Estimation of the Thermal Conductivity of Porous Silicon Using Molecular Dynamics and Machine Learning Methods

V. V. Kuryliuk, O. Ya. Olikh

Taras Shevchenko National University of Kyiv,  
64/13, Volodymyrska Str., Kyiv, 03028, Ukraine

e-mail: olegolikh@knu.ua

Porous media are widely used in various applications, such as thermal barriers or thermoelectric materials for enhancing the figure of merit. In particular, porous silicon (p-Si) layers can be effectively utilized in high-efficiency solar-thermal energy storage systems. Efficiently predicting their thermal conductivity (TC) can accelerate the design of porous systems to improve energy efficiency, among other benefits [1]. The accurate TC estimation can be obtained through molecular dynamics (MD) simulation. However, such calculations require high computational and time costs. One way to reduce expenses and obtain results for a vast number of cases is to use artificial intelligence methods [2].

This study will focus on applying machine learning (ML) techniques for predicting the TC values of p-Si from the results of data from MD simulation. MD calculation проводилися з використанням пакету. Значення ТС визначалися за кінетичними кривими кореляції теплового потоку. При цьому взаємодія атомів розглядалася в наближенні Tersoff потенціалу. We calculated the temperature dependence of TS for porosities of 0%, 20%, and 40%, as well as the dependence of TS on porosity at 300 K – about 30 values in total (see Fig).

The application of Machine Learning (ML) involved two distinct approaches. The first one utilized Symbolic Regression (SR), a method capable of generating analytical equations without prior constraints. The resulting expression obtained using the PySR package is as follows:

 (1)

where *p* is the porosity (0 ≤ *p* < 1), *Tn* = *T* / 300 is the normalized temperature. The mean percentage error for Eq. (1), compared to MD calculations, is less than 3%.

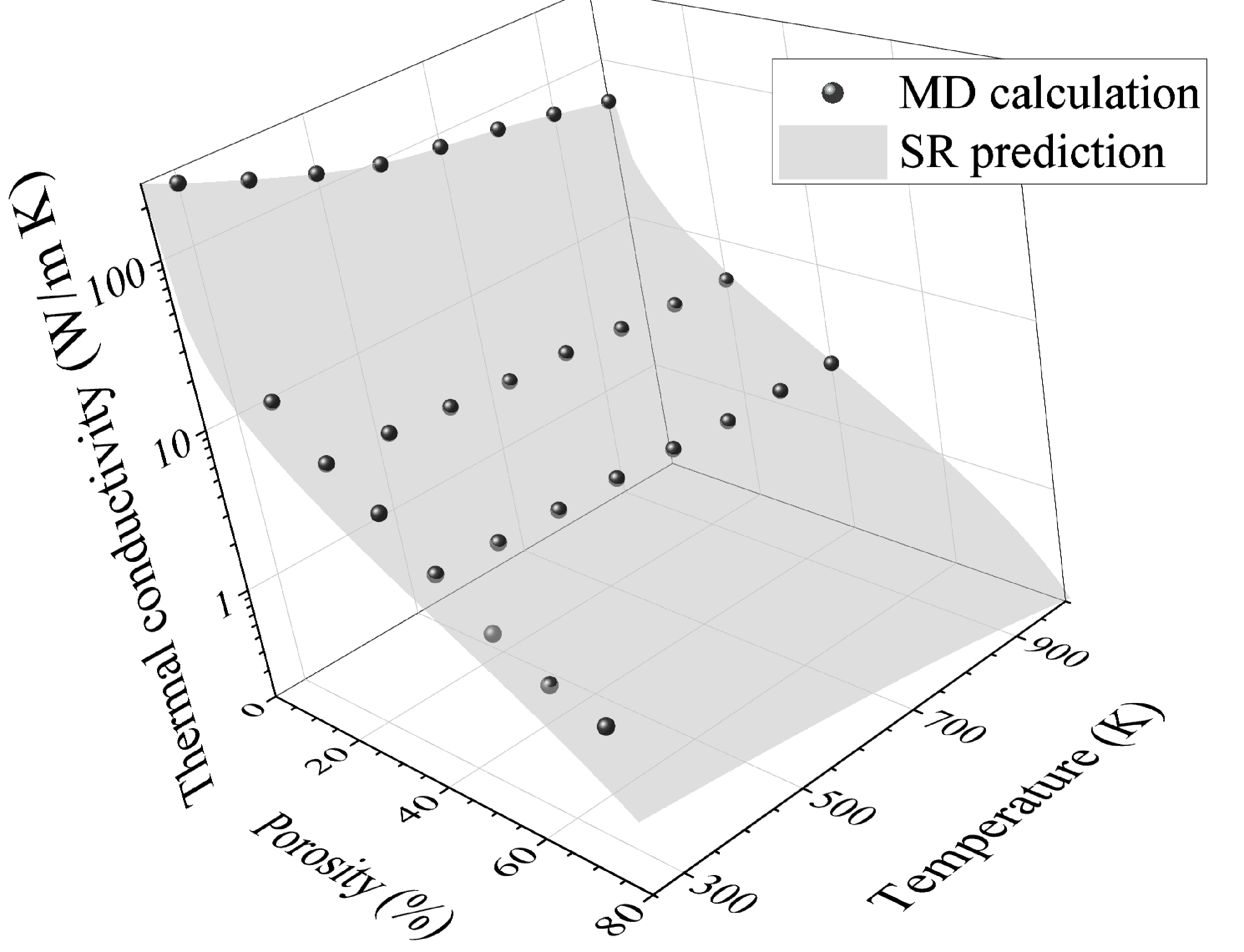


Fig. Dependence of the TC of porous silicon on temperature and porosity. The balls represent the results of MD calculations, while the surface corresponds to Eq. (1).

The second approach employed Random Forest (RF), Gradient Boosting (GB), and Support Vector Regression (SVR) algorithms to construct kinetic correlation curves for heat flux. The resulting thermal conductivity values, predicted for *T* and *p* values not covered by MD calculations, exhibited median percentage errors of 14% for RF, 15% for GB, and 45% for SVR, compared to the values predicted by Eq. (1). Notably, prediction with a trained ML model takes a split second on a laptop, while the MD calculation takes about a week on a cluster.

The work was supported by National Research Foundation of Ukraine (project No. 2023.05/0024).

[1] J. Xu, H. Wei, and H. Bao, Int. J. Heat Mass Transf. 217, 124671 (2023). https://doi.org/10.1016/j.ijheatmasstransfer.2023.124671.

[2] H.-G. Nguyen, T.-D. Le, H.-G. Nguyen, and T.-H. Fang. Mater. Sci. Eng. R 160, 100833 (2024). https://doi.org/10.1016/j.mser.2024.100833.