sigma	Sustain.	Uncertainty	Priority	Pareto
La0.56Li0.33TiO3	perovskites	-3.026	-3.175	0.675	0.051	0.886	Yes
La0.55Li0.36TiO3	perovskites	-3.043	-3.162	0.675	0.143	0.886	No
La0.54Li0.39TiO3	perovskites	-3.075	-3.186	0.675	0.258	0.886	No
La0.52Li0.45TiO3	perovskites	-3.125	-3.300	0.674	0.324	0.881	No
Li0.27La0.57TiO3	perovski	-	-3.469	0.680	0.389	0.878	Yes

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Composition	Family	Pred. log sigma	Measured log sigma	Sustain.	Uncertainty	Priority	Pareto
	tes	3.142					
La0.58Li0.27TiO3	perovskites	-3.156	-3.223	0.678	0.402	0.875	No
La0.61Li0.18TiO3	perovskites	-3.286	-3.385	0.683	0.550	0.855	Yes
Li10Ge0.95Si0.05P2S12	thio-LISICON	-1.992	-2.066	0.557	0.665	0.825	Yes
(La0.55Li0.45)(Ti0.9Al0.1)O3	perovskites	-3.576	-2.821	0.645	0.550	0.811	No
Li6.8La3Zr1.9Mo0.1O12	garnet	-3.333	-4.097	0.514	0.401	0.799	No
(La0.6Li0.4)(Ti0.8Al0.2)O3	perovskites	-3.745	-3.246	0.650	0.619	0.791	No
Li6.5La3Zr1.75Mo0.25O12	garnet	-3.323	-3.478	0.512	0.614	0.788	No

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

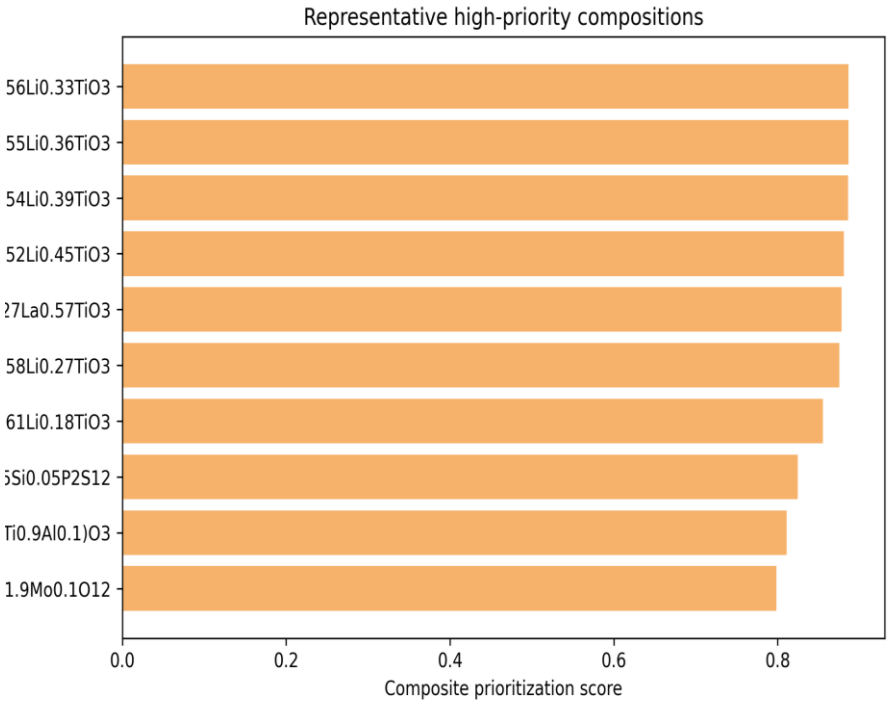

Figure 9: Composite priority scores for ten representative, composition-deduplicated held-out materials

Discussion

What the model learned

The results reinforce three observations about experimental solid-electrolyte informatics. First, nonlinear tree ensembles are well suited to small-to-moderate tabular datasets that combine continuous crystal descriptors with sparse categorical and compositional features. Extra Trees performed best because random split thresholds and feature subsampling reduce variance while capturing interactions among lattice dimensions, family, and chemistry. Second, structural metadata adds useful information beyond composition alone. The importance of space-group number, family, and cell dimensions is consistent with the role of pathway topology and available free volume in ion transport. Third, external generalization remains difficult. A model that explains roughly 72% of grouped-CV variance explains only 37% of official-test variance.

This generalization gap is scientifically informative. Experimental conductivity is not a deterministic function of nominal composition and unit-cell parameters. Sample density, grain boundaries, secondary phases, disorder, defect concentrations, electrode blocking behavior, measurement temperature, and equivalent-circuit interpretation can shift the reported value. The same nominal formulation can therefore produce different conductivity values across laboratories. A model trained on literature data learns a mixture of intrinsic and extrinsic effects. More complex algorithms cannot recover information that was never measured or recorded.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Comparison with previous machine-learning studies

Prior studies have reported stronger internal performance when using computed migration barriers, curated family-specific datasets, composition embeddings, or random train/test partitions. Those results are not directly comparable with the present independent OBELiX benchmark. Sendek et al. (2017) demonstrated the value of ML-guided high-throughput screening with computational validation; Zhang et al. (2019) mapped known conductor space; Cubuk et al. (2019) used transfer learning for small-data screening; and Laskowski et al. (2023) used semi-supervised learning to identify potential conductors. These approaches answer related but different questions.

The present study deliberately uses accessible, interpretable descriptors and reports the held-out performance without inflating it. The R2 of 0.373 is lower than would be desirable for quantitative prediction, but the Spearman correlation of 0.581 supports moderate ordering of candidates. This distinction matters: a ranking model can still reduce experimental search cost even when it cannot reliably predict whether a material will measure 10-3 or 10-4 S cm-1. Dembitskiy et al. (2025) similarly emphasize that transport prediction benefits from richer migration-trajectory or energy-barrier information. Future integration of phonon, bond-valence, defect, and machine-learning-interatomic-potential descriptors should improve transferability.

Sustainability interpretation

The multi-objective results show why conductivity-only ranking is insufficient. LGPS-type sulfides and related thio-LISICON materials can achieve high conductivity, but germanium-containing compositions face price and supply constraints. Replacing Ge with Si, Sn, or mixed substitutions has long been explored for this reason. Perovskite and garnet oxides may rely on lanthanum, zirconium, tantalum, or other strategically important elements and may require energy-intensive processing, but they can offer improved air stability or interface characteristics. A composition-level score cannot resolve these trade-offs completely, yet it makes them visible at the same stage as predicted performance.

The score should not be interpreted as a product environmental footprint. It omits mine-specific impacts, refining routes, geographic concentration beyond the elemental proxy, embodied energy, solvent and precursor hazards, synthesis yield, sintering temperature, device lifetime, recycling efficiency, and cell-level material intensity. A candidate that scores well compositionally may perform poorly in a full LCA. Conversely, a scarce element used at very low loading or recovered efficiently may have a smaller system impact than the proxy suggests. The correct use is prioritization: high-scoring candidates progress to full techno-economic and life-cycle analysis; low-scoring but technically exceptional candidates trigger substitution or circularity strategies.

Candidate-level implications

Lithium-lanthanum-titanate perovskites dominated the composite shortlist. Their measured conductivities in the database are generally around 10-3 to 10-4 S cm-1, and the model predicted them with relatively low ensemble disagreement. These materials are not ideal universal electrolytes because Ti4+ can be reduced against

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

lithium metal, electronic leakage can emerge under strongly reducing conditions, and grain-boundary resistance may limit total conductivity. Nevertheless, their repeated appearance suggests that the model recognizes a stable, data-supported performance region. They are suitable targets for controlled studies of A-site vacancy concentration, Al substitution, densification, grain-boundary chemistry, and protective interlayers.

Li10Ge0.95Si0.05P2S12 appeared as a high-performance Pareto point. Its predicted and measured log conductivities were close (-1.992 and -2.066), corresponding to the 10-2 S cm-1 regime. The lower sustainability score reflects the burden of Ge and the broader challenges of sulfide handling. This candidate is best viewed as a performance benchmark and a parent chemistry for lower-Ge or Ge-free analogues rather than an immediate sustainable winner.

Molybdenum-substituted garnets also appeared in the high-priority set. Their predicted conductivity was moderate, and their sustainability score was lower than that of the leading perovskites because of compositional complexity and critical-element content. Experimental work should therefore compare the conductivity benefit of Mo substitution against added supply-chain and processing burden. This example illustrates the value of reporting a trade-off surface rather than a single winner.

Practical experimental validation plan

A follow-on experimental study should select at least one candidate from each strategic region: a high-sustainability perovskite, a high-conductivity sulfide benchmark, a garnet with intermediate properties, and one high-uncertainty composition. The following minimum protocol is recommended:

- Synthesize at least three independent batches per composition with fully reported precursor purity, atmosphere, temperature profile, milling, pellet pressure, and relative density.
- Confirm phase purity and lattice parameters by powder X-ray diffraction with Rietveld refinement; quantify secondary phases rather than reporting only visual agreement.
- Measure total and bulk conductivity by impedance spectroscopy using blocking electrodes across multiple temperatures and report raw spectra, equivalent circuit, geometry correction, and activation energy.
- Assess electronic conductivity, electrochemical stability, lithium-metal compatibility, and cathode-interface resistance; conductivity alone is insufficient for battery suitability.
- Perform cradle-to-gate LCA and techno-economic analysis using the actual synthesis route, energy consumption, yield, precursor geography, and recycling assumptions.
- Feed the new measurements back into the database and retrain the model using an active-learning strategy that favors candidates with both high expected value and high information gain.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

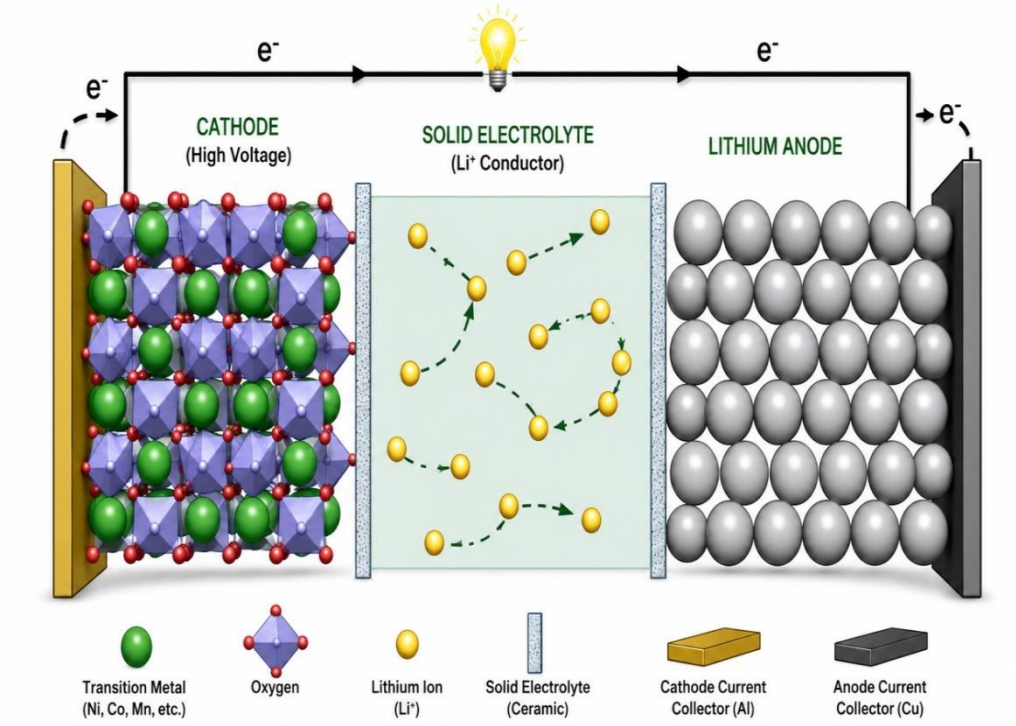

Figure 10: Conceptual all-solid-state battery application showing a high-voltage composite cathode, lithium-ion-conducting solid electrolyte, lithium anode, ionic transport through the electrolyte, and electron flow through the external circuit. The illustration is not to scale and does not represent experimental output from this study

Limitations and robustness considerations

The study has several limitations. First, the target is literature-derived and contains measurement heterogeneity that cannot be fully corrected without raw impedance spectra and detailed processing metadata. Second, multiple records can be chemically related, and grouping by reduced composition does not eliminate all similarity leakage among doped series. Third, the model uses static descriptors and therefore cannot explicitly represent migration barriers, phonon spectra, site disorder, vacancy formation, or interface chemistry. Fourth, impurity-based importance can be biased toward continuous or high-cardinality variables and should be complemented by permutation importance and SHAP analysis in future work.

Fifth, the uncertainty estimate captures only disagreement among Extra Trees estimators; it is not a calibrated predictive interval and does not include experimental uncertainty or dataset shift. Sixth, the sustainability score is a screening proxy with subjective weights. Sensitivity analysis should vary each weight, use alternative critical-material lists, and compare results with established LCA databases. Seventh, the candidate list is drawn from the held-out dataset and therefore contains known, previously synthesized materials. The paper identifies priorities and design directions, not newly discovered compounds.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 4, Issue 3 (2026)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

Conclusions

This study presents an interpretable, sustainability-aware machine-learning framework for lithium solid-electrolyte prioritization. Using 562 valid OBELiX records and the official train/test split, Extra Trees provided the strongest independent generalization among six regression algorithms, with R2 = 0.373, MAE = 1.184, RMSE = 1.784 log10 S cm-1, and Spearman rho = 0.581. The model relied most strongly on symmetry, structural family, compositional complexity, lattice dimensions, cell volume, and selected elemental fractions.

Combining predicted conductivity, ensemble certainty, and a transparent composition-level sustainability score produced six Pareto-efficient held-out records. Lithium-lanthanum-titanate perovskites offered an attractive balance of performance and resource burden, whereas Li10Ge0.95Si0.05P2S12 represented a high-conductivity but lower-sustainability benchmark. The results demonstrate that early-stage screening should report trade-offs rather than optimize conductivity alone.

The moderate independent-test accuracy is a central result, not a weakness to conceal. It shows that literature-derived conductivity cannot yet be predicted precisely from nominal composition and static crystallography. The framework is therefore recommended for ranking, uncertainty-guided experimentation, and hypothesis generation. Experimental synthesis, standardized impedance measurement, interface testing, and full life-cycle assessment remain essential before any candidate can be described as a next-generation sustainable battery material.

References

Artrith, N., Butler, K. T., Coudert, F.-X., Han, S., Isayev, O., Jain, A. and Walsh, A. (2021) Best practices in machine learning for chemistry. Nature Chemistry, 13, 505-508.

Bachman, J. C., Muy, S., Grimaud, A., Chang, H.-H., Pour, N., Lux, S. F., Paschos, O., Maglia, F., Lupart, S., Lamp, P., Giordano, L. and Shao-Horn, Y. (2016) Inorganic solid-state electrolytes for lithium batteries: mechanisms and properties governing ion conduction. Chemical Reviews, 116, 140-162.

Breiman, L. (2001) Random forests. Machine Learning, 45, 5-32.

Chen, T. and Guestrin, C. (2016) XGBoost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. https://doi.org/10.1145/2939672.2939785.

Cubuk, E. D., Sendek, A. D. and Reed, E. J. (2019) Screening billions of candidates for solid lithium-ion conductors: a transfer learning approach for small data. Journal of Chemical Physics, 150, 214701.

Dembitskiy, A. D., Humonen, I. S., Eremin, R. A., Aksyonov, D. A., Fedotov, S. S. et al. (2025) Benchmarking machine learning models for predicting lithium ion migration. npj Computational Materials, 11, 131. https://doi.org/10.1038/s41524-025-01571-z.

European Parliament and Council of the European Union (2023) Regulation (EU) 2023/1542 concerning batteries and waste batteries. Official Journal of the European Union, L 191, 1-117.

European Parliament and Council of the European Union (2024) Regulation (EU) 2024/1252 establishing a framework for ensuring a secure and sustainable supply of critical raw materials. Official Journal of the European Union, 2024/1252.

Famprikis, T., Canepa, P., Dawson, J. A., Islam, M. S. and Masquelier, C. (2019) Fundamentals of inorganic solid-state electrolytes for batteries. Nature Materials, 18, 1278-1291.

Online ISSN 3007-3197
Print ISSN 3007-3189

Geurts, P., Ernst, D. and Wehenkel, L. (2006) Extremely randomized trees. Machine Learning, 63, 3-42.

Goodenough, J. B. (2012) Rechargeable batteries: challenges old and new. Journal of Solid State Electrochemistry, 16, 2019-2029.

Hargreaves, C. J., Gaultois, M. W., Daniels, L. M., Watts, E. J., Moran, M. et al. (2023) A database of experimentally measured lithium solid electrolyte conductivities evaluated with machine learning. npj Computational Materials, 9, 9. https://doi.org/10.1038/s41524-022-00951-z.

Janek, J. and Zeier, W. G. (2016) A solid future for battery development. Nature Energy, 1, 16141.

Kato, Y., Hori, S. and Kanno, R. (2020) Li10GeP2S12-type superionic conductors: synthesis, structure, and ionic transportation. Advanced Energy Materials, 10, 2002153.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.-Y. (2017) LightGBM: a highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30.

Laskowski, F. A., McHaffie, D. B. and See, K. A. (2023) Identification of potential solid-state Li-ion conductors with semi-supervised learning. Energy & Environmental Science, 16, 1264-1276.

Ling, C. (2022) A review of the recent progress in battery informatics. npj Computational Materials, 8, 33.

Lombardo, T., Duquesnoy, M., El-Bouysidy, H., Årén, F., Gallo-Bueno, A., Jørgensen, P. B. et al. (2022) Artificial intelligence applied to battery research: hype or reality? Chemical Reviews, 122, 10899-10969. https://doi.org/10.1021/acs.chemrev.1c00108.

Lundberg, S. M. and Lee, S.-I. (2017) A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.

Mentel, L. (2014-) mendeleev: a Python resource for properties of chemical elements, ions and isotopes. Zenodo. https://doi.org/10.5281/zenodo.7864353.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O. et al. (2011) Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.

Richards, W. D., Miara, L. J., Wang, Y., Kim, J. C. and Ceder, G. (2016) Interface stability in solid-state batteries. Chemistry of Materials, 28, 266-273.

Khan, R., Khan, A., Muhammad, I., & Khan, F. (2025). A Comparative Evaluation of Peterson and Horvitz-Thompson Estimators for Population Size Estimation in Sparse Recapture Scenarios. *Journal of Asian Development Studies*, *14*(2), 1518-1527.

Sendek, A. D., Yang, Q., Cubuk, E. D., Duerloo, K.-A. N., Cui, Y. and Reed, E. J. (2017) Holistic computational structure screening of more than 12,000 candidates for solid lithium-ion conductor materials. Energy & Environmental Science, 10, 306-320.

Therrien, F., Abou Haibeh, J., Sharma, D., Hendley, R., Wairimu Mungai, L., Sun, S., Tchagang, A., Su, J., Huberman, S., Bengio, Y., Guo, H., Hernández-García, A. and Shin, H. (2026) OBELiX: a curated dataset of crystal structures and experimentally measured ionic conductivities for lithium solid-state electrolytes. Digital Discovery, 5, 910-918. https://doi.org/10.1039/D5DD00441A.

Wang, J., He, T., Yang, X., Cai, Z., Wang, Y. et al. (2023) Design principles for NASICON super-ionic conductors. Nature Communications, 14, 5210.

Wang, Y., Richards, W. D., Ong, S. P., Miara, L. J., Kim, J. C., Mo, Y. and Ceder, G. (2015) Design principles for solid-state lithium superionic conductors. Nature Materials, 14, 1026-1031.

Ward, L., Dunn, A., Faghaninia, A., Zimmermann, N. E. R., Bajaj, S., Wang, Q. et al. (2018) Matminer: an open source toolkit for materials data mining. Computational Materials

Online ISSN 3007-3197
Print ISSN 3007-3189

Science, 152, 60-69.

Xie, S. R., Honrao, S. J. and Lawson, J. W. (2024) High-throughput screening of Li solid-state electrolytes with bond valence methods and machine learning. Chemistry of Materials, 36, 9320-9329.

Zhang, Y., He, X., Chen, Z., Bai, Q., Nolan, A. M., Roberts, C. A., Banerjee, D., Matsunaga, T., Mo, Y. and Ling, C. (2019) Unsupervised discovery of solid-state lithium ion conductors. Nature Communications, 10, 5260.

Zhu, Y., He, X. and Mo, Y. (2016) First principles study on electrochemical and chemical stability of solid electrolyte-electrode interfaces in all-solid-state Li-ion batteries. Journal of Materials Chemistry A, 4, 3253-3266.