# Supplementary Material: Two-Stage Transfer Learning for Deep Learning-based Prediction of Lattice Thermal Conductivity

L. Klochko and M. d'Aquin

*Université de Lorraine, LORIA, Nancy F-54000, France* [a)]

[a)]Electronic mail: liudmyla.klochko@loria.fr

## I. DATA

As mentioned in the main paper, the accuracy and reliability of any machine learning model is highly dependent on the quality of the data used for training and validation. In addition to the two datasets (Togo15 and AFLOW) discussed in the main paper, we include here results obtained with another dataset (Miyazaki2021) that we used to test was derived from calculations of the first principle of anharmonic lattice dynamics[1,2], looking at a small set of materials with specific structures in the same way as Togo15. A fourth, MIX dataset, is a combination of the first two, used to obtain a slightly larger and slightly more diverse dataset (see Fig. S14).

**Miyazaki2021:** This dataset includes thermal conductivity data for 143 half-Heusler compounds, as reported in[3]. This dataset adds another set of materials, while being itself significantly more specific than the previous one: it focuses only on one specific structure, for which the range of LTC values is significantly narrower ([2.17, 34.51] W/mK, compared to [0.51,1769.00] W/mK in Togo15).

**MIX dataset:** This dataset is a combination of the two datasets mentioned above, integrating the properties of Togo15 and Miyazaki2021 to enhance the diversity and scope of our model training.

The data processing steps applied to the two first datasets as described in the main article were also applied to the two additional ones.

## II. MODEL PARAMETERS AND PERFORMANCE METRICS FOR EACH STEP

We measured the performance of the model with different torch metrics.

Mean Absolute Percentage Error (MAPE) $= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right|$, where $y_i$ is the predicted value, and $\hat{y}$ is the target value.

R2 score: $(\mathrm{R}^2) = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$, where $y_i$ is the predicted value, and $\hat{y}$ is the target value.

Mean Absolute Error (MAE) $= \frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|$, where $y_i$ is the predicted value, and $\hat{y}$ is the target value.

| **Parameter** | **Value** |
|---|---|
| Activation Function | ReLU() |
| Learning Rate | $1 \times 10^{-3}$ |
| Optimizer | Adam |
| Loss Function | MAPE |
| Number of Hidden Layers | 3 |
| Neurons per Layer | $450 \rightarrow 350 \rightarrow 350$ |
| Train/Test Split | 80/20 |
| Batch Size | 8 |

TABLE S1: ParAIsite model Configuration Parameters

Root Mean Squared Error (RMSE) = $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$, where $y_i$ is the predicted value, and $\hat{y}$ is the target value.

TABLE S2: MAPE Results

| **Model** | **Tested on dataset** | | | |
|---|---|---|---|---|
| | **Togo15** | **Miyazaki2021** | **MIX** | **AFLOW** |
| **Step 1** | | | | |
| RWTG15 | 0.55 (0.20) | 2.24 (1.15) | 1.57 (0.64) | 2.28 (1.12) |
| RWMZ21 | 0.50 (0.08) | 0.38 (0.05) | 0.43 (0.05) | 0.48 (0.04) |
| RWMIX | 0.70 (0.15) | 0.75 (0.14) | 0.73 (0.11) | 1.10 (0.56) |
| RWAF | 0.58 (0.33) | 1.13 (0.28) | 0.92 (0.27) | 0.65 (0.33) |
| **Step 2** | | | | |
| FETG15 | 0.53 (0.21) | 3.20 (2.55) | 2.14 (1.50) | 3.27 (3.96) |
| FEMZ21 | 0.50 (0.08) | 0.37 (0.08) | 0.42 (0.05) | 0.49 (0.09) |
| FEMIX | 0.69 (0.15) | 0.73 (0.13) | 0.71 (0.12) | 0.97 (0.20) |
| FEAF | 0.55 (0.38) | 1.18 (0.27) | 0.93 (0.26) | 0.61 (0.34) |
| **Step 3** | | | | |
| FEAFTG15 | 0.28 (0.10) | 1.27 (0.39) | 0.88 (0.24) | 0.66 (0.18) |
| FEAFMZ21 | 0.56 (0.13) | 0.64 (0.21) | 0.61 (0.15) | 0.65 (0.07) |
| FEAFMIX | 0.34 (0.13) | 0.69 (0.25) | 0.55 (0.17) | 0.65 (0.08) |
| **Additional Step** | | | | |
| RWAFTG15 | 0.43 (0.39) | 1.32 (0.44) | 0.96 (0.33) | 0.83 (0.37) |

TABLE S3: R2 Score Results

| **Model** | **Tested on dataset** | | | |
|---|---|---|---|---|
| | **Togo15** | **Miyazaki2021** | **MIX** | **AFLOW** |
| **Step 1** | | | | |
| RWTG15 | 0.21 (0.61) | -0.58 (0.49) | -0.22 (0.53) | -2.26 (2.65) |
| RWMZ21 | -1.01 (0.43) | -0.51 (0.39) | -0.62 (0.25) | -0.80 (0.16) |
| RWMIX | -0.48 (0.61) | -0.48 (0.42) | -0.40 (0.44) | -0.89 (0.60) |
| RWAF | -0.08 (1.11) | -0.81 (0.68) | -0.42 (0.58) | -0.14 (0.83) |
| **Step 2** | | | | |
| FETG15 | 0.40 (0.47) | -0.51 (0.38) | -0.14 (0.33) | -1.89 (1.43) |
| FEMZ21 | -0.77 (0.49) | -0.34 (0.35) | -0.49 (0.28) | -0.78 (0.37) |
| FEMIX | -0.52 (0.44) | -0.51 (0.30) | -0.45 (0.24) | -0.58 (0.17) |
| FEAF | 0.03 (0.99) | -0.41 (0.73) | -0.17 (0.71) | -0.03 (0.83) |
| **Step 3** | | | | |
| FEAFTG15 | 0.85 (0.16) | -0.29 (0.38) | 0.19 (0.26) | 0.06 (0.47) |
| FEAFMZ21 | -0.03 (0.36) | -0.24 (0.53) | -0.17 (0.43) | -0.28 (0.27) |
| FEAFMIX | 0.62 (0.50) | -0.16 (0.49) | 0.17 (0.43) | -0.10 (0.25) |
| **Additional Step** | | | | |
| RWAFTG15 | 0.47 (1.19) | -0.64 (0.55) | -0.11 (0.60) | -0.28 (0.87 ) |

TABLE S4: MAE Results

| **Model** | **Tested on dataset** | | | |
|---|---|---|---|---|
| | **Togo15** | **Miyazaki2021** | **MIX** | **AFLOW** |
| **Step 1** | | | | |
| RWTG15 | 75.37 (33.64) | 101.67 (37.00) | 91.26 (29.56) | 287.70 (43.83) |
| RWMZ15 | 5.95 (1.38) | 4.37 (0.84) | 5.00 (0.82) | 5.50 (0.21) |
| RWMIX | 29.37 (11.13) | 25.75 (8.31) | 27.18 (8.46) | 44.50 (3.79) |
| RWAF | 4.25 (2.67) | 5.25 (1.42) | 4.86 (1.79) | 4.95 (1.72) |
| **Step 2** | | | | |
| FETG15 | 79.76 (40.04) | 102.68 (37.15) | 93.61 (23.15) | 286.44 (47.86) |
| FEMZ21 | 5.50 (1.01) | 3.92 (1.29) | 4.55 (0.99) | 5.34 (0.52) |
| FEMIX | 31.86 (9.51) | 26.10 (7.10) | 28.38 (6.46) | 44.40 (4.05) |
| FEAF | 3.70 (2.37) | 4.92 (1.60) | 4.44 (1.79) | 4.58 (1.48) |
| **Step 3** | | | | |
| FEAFTG15 | 1.24 (0.47) | 4.82 (1.07) | 3.41 (0.74) | 4.35 (0.48) |
| FEAFMZ21 | 4.57 (1.31) | 4.42 (1.58) | 4.48 (1.22) | 5.54 (0.56) |
| FEAFMIX | 2.44 (1.95) | 4.19 (1.53) | 3.50 (1.57) | 4.95 (0.67) |
| **Additional Step** | | | | |
| RWAFTG15 | 2.40 (3.19) | 5.25 (1.53) | 4.12 (2.07) | 4.96 (1.65) |

TABLE S5: RMSE Results

| **Model** | **Tested on dataset** | | | |
|---|---|---|---|---|
| | **Togo15** | **Miyazaki2021** | **MIX** | **AFLOW** |
| **Step 1** | | | | |
| RWTG15 | 152.49 (66.44) | 191.36 (92.18) | 176.46 (61.72) | 702.37 (116.27) |
| RWMZ21 | 8.22 (1.64) | 6.46 (1.22) | 7.12 (1.06) | 8.16 (0.33) |
| RWMIX | 50.77 (19.10) | 44.43 (16.74) | 46.43 (14.82) | 93.11 (8.65) |
| RWAF | 6.06 (3.34) | 7.82 (2.10) | 7.15 (2.35) | 8.73 (2.17) |
| **Step 2** | | | | |
| FETG15 | 162.46 (87.98) | 189.05 (92.71) | 182.12 (50.58) | 699.51 (121.79) |
| FEMZ21 | 7.55 (1.42) | 5.86 (1.88) | 6.56 (1.36) | 7.92 (0.62) |
| FEMIX | 55.89 (16.18) | 45.30 (15.28) | 49.12 (11.24) | 93.56 (8.88) |
| FEAF | 5.66 (3.75) | 7.17 (2.18) | 6.59 (2.43) | 8.15 (1.75) |
| **Step 3** | | | | |
| FEAFTG15 | 1.77 (0.74) | 7.01 (1.75) | 5.01 (1.20) | 7.86 (0.64) |
| FEAFMZ21 | 6.89 (2.15) | 7.06 (2.45) | 6.97 (1.65) | 9.61 (0.97) |
| FEAFMIX | 3.65 (3.13) | 6.70 (2.47) | 5.51 (2.41) | 8.72 (1.14) |
| **Additional Step** | | | | |
| RWAFTG15 | 0.47 (1.20) | 7.70 (2.26) | 6.06 (2.74) | 8.60 (2.05) |

## III. VALIDATION LOSS PER RUN AND STEP

Similar conclusions as in the main paper can be drawn on the MIX dataset, even if those are less strong: The models trained on their training subsets and tested on their validation subsets reach 73%, 71% and 55% in steps 1, 2 and 3 respectively. As seen in Figure S1, MAPE also reaches lower values in Step 2 compared to Step 2 and significantly lower ones yet in Step 3. However, in this case, overfitting appears early in the training process (around epoch 50) in all steps. This is probably explained by the fact that the MIX dataset combines Togo15, which is well predicted through the transfer learning process, and Miyazaki2021 which, as we discuss below, did not work as well.

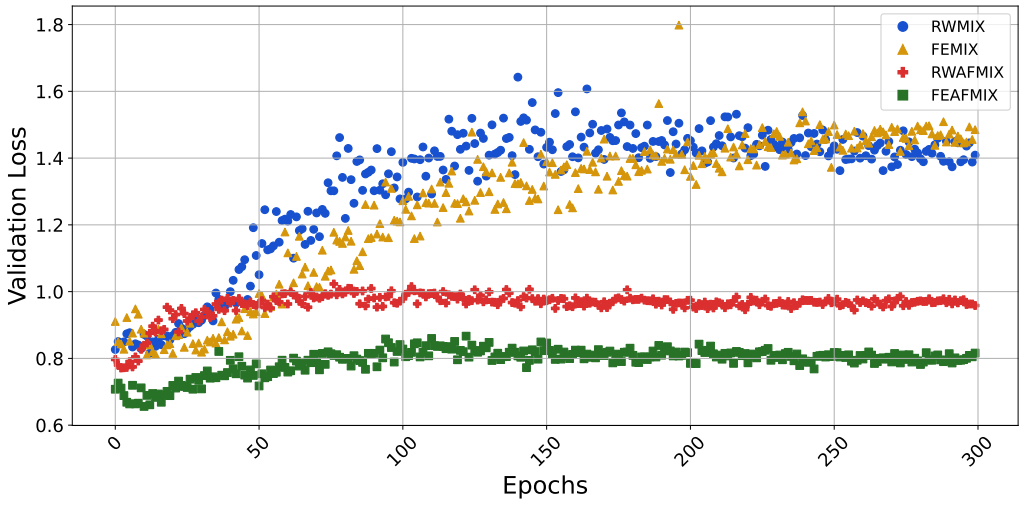


FIG. S1: Average validation loss for models trained and tested on MIX across training epochs.

Regarding Miyazaki2021, the results shown in Table S2 are significantly different from the ones obtained for the other two datasets. In fact, when training and testing on subsets of Miyazaki2021, the best average MAPE obtained where 38%, 37% and 64% in steps 1, 2 and 3 respectively. In other words, the third step, the double pre-training, significantly worsens the results obtained. This can be explained by the fact that this dataset contains a very restricted range of LTCs compared to the others, in particular AFLOW. In other words, having learned in Step 2 to predict a wide range of LTC values through the AFLOW dataset, the model did not adapt well to the very specific set of materials in Miyazaki2021. We can observe this issue in Figure S2 where, in Step 3, the training starts with a high error

rate and is overfit almost immediately.

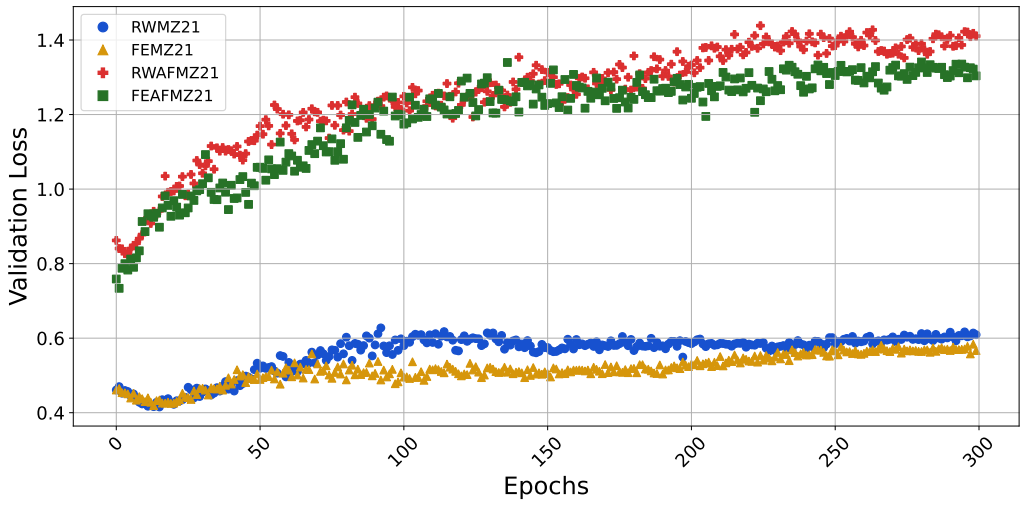


FIG. S2: Average validation loss for models trained and tested on Miyazaki2021 across training epochs.

The issue mentioned above, of the lack of variety in Miyazaki2021, is part of the motivation for integrating the MIX dataset as well. One final conclusion that can be drawn from the results shown in Tabs. S2– S5 is that, in accordance with our assumption, a more varied dataset (MIX) tends to generalize better, but also that the double pre-training process applied here helps support this ability to generalize. This is visible in the last line of the table, which shows that models trained on the MIX dataset are better on average when tested on all other datasets (including AFLOW) than those trained on the other two datasets in the same step. In most cases, models trained on MIX or Togo15 achieved better results when tested with other datasets at Step 3 than at other steps.

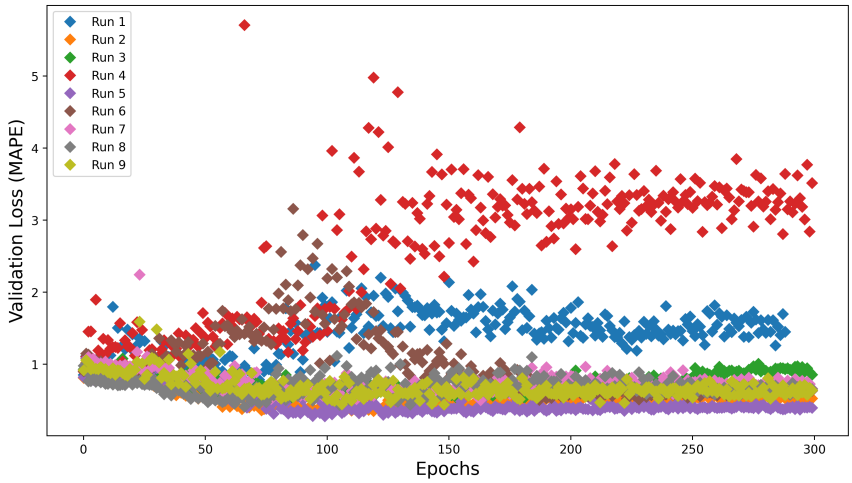


FIG. S3: Step 1: Validation loss (MAPE) for model trained and tested on Togo15 across all runs and training epochs.

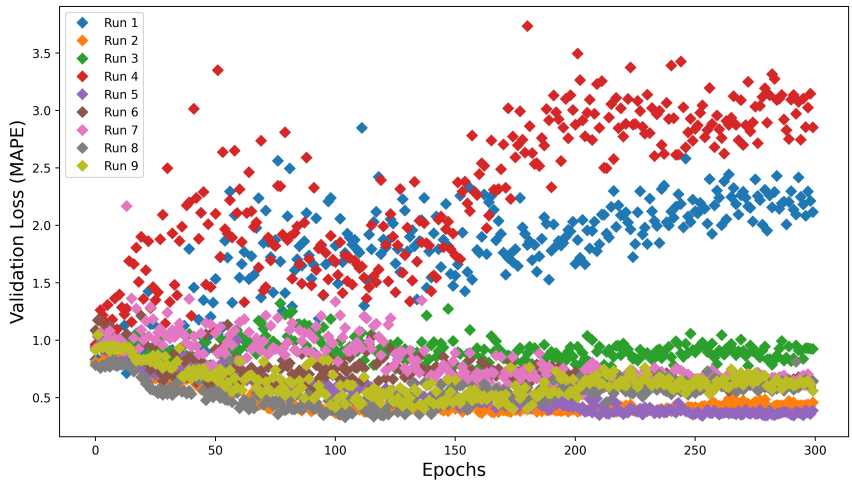


FIG. S4: Step 2: Validation loss (MAPE) for model trained and tested on Togo15 across all runs and training epochs.

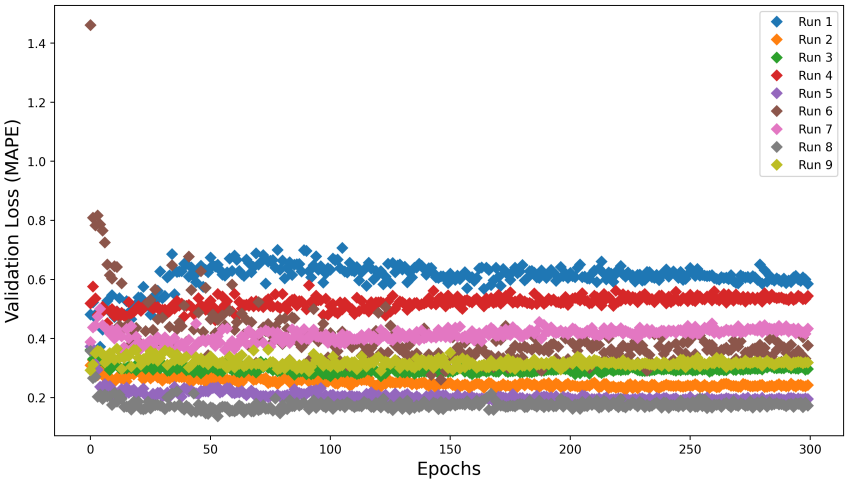


FIG. S5: Step 3: Validation loss (MAPE) for model trained and tested on Togo15 across all runs and training epochs.

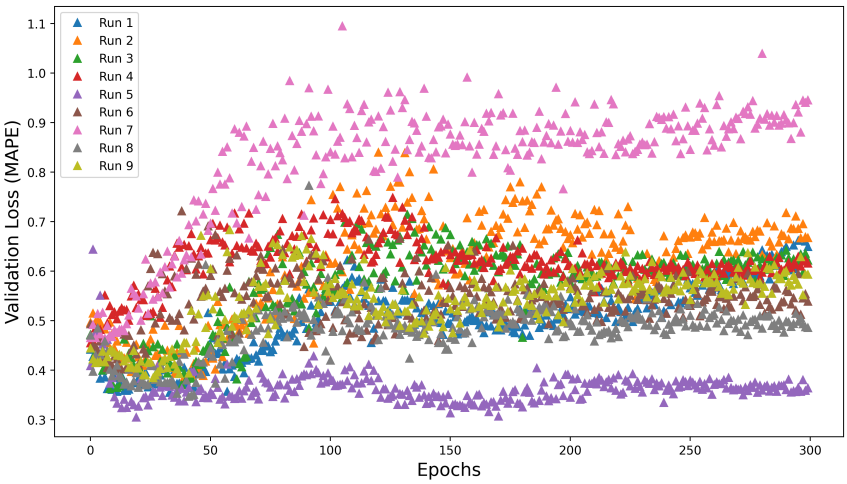


FIG. S6: Step 1: Validation loss (MAPE) for model trained and tested on Miyazaki2021 across all runs and training epochs.

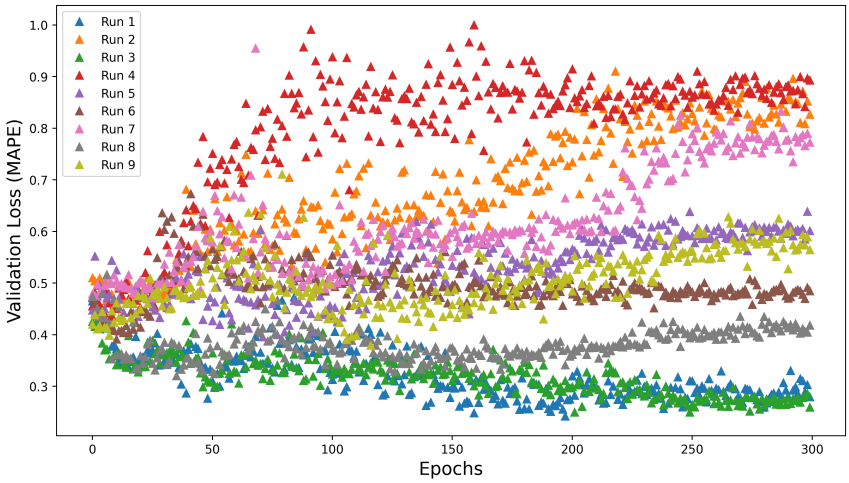


FIG. S7: Step 2: Validation loss (MAPE) for model trained and tested on Miyazaki2021 across all runs and training epochs.

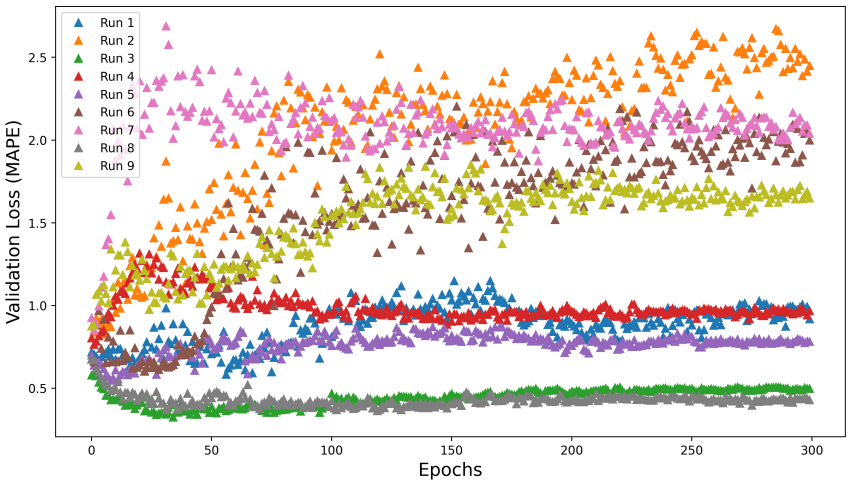


FIG. S8: Step 3: Validation loss (MAPE) for model trained and tested on Miyazaki2021 across all runs and training epochs.

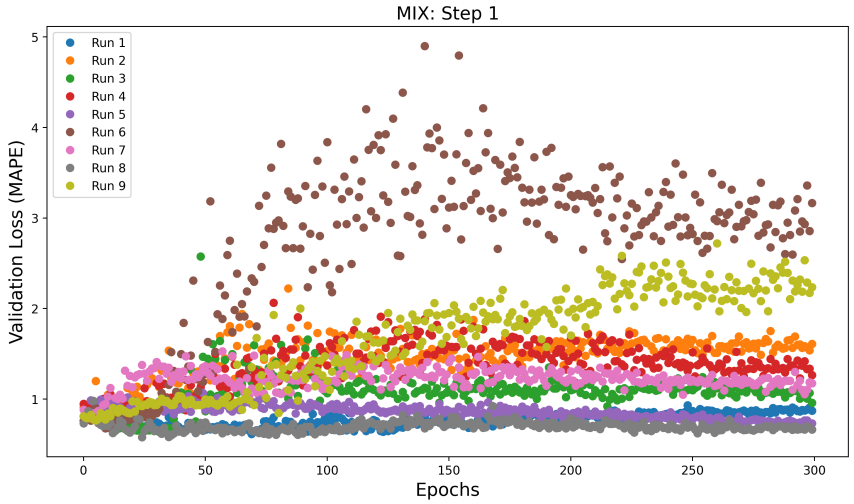


FIG. S9: Step 1: Validation loss (MAPE) for model trained and tested on MIX across all runs and training epochs.

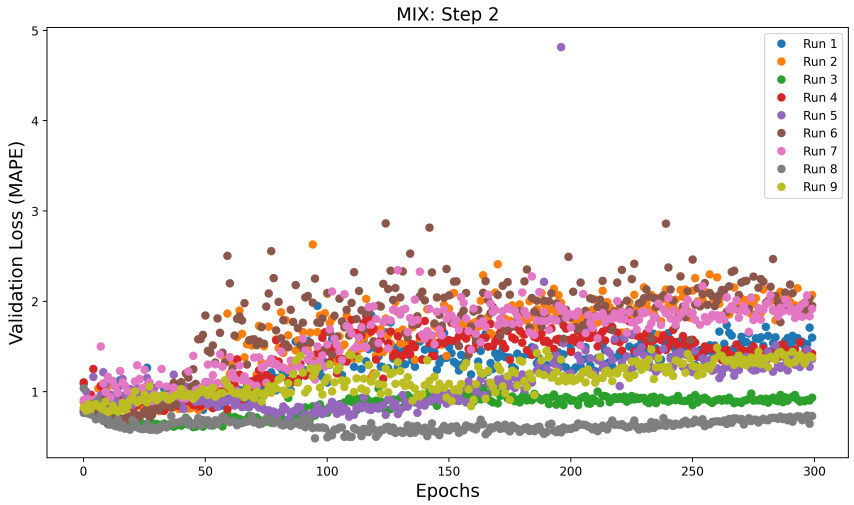


FIG. S10: Step 2: Validation loss (MAPE) for model trained and tested on MIX across all runs and training epochs.

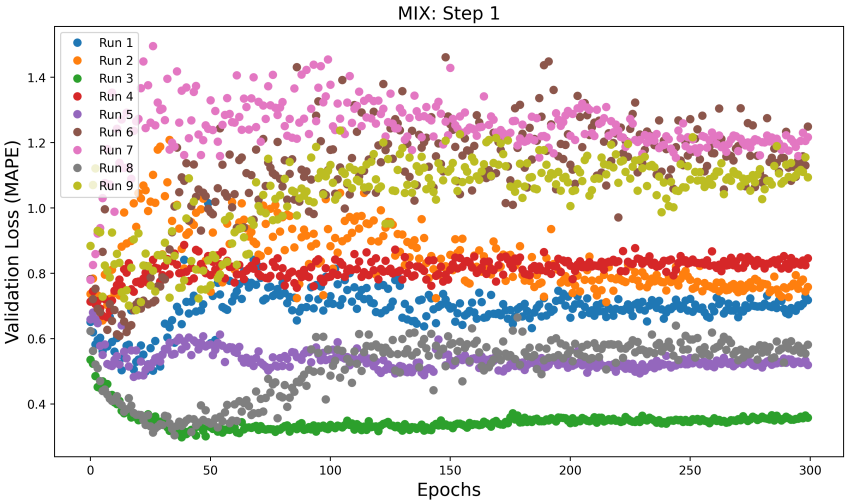


FIG. S11: Step 3: Validation loss (MAPE) for model trained and tested on MIX across all runs and training epochs.

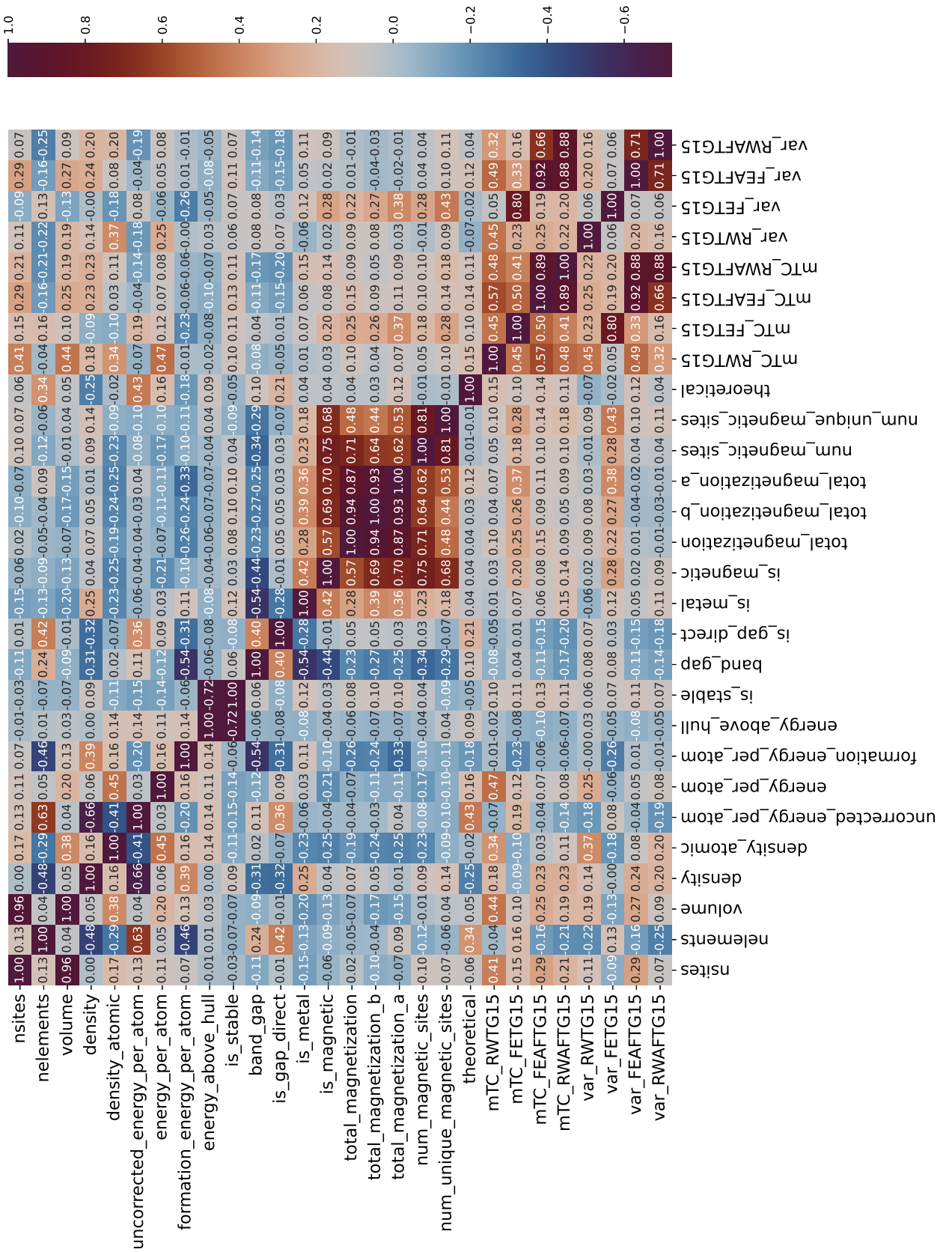


FIG. S12: Correlation matrix between all descriptors from the Materials Project for models trained on Togo15 dataset.

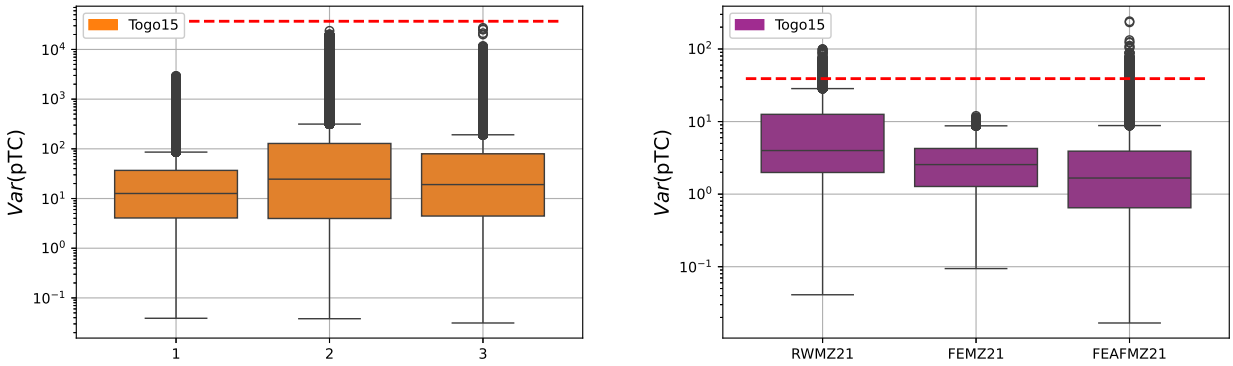


FIG. S13: Variance change at each step for predicted thermal conductivity ($pTC$) of models trained on Miyazaki2021 and MIX. The dashed line above box-plots represents the variance of LTC values in the original dataset.

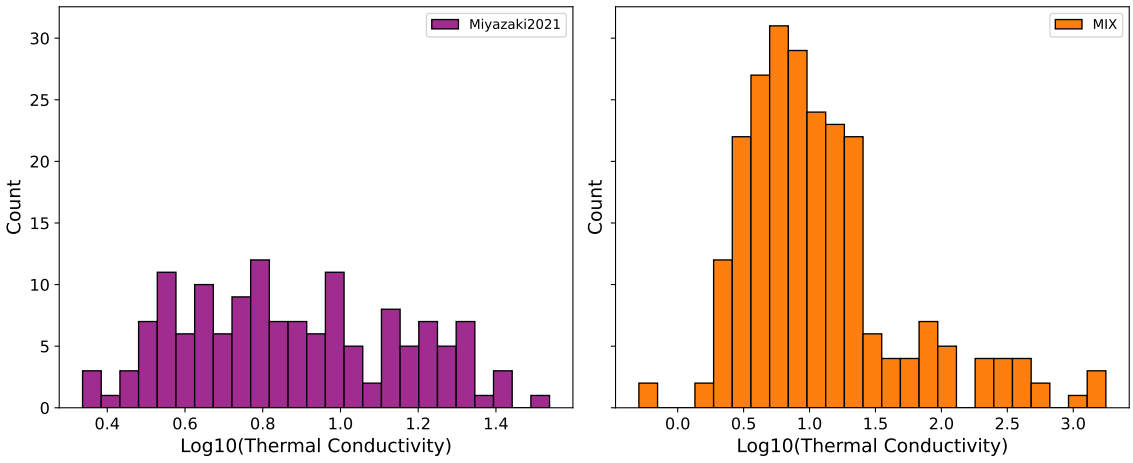


FIG. S14: The distribution of the thermal conductivity for each selected dataset.

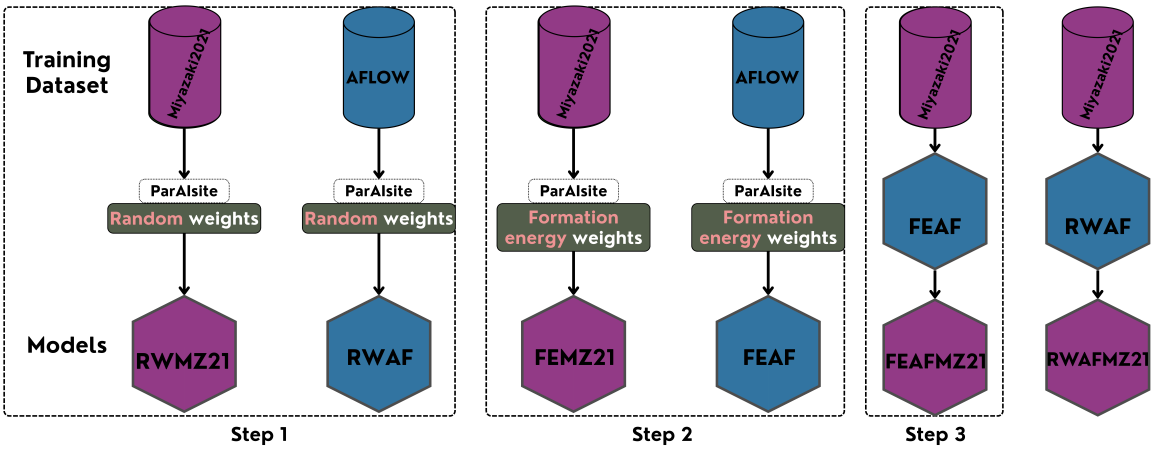


FIG. S15: Sketch representing our proposed approach. Training datasets are illustrated as cylinders, and the resulting models after training ParAIsite are represented as cubes. The models are labeled from left to right as follows: Step1 (no pre-training): Random Weights Miyazaki2021 (RWMZ21), Random Weights AFLOW (RWAF). Step 2 (using a pre-trained weights of MEGNet on the formation energy): Formation Energy Miyazaki2021 (FEMZ21), Formation Energy AFLOW (FEAF). Step 3 (transfer learning over model fine-tuned with AFLOW): Formation Energy AFLOW Miyazaki2021 (FEAFMZ21), Random Weights AFLOW Miyazaki2021 (RWAFMZ21).

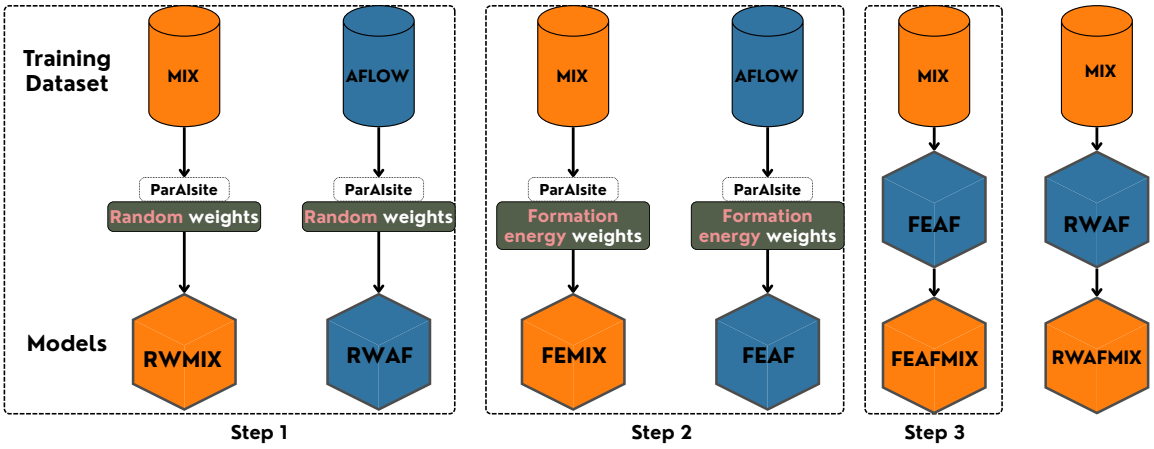


FIG. S16: Sketch representing our proposed approach. Training datasets are illustrated as cylinders, and the resulting models after training ParAIsite are represented as cubes. The models are labeled from left to right as follows: Step1 (no pre-training): Random Weights MIX (RWMIX), Random Weights AFLOW (RWAF). Step 2 (using a pre-trained weights of MEGNet on the formation energy): Formation Energy MIX (FEMIX), Formation Energy AFLOW (FEAF). Step 3 (transfer learning over model fine-tuned with AFLOW): Formation Energy AFLOW MIX (FEAFMIX), Random Weights AFLOW MIX (RWAFMIX).

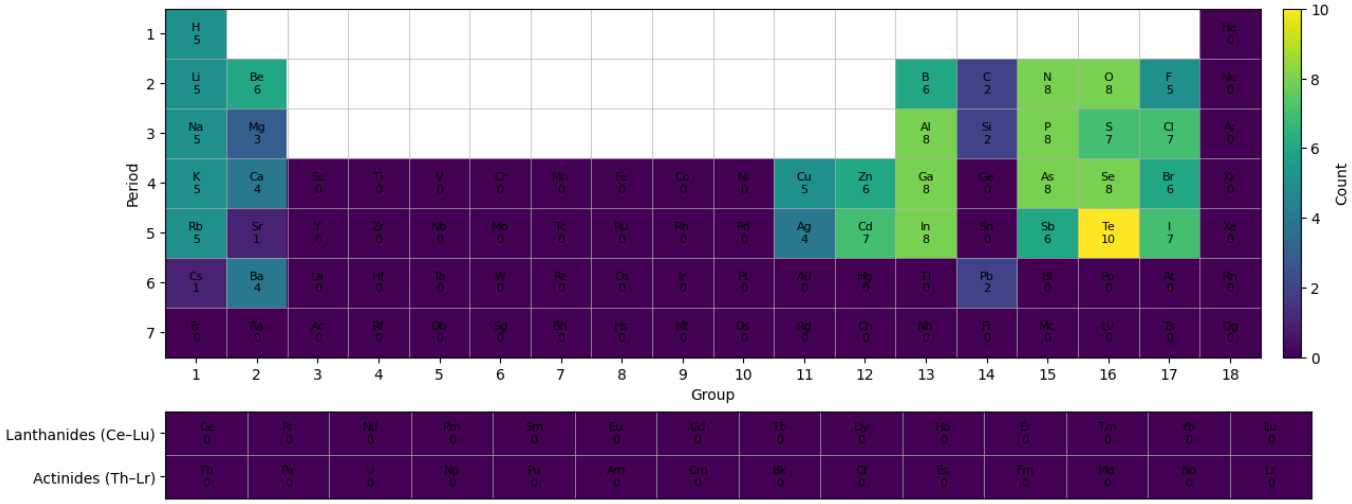


FIG. S17: The representation of element coverage of Togo15 dataset.

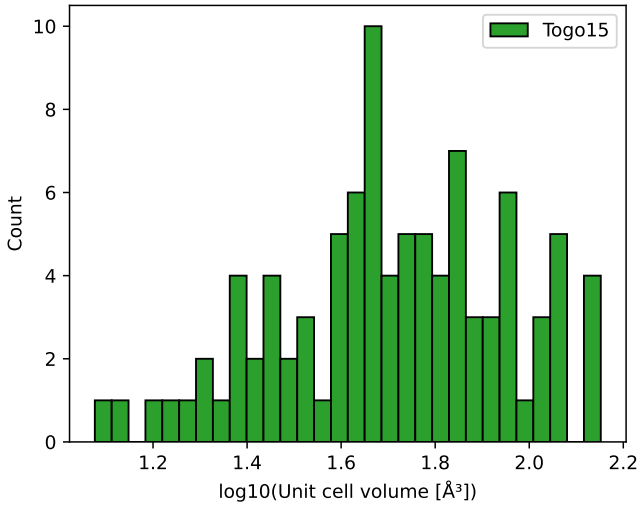


FIG. S18: The representation of unit cell volume distribution of Togo15 dataset.

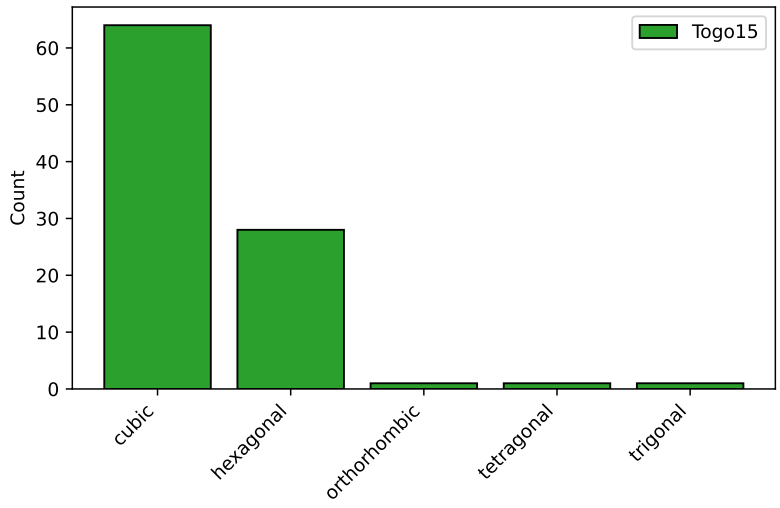


FIG. S19: The representation of crystal system distribution of Togo15 dataset.

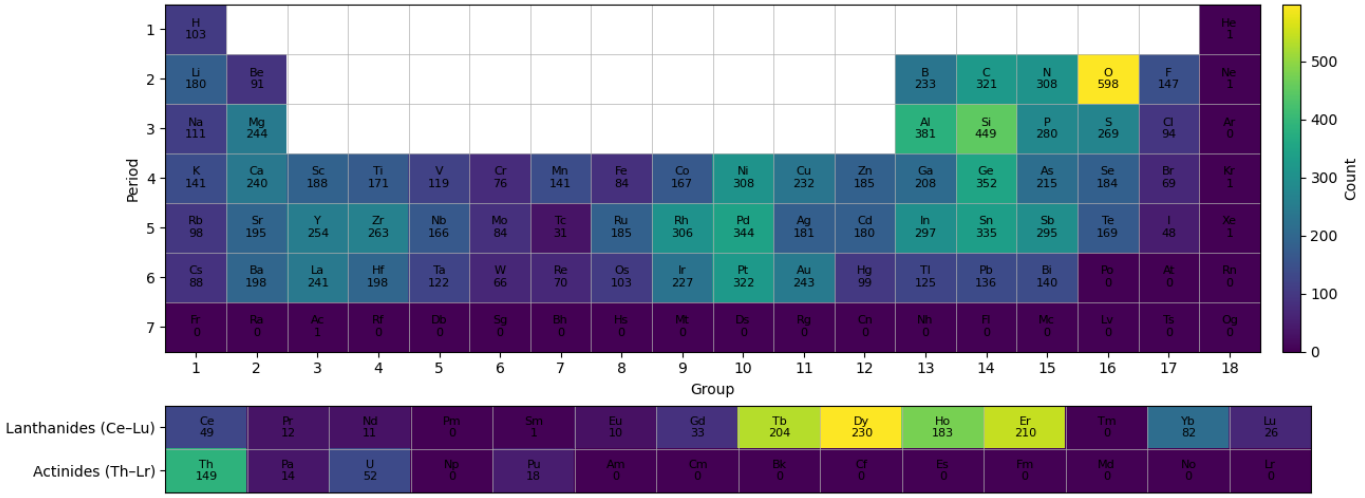


FIG. S20: The representation of element coverage of AFLOW dataset.

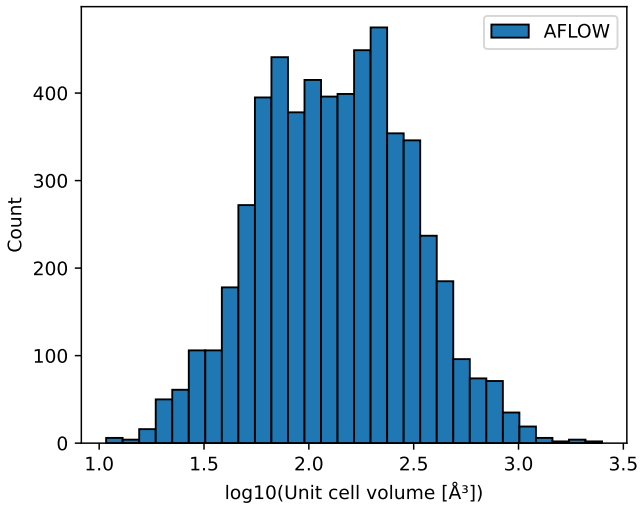


FIG. S21: The representation of unit cell volume distribution of AFLOW dataset.

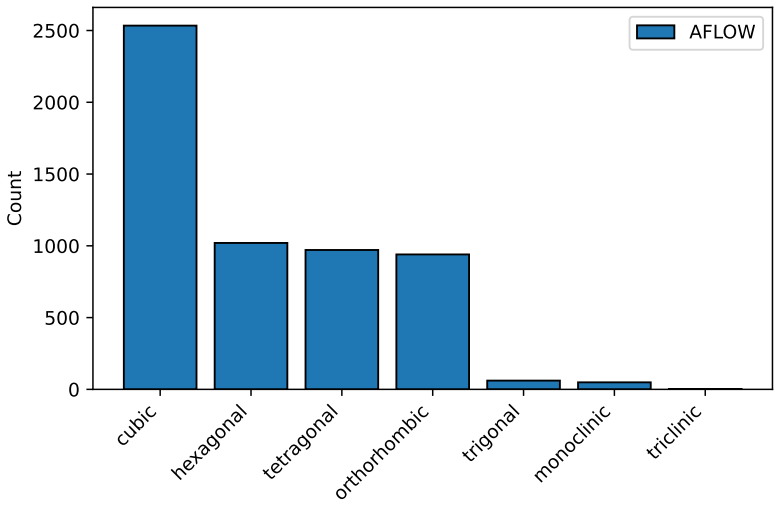


FIG. S22: The representation of crystal system distribution of AFLOW dataset.

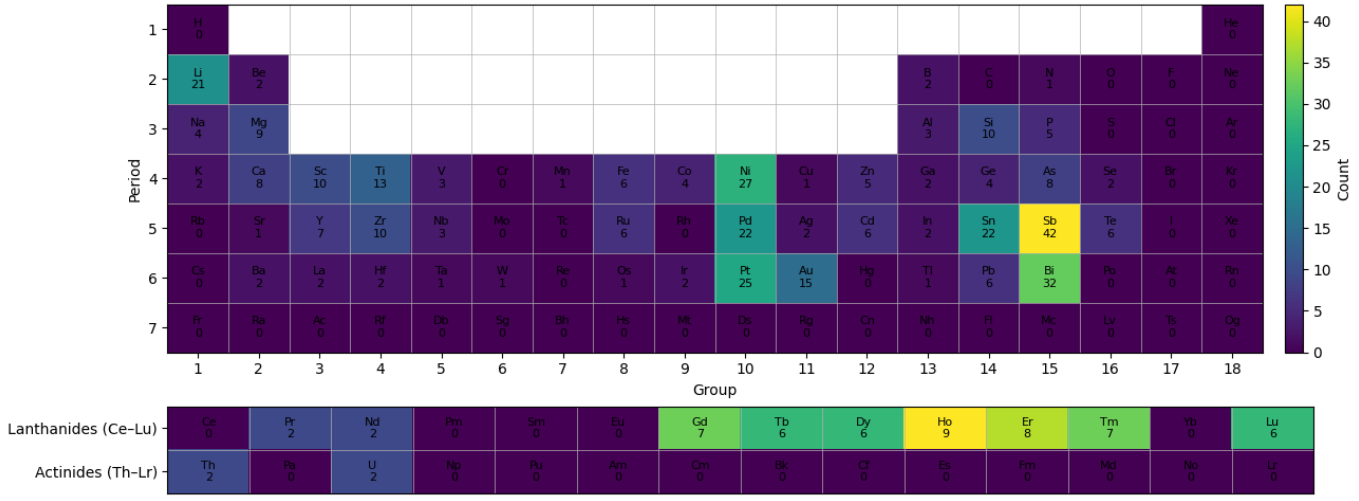


FIG. S23: The representation of element coverage of Miyazaki2021 dataset.

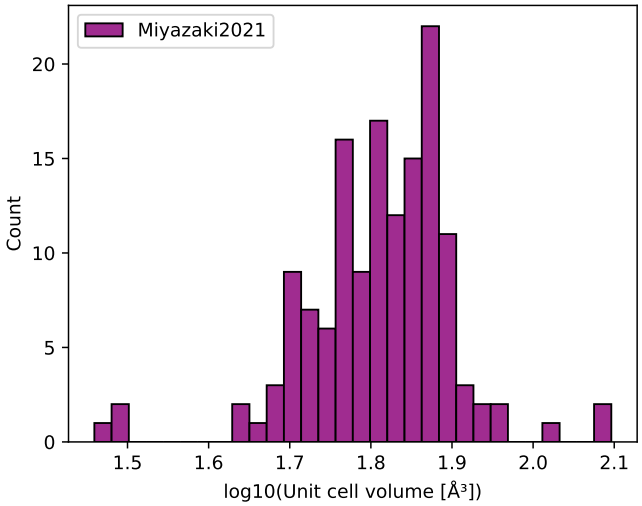


FIG. S24: The representation of unit cell volume distribution of Miyazaki2021 dataset.

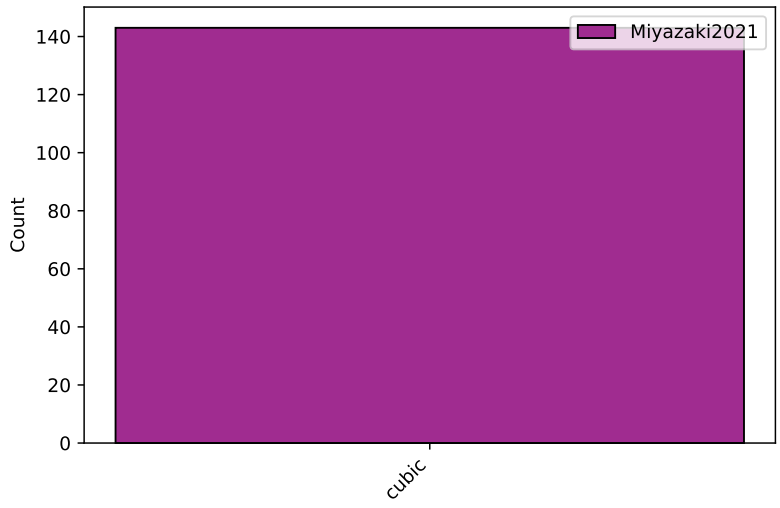


FIG. S25: The representation of crystal system distribution of Miyazaki2021 dataset.

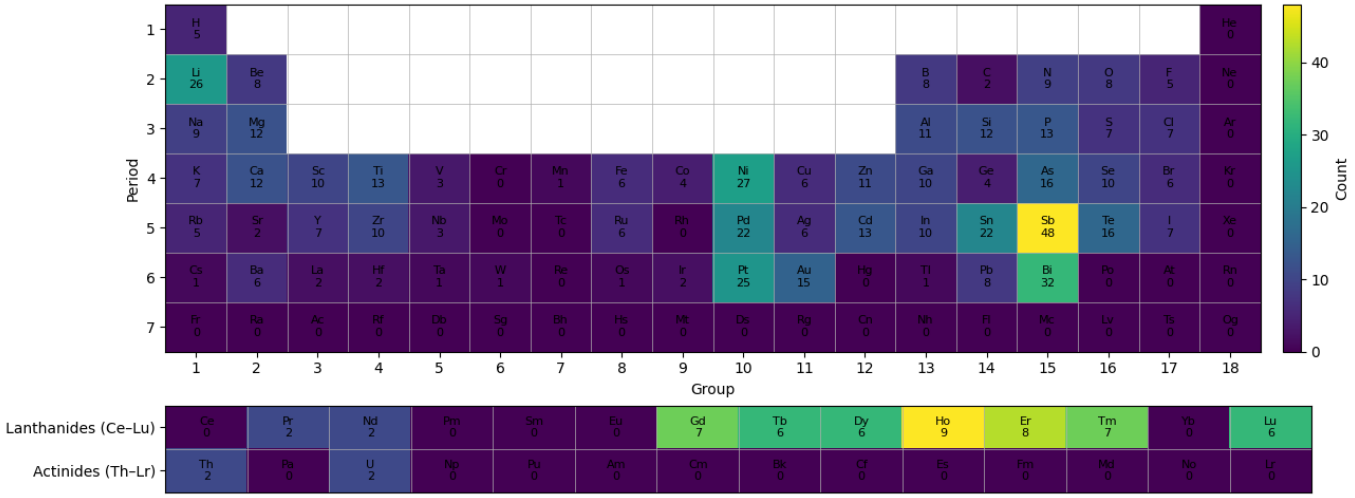


FIG. S26: The representation of element coverage of MIX dataset.

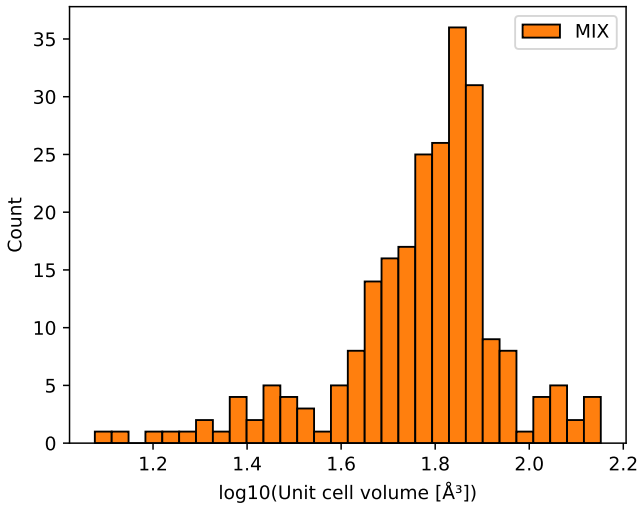


FIG. S27: The representation of unit cell volume distribution of MIX dataset.

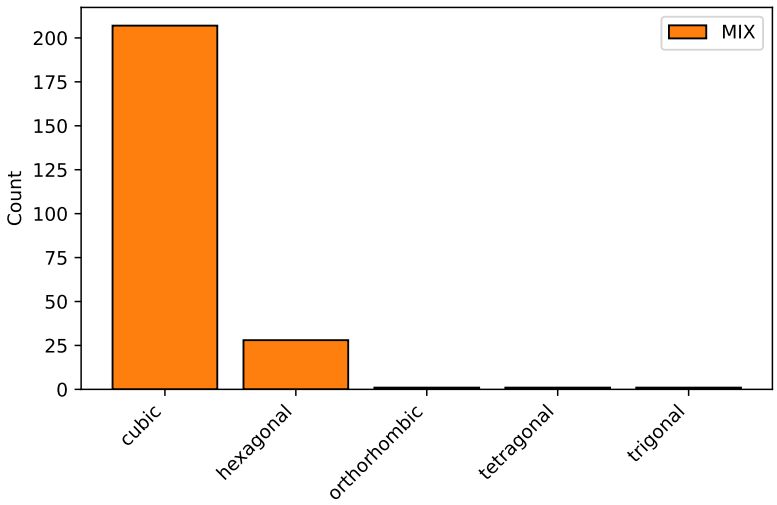


FIG. S28: The representation of crystal system distribution of MIX dataset.

TABLE S6: Results of top 30 materials with low LTC after ParAIsite screening over Material Project Database

| **MP-ID** | **Formula** | **LTC mean** | **std** | **LTC max** | **std** |
|---|---|---|---|---|---|
| mp-865958 | AcTlAg2 | 0.146 | 0.242 | 0.590 | 0.980 |
| mp-12056 | Ba(POs)2 | 0.464 | 0.391 | 1.345 | 1.134 |
| mp-1214877 | AlNi2BO5 | 0.640 | 0.497 | 1.500 | 1.163 |
| mp-568443 | Al2VCl8 | 1.003 | 0.466 | 1.829 | 0.850 |
| mp-2803876 | AgO | 1.200 | 0.429 | 1.927 | 0.689 |
| mp-9127 | Ba(SbO3)2 | 1.227 | 0.589 | 1.931 | 0.927 |
| mp-34418 | AgPF6 | 1.482 | 0.358 | 2.041 | 0.494 |
| mp-30273 | AcClO | 1.389 | 0.552 | 2.046 | 0.813 |
| mp-1194585 | AgP(OF)2 | 1.572 | 0.527 | 2.154 | 0.722 |
| mp-643387 | AlH2PbO2F3 | 1.261 | 0.651 | 2.164 | 1.117 |
| mp-1228984 | AlCu(PSe3)2 | 1.308 | 0.606 | 2.174 | 1.007 |
| mp-1228823 | AsIrS | 1.795 | 0.385 | 2.252 | 0.483 |
| mp-862319 | Ac2CdSn | 0.943 | 0.704 | 2.300 | 1.715 |
| mp-30934 | As(S8I)3 | 0.828 | 0.697 | 2.329 | 1.960 |
| mp-1229012 | AgPd | 1.244 | 0.524 | 2.372 | 0.999 |
| mp-8236 | Ba(PPd)2 | 1.758 | 0.499 | 2.449 | 0.695 |
| mp-18702 | AgPO3 | 1.648 | 0.448 | 2.467 | 0.671 |
| mp-5065 | Al2SiO5 | 1.685 | 0.592 | 2.637 | 0.927 |
| mp-865398 | Ac2ZnIr | 1.927 | 0.454 | 2.671 | 0.630 |
| mp-552185 | AgNO3 | 1.872 | 0.719 | 2.683 | 1.030 |
| mp-1191113 | AsH3OF6 | 1.986 | 0.482 | 2.790 | 0.676 |
| mp-1183136 | Al2OsRh | 1.983 | 0.527 | 2.803 | 0.745 |
| mp-1006278 | AcEuAu2 | 1.852 | 0.617 | 2.804 | 0.934 |
| mp-1183115 | AcAlO3 | 1.563 | 0.749 | 2.819 | 1.351 |
| mp-28967 | Ba(PdS2)2 | 1.202 | 0.878 | 2.820 | 2.060 |
| mp-23218 | AsI3 | 2.479 | 0.221 | 2.842 | 0.253 |
| mp-1183086 | Ac2CdHg | 1.218 | 0.956 | 2.958 | 2.321 |
| mp-707734 | AsHPb4(ClO)4 | 2.007 | 0.737 | 2.977 | 1.094 |
| mp-2593 | AlNi3 | 2.211 | 0.508 | 3.045 | 0.700 |

TABLE S7: Comparison of evaluation metrics for predicting thermal conductivity between 15 machine learning models (as tested in[4]) and our model (evaluated on multiple runs with the Togo15 dataset. Mean and standard deviation are reported for this model)

| **Model** | **RMSE** | **$R^2$** | **MAE** |
|---|---|---|---|
| Linear[4] | 26.4930 | 0.8096 | 13.5803 |
| Ridge[4] | 26.3697 | 0.8103 | 13.5384 |
| SGD[4] | 17.4929 | 0.8241 | 9.5713 |
| Linear SVR[4] | 11.9823 | 0.8479 | 6.8762 |
| SigmoidSVR[4] | 20.6119 | 0.3432 | 9.7734 |
| rbfSVR[4] | 14.6274 | 0.7547 | 6.9582 |
| PolySVR[4] | 12.0221 | 0.7496 | 7.2365 |
| Decision tree[4] | 19.3780 | 0.5348 | 8.6358 |
| DGBT[4] | 10.3623 | 0.8158 | 6.9570 |
| Random forests[4] | 9.6385 | 0.8767 | 6.0574 |
| LightGBM[4] | 12.9994 | 0.7398 | 7.7365 |
| ANN[4] | 8.7211 | 0.8593 | 5.7933 |
| CNN[4] | 8.4061 | 0.8799 | 5.1674 |
| RNN[4] | 8.3726 | 0.8748 | 5.3209 |
| LSTM[4] | 8.3593 | 0.8866 | 5.4011 |
| **FEAFTG15 (step 3)** | **1.77 (0.74)** | **0.85 (0.16)** | **1.24 (0.47)** |

## REFERENCES


[1]A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and I. Tanaka, "Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and bayesian optimization," Physical Review Letters **115**, 205901 (2015).

[2]A. Togo, L. Chaput, and I. Tanaka, "Distributions of phonon lifetimes in brillouin zones," Physical Review B **91**, 094306 (2015).

[3]H. Miyazaki, T. Tamura, M. Mikami, K. Watanabe, N. Ide, O. M. Ozkendir, and Y. Nishino, "Machine learning based prediction of lattice thermal conductivity for half-heusler compounds using atomic information," Scientific Reports **11**, 13410 (2021).

[4]G. Qin, Y. Wei, L. Yu, J. Xu, J. Ojih, A. D. Rodriguez, H. Wang, Z. Qin, and M. Hu, "Predicting lattice thermal conductivity from fundamental material properties using machine learning techniques," Journal of Materials Chemistry A **11**, 5801–5810 (2023).