Editor's Choice

Machine learning models for the lattice thermal conductivity prediction of inorganic materials

Lihua Chen, Huan Tran, Rohit Batra, Chiho Kim, Rampi Ramprasad*

School of Materials Science and Engineering, Georgia Institute of Technology, 771 Ferst Drive NW, Atlanta, GA 30332, United States

ABSTRACT

The lattice thermal conductivity (κ_L) is a critical property of thermoelectrics, thermal barrier coating materials and semiconductors. While accurate empirical measurements of κ_L are extremely challenging, it is usually approximated through computational approaches, such as semi-empirical models, Green-Kubo formalism coupled with molecular dynamics simulations, and first-principles based methods. However, these theoretical methods are not only limited in terms of their accuracy, but sometimes become computationally intractable owing to their cost. Thus, in this work, we build a machine learning (ML)-based model to accurately and instantly predict κ_L of inorganic materials, using a benchmark data set of experimentally measured κ_L of about 100 inorganic materials. We use advanced and universal feature engineering techniques along with the Gaussian process regression algorithm, and compare the performance of our ML model with past theoretical works. The trained ML model is not only helpful for rational design and screening of novel materials, but we also identify key features governing the thermal transport behavior in non-metals.

1. Introduction

The lattice thermal conductivity (κ_L) dictates the ability of a nonmetal to conduct heat, and serves as a critical design parameter for a wide range of applications, including thermoelectrics for power generation [1,2], thermal barrier coatings for integrated circuits [3,4], and semiconductors for microelectronic devices [5]. Depending on the specific application, materials with different ranges of κ_L values are desired. For example, low κ_L is preferred as thermoelectrics (e.g., PbTe and Bi_2Te_3) to maximize the thermoelectric figure of merit, while for semiconductors (e.g., SiC and BP), high κ_L is required to avoid overheating in electronic devices. Motivated by their practical and technological significance, extensive theoretical and empirical efforts have been made to compute κ_L, aimed at discovering materials with targeted thermal conductivity for specific applications.

In one of the early and famous theoretical works, κ_L of inorganic materials was estimated using semi-empirical Slack model [6], which relies on the Debye temperature (Θ_D) and the Gruneisen parameter (Υ) as inputs, obtained from either experimental measurements or first-principles calculations [6,7]. Although the Slack model can provide a quick κ_L estimate, the uncertainty in its input parameters (Θ_D, Υ) severely impacts its prediction accuracy. Slight modifications in the functional form of the Slack model (or its closely related Debye-

Callaway model [8]) have also been attempted by treating certain power coefficients as fitting parameters, which are determined using experimentally measured κ_L values.

However, the underlying problem of Θ_D and Υ uncertainty and their unavailability for new materials persists. Alternatively, the Green-Kubo formalism, combined with non-equilibrium molecular dynamics simulations, has been employed to predict κ_L in semiconductors (e.g., Si) [9–11]. However, this method can only be used for materials for which reliable atomistic force fields are available. With the recent developments of computing power and first-principles implementations, the *ab initio* Green-Kubo approach has been proposed to compute the κ_L of Si and ZrO_2, but it is limited by the high computational cost to achieve the heat flux and system size convergences [12]. Additionally, the phonon Boltzmann transport equation (BTE) can now be solved numerically within the relaxation time approximation [13–16]. In this approach, κ_L is computed from the group velocity, the mode-dependent heat capacity, and the single-mode relaxation time (approximated by the phonon lifetime), all of which rely on either the harmonic or the anharmonic force constants computed at the first-principles level. While BTE calculations could in principle be done for large systems [16], they are generally restricted to small unit cells owing to high computational costs.

Machine learning (ML) based methods, which are emerging in

* Corresponding author.
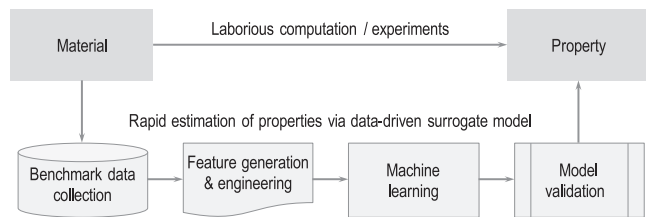 *E-mail address:* rampi.ramprasad@mse.gatech.edu (R. Ramprasad).

Fig. 1. Schematic of the workflow adopted to build data-driven models of κ_L.
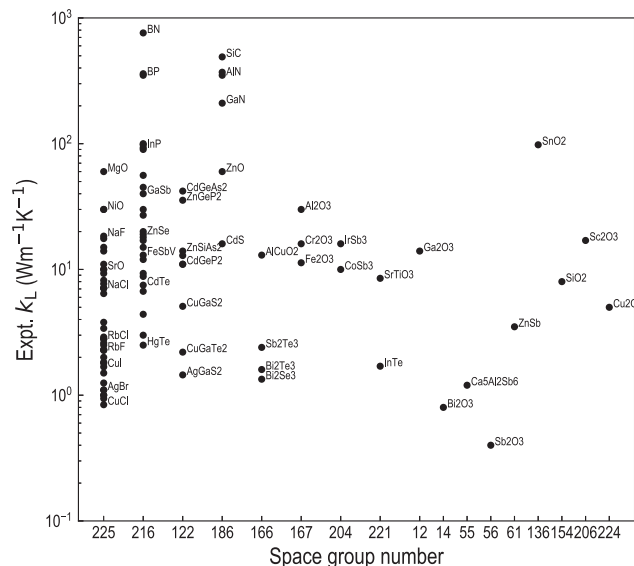

Fig. 2. Experimentally measured κ_L for 100 inorganic compounds with respect to their space group number. For space group 225 and 216, only a few representative cases are labeled.

Materials Science and Engineering [17–22] provide yet another approach to build surrogate models to rapidly predict the thermal conductivity of materials. Seko et al. developed ML models based on κ_L computed for 110 materials (by solving the phonon Boltzmann transport equation as mentioned above) and a set of descriptors characterizing elemental and structural properties [13,23]. The main concern with such ML model is the discrepancy between the DFT computed training data and the actual experimental values (especially for solids with very high κ_L) which directly impacts the accuracy of these models. Furthermore, the identification of key features in determining the κ_L is far from trivial.

To fill the above-mentioned gaps, we have built an ML model for κ_L, starting from a benchmark empirical data set of 100 inorganic compounds. The scheme adopted in this work is illustrated in Fig. 1. First, the recently released Matminer package [24] was used to generate a comprehensive list of 63 features to numerically represent the materials. This step was followed by the recursive feature elimination algorithm, down selecting the relevant features. The Gaussian process regression (GPR) algorithm, with 5-fold cross-validation (CV), was then utilized to build predictive models. The performance of the κ_L models was compared with past studies and validated by 5 unseen materials. The developed ML model, which is trained on the κ_L dataset spanning across 3 orders of magnitude, can be used to instantly predict κ_L of new inorganic materials while the associated GPR uncertainty could indicate whether the new materials are within the training domain or not. It is hoped that the model developed in this work can be used to screen new inorganic materials with targeted κ_L, and it can be systematically improved when new materials are identified and added to the initial dataset.

2. Technical details

2.1. Data set

Fig. 2 and Table 1 summarize the dataset of empirically measured κ_L values (at room temperature) for 100 single crystal inorganic materials collected from the literature [7,6,18,25–54], including 81 binary and 19 ternary compounds. κ_L of single-element materials are excluded since thermal conductivity of individual elements within a compound were used as features. The dataset is significantly diverse in chemical compositions (35 cations and 22 anions), crystal structures (with space group 225, 216, 122, 186, etc.), and the range of κ_L, which spans over 3 orders of magnitude (0.4–760 $Wm^{-1}K^{-1}$). The entire κ_L data set—along with the bulk modulus feature values—is provided in Table S1 of the Supporting Information (SI).

Given the wide range of κ_L, our learning problem was framed in the logarithmic scale, i.e., $\log(\kappa_L)$ was set as the target property, to allow better generalization of the ML models across the entire range. Furthermore, 95 out of 100 cases were used to train (with CV) the ML models, while the remaining 5 data points were held-out separately (completely unseen to the entire training process) to further validate the performance of the learned κ_L model. For cases where multiple κ_L values were reported in the literature, their average was used to train the ML model.

Table 1
Different properties of the κ_L data set utilized in this work, including the class of materials, their chemical composition and space group, and the range of κ_L values.

Classification Area	Category	Examples	Count
Compounds	Binary	$AgBr$, SiO_2, Al_2O_3, …	81
	Ternary	$AgGaS_2$, $HfCoSb$, …	19
Chemical composition	Cations	Na, K, Li, Be, Mg, Al, …	35
	Anions	F, Cl, Br, I, O, S, Se, …	22
Space group	225	$CuCl$, $SnTe$, $NaCl$, …	39
	216	$InSb$, $AlAs$, SiC, …	26
	122, 166, etc.	$CdGeP_2$, Bi_2Se_3, …	35
Expt. κ_L	0.4–10	Sb_2O_3, $AgCl$, Mg_2Sn, …	53
$(Wm^{-1}K^{-1})$	10–100	CoO, ZnS, $CdGeAs_2$, …	40
at room temperature	100 – 760	GaN, BN, BeO, …	7

2.2. Feature set and dimensionality reduction

To build accurate and reliable ML models, it is important to include relevant features that collectively capture the trends in the κ_L values across the different materials. The features should not only uniquely represent each material, but also be readily available to allow instant predictions for new cases. In this regard, Matminer is a good resource to easily and quickly generate features, applicable specifically to the field of materials science [24]. In total, 61 features, belonging to three distinct categories, i.e., elemental, structural and pertaining to valence electrons, were obtained using the Matminer package [24] by providing the chemical formula and the atomic configuration of all compounds. A total of 18 elemental properties were derived, including atomic radius, atomic mass, atom number, periodic table group and row, block, Mendeleev number, covalent radius, volume per atom from ground state, molar volume, coordination number (cn), Pauling electronegativity, first ionization energy, melting point, boiling point, thermal conductivity, average bond length and angle of a specific site with all its nearest neighbors. Since our dataset consists of binaries and ternaries, each of these elemental feature values was obtained by taking the minimum, maximum, and weighted average over the constituting chemical species, resulting in a total of 54 elemental features. For the structural features, volume per atom, packing fraction and density were considered. These quantities were computed for the crystal structure

obtained from the Materials Project database [55]. Moreover, 4 features that capture the average number of valence electrons in the s, p, d, and f shells of the constituting elements were also included. Finally, two additional features, DFT computed bulk modulus and the space group number, were also incorporated, resulting in a 63-dimensional feature vector. The values for bulk modulus of all compounds were obtained from the Material Project database [55]. As per standard ML practices, all features were scaled from 0 to 1 during model training.

To retain only the relevant features, recursive feature elimination (RFE) using linear support vector regression algorithm (with 5-fold CV) was performed on the initial 63-dimensional feature vector and the dataset of 95 training points. RFE eliminates the irrelevant features by recursively ranking feature importance and pruning the least important ones. In our case, it reduced the dimensionality from 63 to 29 (see Table S2 of the SI). We also used random forest algorithm for feature dimensionality reduction. In particular, we trained the data set of 95 points using 100 trees, and used the feature importance/weight to determine the relevance of the features. As discussed in Section 2 of the SI, nearly 40 features were identified to be important using the random forest method, most of which were found to be consistent to those retained from the RFE scheme discussed earlier. This provides more confidence to the RFE based dimensionality reduction step performed in this work. Overall, the 29-dimensional feature vector obtained after RFE resulted in more accurate models than the original 63-dimensional feature, as will be discussed in detail next, while a detailed comparison of the RFE and random forest methods is provided in SI.

2.3. Gaussian process regression

The Gaussian process regression (GPR) with the radial basis function (RBF) kernel was utilized to train the ML models. In this case, the co-variance function between two materials with features $\boldsymbol{x}$ and $\boldsymbol{x}'$ is given by

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f \exp\left(-\frac{1}{2\sigma_l^2}||\boldsymbol{x} - \boldsymbol{x}'||^2\right) + \sigma_n^2. \tag{1}$$

Here, three hyper parameters σ_f, σ_l and σ_n signify the variance, the length-scale parameter and the expected noise in the data, respectively. These hyper parameters were determined during the training of the models by maximizing the log-likelihood estimate. Further, 5-fold CV was adopted to avoid overfitting. Two error metrics, namely, the root mean square error (RMSE) and the coefficient of determination (R^2), were used to evaluate the performance of the ML models. To estimate the prediction errors on unseen data, learning curves were generated by varying the size of the training and the test sets. We note that the test sets were obtained by excluding the training points from the data set of 95 points. The left-out set of 5 points was completely separated from the learning process, and was used for just evaluation purposes on a few "extrapolative" material cases. Additionally, for each case, statistically meaningful results were obtained by averaging RMSE results over 100 runs with random training and test splits.

3. Results and discussion

It is worth analyzing the correlation between these 29 features and the empirically measured κ_L to see how much trend is captured by these elemental, structural and chemical attributes. While in Fig. 3 we plot the $\log(\kappa_L)$ vs four important features, the corresponding plots for the remaining cases are provided in Fig. S3 of SI. A strong positive correlation between $\log(\kappa_L)$ and bulk modulus, and a strong inverse relation between $\log(\kappa_L)$ and the mean average bond length are evident from the figure. While density alone does not show a strong correlation with $\log(\kappa_L)$, the combined feature $\sqrt{\text{bulkmodulus/density}}$ does indeed show a very strong linear relation. This is in-line with the physical understanding that group velocity, which is an integral part of the semi-

empirical models discussed earlier, is related to the lattice anharmonic force constants, and can be approximated as $\sqrt{\text{bulkmodulus/density}}$. Thus, bulk modulus can be considered to play a critical role in influencing the κ_L of different inorganic non-metals. Similarly, the inverse relationship between $\log(\kappa_L)$ and the mean average bond length is also physically meaningful as when the bonds are shorter, the bond-strength anharmonicity are stronger, and the resulting κ_L is larger. For the case of mean atomic mass (a common feature used in the past ML model works), a slightly dispersed relationship is observed, indicating that it may be less important in governing κ_L, as was the case with the rest of the 25 features illustrated in Fig. S3 of the SI.

Next, the performance of the ML models can be evaluated from the learning curves presented in Fig. 4(a), wherein average RMSE on the training and the test sets as a function of training set size are included. The error bars denote the 1σ deviation in the reported RMSE values over 100 runs. Results using both the initial set of 63 features (GPR), and those for the reduced 29 features (GPR-RFE) are included. Clearly, the RFE dimensionality reduction step leads to improved model performance with lower test errors, which signify better generalization of these models for unseen data. As expected, the test RMSE of both the GPR and the GPR-RFE models decreases with increase in training set size, reaching a convergence of 0.28 in test error and of 0.18 in train error for GPR-RFE models when the training set is about 80 % of the data (i.e. 76 points). Fig. 4(b) and (c) show the performance of GPR-RFE models via the example parity plots (i.e., ML predicted vs experimental $\log(\kappa_L)$), using 76 and 95 train points, respectively. The error bars in these cases represent the GPR uncertainty. Pretty high R^2 coefficient ($\geqslant 0.93$) on the test set in both these cases suggest a good κ_L model has indeed been developed.

We compared the performance of our ML model with other semi-empirical models by computing the average factor difference (AFD) [8], using the definition $\text{AFD} = 10^a$, where $a = \frac{1}{N}\sum_{i=1}^{N} |\log(\kappa_L)^{\text{expt.}} - \log(\kappa_L)^{\text{model}}|$, with N being the number of data points. As shown in Table 2, the computed AFD of GPR-RFE models using the entire set of 95 points is 1.36 ± 0.03, which is comparable to the reported values of 1.38 and 1.46, respectively, obtained using the Slack [56] and Debye-Callaway [8] models. More importantly, the latter two ML models rely on the experimental/computed features that are much more difficult to obtain owing to their dependence on the use of the Slack or Debye-Callaway models. The ML model presented here uses easily and rapidly accessible chemical and structural features derived directly from the identity of the material, making it more inexpensive and flexible. In addition, the predicted GPR uncertainty can be used to guide the next experiments via active learning [57]. Further, we note that the possibility of further diversifying our ML model with data from first-principles or semi-empirical methods using multi-fidelity fusion approaches also exists [58,59].

In order to further validate the generality and the accuracy of our ML models, we used the GPR-RFE models trained on the entire set of 95 points (see Fig. 4(a)) to predict the $\log(\kappa_L)$ of 5 unseen inorganic solids with various space group numbers present in the hold-out set. These include Sc_2O_3 (206), Ga_2O_3 (12), MnO (225), $AlCuO_2$(166), and $Ca_5Al_2Sb_6$ (55), where the number within brackets is the space group number. Fig. 4(c) shows the comparison between the predicted and the experimental $\log(\kappa_L)$, with error bars capturing the GPR uncertainty. A good performance for these 5 unseen data points is clearly evident. The high GPR uncertainty in the case of $Ca_5Al_2Sb_6$ correctly signals its space group number differences from that of the majority training data, and the application of the ML model in the "extrapolative" regime. Overall, the results presented here strongly advocate the good performance of the GPR-RFE models developed, which can be used to provide an inexpensive and accurate κ_L prediction for other inorganic materials, especially for materials with rock-salt or zincblende structures.

Additionally, it is noteworthy that the space group number is one of the important features in our ML model, although it has little physical meaning beyond allowing the model to distinguish between different
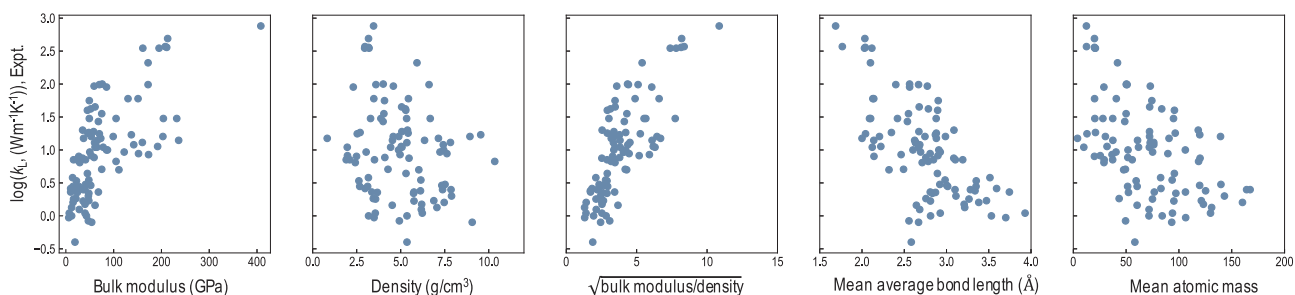
Fig. 3. The correlation between experimental κ_L and four representative features employed in this study. $\sqrt{\text{bulk modulus/density}}$ is derived from bulk modulus and density features.
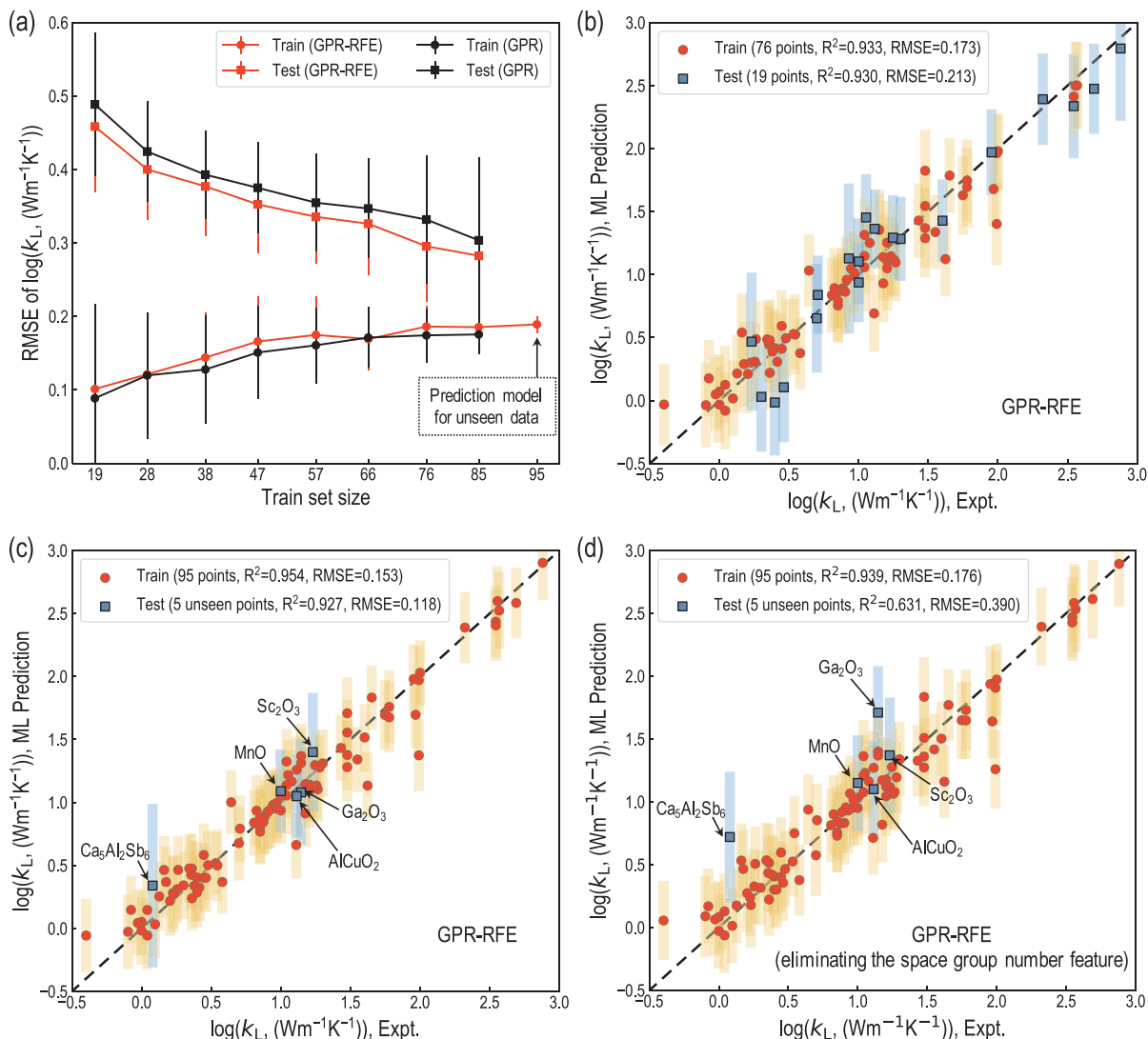

Fig. 4. (a) Prediction accuracy for GPR and GPR-RFE models trained using different train set sizes, averaged over 100 runs. The corresponding test sets in (a) is the difference between total data and train sets. (b) illustrates example parity plot obtained from the GPR-RFE model (29 features) with train and test set of 76 and 19 points, respectively. Parity plots obtained from the GPR-RFE model with 95 train points and 5 unseen test points including, Sc_2O_3, Ga_2O_3, MnO, $AlCuO_2$, and $Ca_5Al_2Sb_6$, using (c) 29 features and (d) 28 features, eliminating the space group number feature from the 29 features.

structure types. If we intentionally eliminated it from the 29 features, the test RMSE of 5 unseen materials increases from 0.12 to 0.39, as shown in Fig. 4(d). This issue is due to the limitation of our present dataset, most of which belong to space groups 225 and 216. As a result, the space group number is required to distinguish materials in terms of their structures in the ML model. However, this problem can be solved when more data with more diverse space groups are included in the training dataset. Furthermore, our present ML model is more suitable for defect-free inorganic materials. There are some accuracy limitations of our model to predict κ_L of materials with defects, allotropic materials and intermetallic compounds. However, the predictive ML model can be more easily improved by actively learning on more diverse training

Table 2

Comparison of this work and other semi-empirical models. 1.38 and 1.46 are reported values from Slack [56] and Debye-Callaway [8] models, respectively.

	This work	Slack	Debye-Callaway
Data set	95	93	55
Cross validation	5-fold	leave-one-out	4-fold
Regression method	GPR-RFE	Kernel ridge	–
AFD of κ_L	1.36 ± 0.03	1.38	1.46

(even temperature-dependent) data sets compared with previous semi-empirical models, due to the easily accessible features.

4. Conclusion

In conclusion, we have developed a simple and general ML model to predict κ_L of inorganic solid materials. This model is faster, and at par or more accurate than traditional physics-based computational methods. This work involves curating a benchmark dataset of experimental values of κ_L of 100 inorganic compounds, generating and optimizing a comprehensive set of features (using the Matminer package), and training the Gaussian Process Regression model on the data prepared. The accuracy of the developed ML models was found to be comparable to past semi-empirical models. Additionally, key features in determining κ_L were identified. Overall, this present work would be useful for rational design and screening of new materials with desired κ_L for specific applications, and fundamentally understanding the heat transport in inorganic solid materials.

Data Availability

The entire experimental κ_L data set and DFT computed bulk modulus are available in Table S1 of the Supporting Information.

CRediT authorship contribution statement

Lihua Chen: Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Huan Tran:** Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Rohit Batra:** Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Chiho Kim:** Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Rampi Ramprasad:** Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare no competing financial interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.commatsci.2019.109155.

References

[1] C. Gayner, K.K. Kar, Recent advances in thermoelectric materials, Prog. Mater Sci. 83 (2016) 330–382.

[2] X. Zhang, L.-D. Zhao, Thermoelectric materials: energy conversion between heat and electricity, J. Materiomics 1 (2) (2015) 92–105.

[3] X. Cao, R. Vassen, D. Stoever, Ceramic materials for thermal barrier coatings, J. Eur. Ceram. Soc. 24 (1) (2004) 1–10.

[4] R. Darolia, Thermal barrier coatings technology: critical review, progress update, remaining challenges and prospects, Int. Mater. Rev. 58 (6) (2013) 315–348.

[5] G. Nolas, H. Goldsmid, Thermal conductivity of semiconductors, Thermal Conductivity, Springer, 2004, pp. 105–121.

[6] D.T. Morelli, G.A. Slack, High lattice thermal conductivity solids, High Thermal Conductivity Materials, Springer, 2006, pp. 37–68.

[7] C. Toher, J.J. Plata, O. Levy, M. de Jong, M. Asta, M.B. Nardelli, S. Curtarolo, High-throughput computational screening of thermal conductivity, debye temperature, and grüneisen parameter using a quasiharmonic debye model, Phys. Rev. B 90 (17) (2014) 174107.

[8] S.A. Miller, P. Gorai, B.R. Ortiz, A. Goyal, D. Gao, S.A. Barnett, T.O. Mason, G.J. Snyder, Q. Lv, V. Stevanovic, Capturing anharmonicity in a lattice thermal conductivity model for high-throughput predictions, Chem. Mater. 29 (6) (2017) 2494–2501.

[9] S.G. Volz, G. Chen, Molecular-dynamics simulation of thermal conductivity of silicon crystals, Phys. Rev. B 61 (2000) 2651–2656.

[10] D.P. Sellan, E.S. Landry, J.E. Turney, A.J.H. McGaughey, C.H. Amon, Size effects in molecular dynamics thermal conductivity predictions, Phys. Rev. B 81 (2010) 214305.

[11] N. Papanikolaou, Lattice thermal conductivity of SiC nanowires, J. Phys.: Condens. Matter 20 (13) (2008) 135201.

[12] C. Carbogno, R. Ramprasad, M. Scheffler, Ab initio green-kubo approach for the thermal conductivity of solids, Phys. Rev. Lett. 118 (2017) 175901.

[13] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, I. Tanaka, Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and bayesian optimization, Phys. Rev. Lett. 115 (20) (2015) 205901.

[14] D.A. Broido, M. Malorny, G. Birner, N. Mingo, D. Stewart, Intrinsic lattice thermal conductivity of semiconductors from first principles, Appl. Phys. Lett. 91 (23) (2007) 231922.

[15] A. Togo, L. Chaput, I. Tanaka, Distributions of phonon lifetimes in brillouin zones, Phys. Rev. B 91 (2015) 094306.

[16] T. Tadano, S. Tsuneyuki, Quartic anharmonicity of rattlers and its effect on lattice thermal conductivity of clathrates from first principles, Phys. Rev. Lett. 120 (2018) 105901.

[17] M.W. Gaultois, T.D. Sparks, C.K. Borg, R. Seshadri, W.D. Bonificio, D.R. Clarke, Data-driven review of thermoelectric materials: performance and resource considerations, Chem. Mater. 25 (15) (2013) 2911–2920.

[18] J. Yan, P. Gorai, B. Ortiz, S. Miller, S.A. Barnett, T. Mason, V. Stevanović, E.S. Toberer, Material descriptors for predicting thermoelectric performance, Energy Environ. Sci. 8 (3) (2015) 983–994.

[19] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, Machine learning in materials informatics: recent applications and prospects, NPJ Comput. Mater. 3 (1) (2017) 54.

[20] A. Mannodi-Kanakkithodi, G.M. Treich, T.D. Huan, R. Ma, M. Tefferi, Y. Cao, G.A. Sotzing, R. Ramprasad, Rational co-design of polymer dielectrics for energy storage, Adv. Mater. 28 (30) (2016) 6277–6291.

[21] T.D. Huan, A. Mannodi-Kanakkithodi, C. Kim, V. Sharma, G. Pilania, R. Ramprasad, A polymer dataset for accelerated property prediction and design, Sci. Data 3 (2016) 160012.

[22] A. Mannodi-Kanakkithodi, A. Chandrasekaran, C. Kim, T.D. Huan, G. Pilania, V. Botu, R. Ramprasad, Scoping the polymer genome: a roadmap for rational polymer dielectrics design and beyond, Mater. Today 21 (7) (2018) 785–796.

[23] A. Seko, H. Hayashi, K. Nakayama, A. Takahashi, I. Tanaka, Representation of compounds for machine-learning prediction of physical properties, Phys. Rev. B 95 (2017) 144110.

[24] L. Ward, A. Dunn, A. Faghaninia, N.E. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, Matminer: an open source toolkit for materials data mining, Comput. Mater. Sci. 152 (2018) 60–69.

[25] E.S. Toberer, A. Zevalkink, G.J. Snyder, Phonon engineering through crystal chemistry, J. Mater. Chem. 21 (40) (2011) 15843–15852.

[26] D.R. Lide, CRC Handbook of Chemistry and Physics vol. 85, CRC Press, 2004.

[27] T. Takahashi, T. Kikuchi, Porosity dependence on thermal diffusivity and thermal conductivity of lithium oxide Li2O from 200 to 900 C, J. Nucl. Mater. 91 (1) (1980) 93–102.

[28] N. Surplice, R. Jones, The thermal conductivity of alkaline earth oxides, Brit. J. Appl. Phys. 14 (10) (1963) 720.

[29] G. Slack, R. Newman, Thermal conductivity of MnO and NiO, Phys. Rev. Lett. 1 (10) (1958) 359.

[30] J.-G. Li, T. Ikegami, T. Mori, Fabrication of transparent Sc_2O_3 ceramics with powders thermally pyrolyzed from sulfate, J. Mater. Res. 18 (8) (2003) 1816–1822.

[31] M. Massot, A. Oleaga, A. Salazar, D. Prabhakaran, M. Martin, P. Berthet, G. Dhalenne, Critical behavior of CoO and NiO from specific heat, thermal conductivity, and thermal diffusivity measurements, Phys. Rev. B 77 (13) (2008) 134438.

[32] S.-I. Yanagiya, N. Van Nong, J. Xu, N. Pryds, The effect of (Ag, Ni, Zn)-addition on the thermoelectric properties of copper aluminate, Materials 3 (1) (2010) 318–328.

[33] N.T. Tinh, T. Tsuji, Thermoelectric properties of heavily doped polycrystalline SrTiO_3, in: D.T. Cat, A. Pucci, K. Wandelt (Eds.), Physics and Engineering of New Materials, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 209–217.

[34] E.S. Toberer, A. Zevalkink, N. Crisoto, G.J. Snyder, The zintl compound $Ca_5Al_2Sb_6$ for low-cost thermoelectric power generation, Adv. Funct. Mat. 20 (24) (2010) 4375–4380.

[35] E.G. Víllora, K. Shimamura, Y. Yoshikawa, T. Ujiie, K. Aoki, Electrical conductivity and carrier concentration control in $\beta - Ga_2O_3$ by Si doping, Appl. Phys. Lett. 92 (20) (2008) 202120.

[36] L. Bjerg, B.B. Iversen, G.K.H. Madsen, Modeling the thermal conductivities of the

zinc antimonides ZnSb and Zn_4Sb_3, Phys. Rev. B 89 (2014) 024304.

[37] Ioffe physico-technical institute, http://www.ioffe.ru/SVA/NSM/Semicond/SiC/thermal.html#Thermal%20conductivity.

[38] J.D. Beasley, Thermal conductivities of some novel nonlinear optical materials, Appl. Opt. 33 (6) (1994) 1000–1003.

[39] C.R. Whitsett, D.A. Nelson, J.G. Broerman, E.C. Paxhia, Lattice thermal conductivity of mercury selenide, Phys. Rev. B 7 (1973) 4625–4640.

[40] G.A. Slack, Thermal conductivity of MgO, Al_2O_3, $MgAl_2O_4$, and Fe_3O_4 crystals from 3 to 300 K, Phys. Rev. 126 (1962) 427–441.

[41] P. Turkes, C. Pluntke, R. Helbig, Thermal conductivity of SnO_2 single crystals, J. Phys. C: Solid State Phys. 13 (26) (1980) 4941–4951.

[42] T. Sekimoto, K. Kurosaki, H. Muta, S. Yamanaka, Thermoelectric properties of (Ti, Zr, Hf)CoSb type half-heusler compounds, Mater. Trans. 46 (7) (2005) 1481–1484.

[43] Y. Kawaharada, K. Kurosaki, H. Muta, M. Uno, S. Yamanaka, High temperature thermoelectric properties of CoTiSb half-heusler compounds, J. Alloys Compd. 384 (1) (2004) 308–311.

[44] Y. Xia, S. Bhattacharya, V. Ponnambalam, A. Pope, S. Poon, T. Tritt, Thermoelectric properties of semimetallic (Zr, Hf) CoSb half-heusler phases, J. Appl. Phys. 88 (4) (2000) 1952–1955.

[45] D. Young, P. Khalifah, R. Cava, A. Ramirez, Thermoelectric properties of pure and doped FeMSb (M = V, Nb), J. Appl. Phys. 87 (1) (2000) 317–321.

[46] H. Hohl, A.P. Ramirez, C. Goldmann, G. Ernst, B. Wlfing, E. Bucher, Efficient dopants for ZrNiSn-based thermoelectric materials, J. Phys.: Condens. Matter 11 (7) (1999) 1697–1709.

[47] Caesium bromide (CsBr), https://www.korth.de/index.php/162/items/12.html.

[48] M. Sist, K.F.F. Fischer, H. Kasai, B.B. Iversen, Low-temperature anharmonicity in cesium chloride (CsCl), Angew. Chem. Int. Ed. Engl. 129 (13) (2017) 3679–3683.

[49] Caesium iodide (CsI), https://www.janis.com/Libraries/Window_Transmissions/CaesiumIodide_CsI_TransmissionCurveDataSheet.sflb.ashx.

[50] K. Kamran, M.A. ur Rehman, A. Maqsood, Thermal and electrical properties of crystalline silver bromide, J. Phys. D: Appl. Phys. 40 (3) (2007) 869–873.

[51] B. Hakansson, R.G. Ross, Thermal conductivity and heat capacity of solid LiBr and RbF under pressure, J. Phys.: Condens. Matter 1 (25) (1989) 3977–3985.

[52] M. Pawlak, F. Firszt, S. Łegowski, H. Meczyńska, J. Gibkes, J. Pelzl, Thermal transport properties of $Cd_{1-x}Mg_{xSe}$ mixed crystals measured by means of the photopyroelectric method, Int. J. Thermophys. 31 (1) (2010) 187–198.

[53] J.J. Martin, Thermal conductivity of Mg_2Si, Mg_2Ge and Mg_2Sn, J. Phys. Chem. Solids 33 (1972) 1139–1148.

[54] R.J. LaBotz, D.R. Mason, The thermal conductivities of Mg_2Si and Mg_2Ge, J. Electrochem. Soc. 110 (1963) 121–126.

[55] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, The Materials Project: A materials genome approach to accelerating materials innovation, APL Mater. 1 (1) (2013) 011002.

[56] Y. Zhang, C. Ling, A strategy to apply machine learning to small datasets in materials science, NPJ Comput. Mater. 4 (1) (2018) 25.

[57] C. Kim, A. Chandrasekaran, A. Jha, R. Ramprasad, Active-learning and materials design: the example of high glass transition temperature polymers, MRS Commun. (2019) 1–7.

[58] R. Batra, G. Pilania, B.P. Uberuaga, R. Ramprasad, Multifidelity information fusion with machine learning: a case study of dopant formation energies in hafnia, ACS Appl. Mater. Interfaces 11 (28) (2019) 24906–24918.

[59] G. Pilania, J. Gubernatis, T. Lookman, Multi-fidelity machine learning models for accurate bandgap predictions of solids, Comput. Mater. Sci. 129 (2017) 156–163.