

ORIGINAL RESEARCH ARTICLE

Machine-learned molecular modeling of ruthenium: A Kolmogorov-Arnold Network approach

 Zhiyu An¹  and Jingjie Yeo^{2*} 
¹Department of System Engineering, Cornell University, Ithaca, New York, United States of America

²Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York, United States of America

 (This article belongs to the *Special Issue: Applications of Deep Learning in Advanced Materials Processing*)

Abstract

Developing refractory high-entropy superalloys (RSAs) with performance advantages over nickel-based alloys is a critical frontier in materials science. Body-centered cubic (bcc)-based RSAs have attracted significant attention, with ruthenium (Ru) playing a key role in forming two-phase regions of A2 (disordered bcc) + B2 (ordered bcc), which could lead to superalloy-like microstructures. This study introduces the application of the Kolmogorov-Arnold Network (KAN) model to predict the mechanical and thermodynamic properties of Ru while comparing its performance against other commonly used machine-learned models. Utilizing density functional theory calculations as training data, the KAN model demonstrates superior accuracy and computational efficiency compared to conventional methods, while reducing descriptor complexity. The model accurately predicts a range of properties, including elastic constants, thermal expansion coefficients, and various moduli, with discrepancies within 6% of experimental reference data. Molecular dynamics simulations further validate the model's efficacy, accurately capturing Ru's phase transitions from hexagonal close-packed (hcp) to face-centered cubic structure and the melting point. This work presents the first application of KAN in materials science, demonstrating how its balanced performance and efficiency provide a new pathway for designing advanced materials, with unique advantages over conventional machine learning approaches in predicting material properties.

Keywords: Ruthenium; Kolmogorov-Arnold Network; Machine learning; Mechanical properties; Thermodynamic properties; Density-functional theory; Molecular dynamics

*Corresponding author:

 Jingjie Yeo
 (jingjieyeo@cornell.edu)

Citation: An Z, Yeo J. Machine-learned molecular modeling of ruthenium: A Kolmogorov-Arnold Network approach. *Int J AI Mater Design*. 2025;2(1):21-38. doi: 10.36922/ijamd.8291

Received: December 30, 2024

Revised: January 26, 2025

Accepted: February 7, 2025

Published online: February 25, 2025

Copyright: © 2025 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Researchers have long sought to improve nickel-based superalloys for greater technical, economic, and social benefits by reducing their weight while improving their material properties under extreme loads. Naka and Khan¹ proposed a promising direction by combining B2-ordered NiAl compounds with transition metals having disordered body-centered cubic (bcc) structures (A2). This approach creates an A2 + B2 microstructure like the γ - γ' structure in traditional superalloys. These alloys, which meet the criteria for

high-entropy alloys (HEAs)^{2,3} and complex concentrated alloys,⁴ demonstrate potential as refractory high-entropy superalloys (RSAs). In A₂+B₂ systems, the ordered intermetallic B₂ phase is particularly important, as it often acts as a strengthening precipitate, similar to the role of γ' in traditional superalloys.⁵ In addition, B₂ phases often exhibit excellent strength retention at elevated temperatures and creep resistance compared to disordered phases, enhancing the alloy's performance under sustained loads at high temperatures.⁶ Many B₂ phases, particularly those containing Al, demonstrate superior oxidation resistance by forming protective oxide scales. However, some elements discovered by Naka and Khan are still not fully represented in RSAs that are currently synthesized, especially Pd, Pt, and ruthenium (Ru), which all form stable B₂ phases with Al.⁷

Among the elements critical to RSA development, Ru has garnered significant attention due to its unique properties and potential ability to form stable ordered phases in multi-component systems. Cerba *et al.*⁸ demonstrated that AlRu alloys melt at 2060°C and interact with Ru through eutectic reactions at 1920°C, suggesting potential routes for exploring A₂ + B₂ microstructures. Moreover, the B₂ phase exists in some binary refractory metal systems involving Ru, such as Nb-Ru and Ru-Zr,⁷ indicating its versatility in phase formation. Despite its importance, there remains a significant gap in our understanding of Ru's molecular structure and its influence on mechanical and thermodynamic properties, especially in the context of computational materials science. Specifically, there is currently no machine learning (ML) model capable of accurately predicting Ru's molecular structures along with their corresponding mechanical properties in a wide range of temperatures. This lack of comprehensive data and predictive tools hinders our ability to engineer the structures and material properties of RSAs where Ru is used. The corresponding knowledge gap hinders the efficient design and development of advanced RSAs and other high-temperature materials that incorporate Ru.

Ru is a platinum-group metal that crystallizes in a hexagonal close-packed (hcp) structure at ambient conditions.⁹ Experimentally, Ru displays superconductivity with a critical temperature below 1699 K.¹⁰ A hcp Ru also maintains its structural stability up to pressures of 150 GPa and temperatures of 960 K.¹¹ Güler *et al.*¹² used molecular dynamics (MD) simulations with an empirical interatomic potential to explore the temperature-dependent elastic, mechanical, and anisotropic properties of hcp Ru. Likewise, Lu *et al.*¹³ used density functional theory (DFT) to calculate the high-pressure phase transition of Ru. However, computational approaches, such as DFT and

MD, are often computationally intensive and limited in their ability to rapidly explore the vast parameter space of Ru's structure-property relationships. The development of an ML model that can accurately and rapidly predict Ru's molecular structures and associated mechanical properties would be a significant advancement. Such a model could potentially allow for a faster and more comprehensive exploration of Ru's structural and mechanical behavior under various environmental conditions.

High-throughput *ab initio* simulations are increasingly used for exploring the basic properties of engineering materials, such as determining the material properties.¹⁴ When combined with materials informatics, ML models accelerate the exploration of these material properties.¹⁵ The goal of the ML model is to supplement or replace high-fidelity modeling methods at the quantum chemical or classical level to predict the properties of molecules or materials directly from their structure or their chemical composition. ML methods usually use available experimental and *ab initio* data to build accurate statistical models that can be used to predict the properties of materials. This innovative combination can significantly enhance predictive capabilities and has applications in many fields of materials research, such as dielectric polymers,¹⁶ critical temperatures of superconducting materials,¹⁷ crystal structures,¹⁸ perovskites,¹⁹ and nanostructures.²⁰ Many of these properties were measured with high accuracy using ML approaches. For instance, de Jong *et al.*²¹ developed a model using supervised ML with gradient-boosting regression to predict elastic properties of inorganic polycrystalline compounds, such as the bulk modulus K and shear modulus G . For metallic systems, Song *et al.*²² constructed a general-purpose neural network potential (NNP) for 16 elemental metals and their alloys that achieves superior accuracy compared to traditional embedded-atom method potentials while maintaining computational efficiency. Their potential accurately captures complex interface behavior, phase stability, and mechanical properties across diverse metallic systems. Liyanage *et al.*²³ developed a NNP for the Cu-W system using the Behler-Parrinello framework. Their potential accurately reproduces metallurgically relevant properties, including elasticity, stacking faults, dislocations, and thermodynamic behavior in both elemental Cu and W, as well as Cu-W interfaces and solid solutions. This NNP enables large-scale atomistic simulations to investigate phenomena such as the influence of interface stress on mechanical properties in Cu-W nano-multilayer systems. Another advantage of ML models is their generalizability across different material systems, as it is easier to incorporate crystal data with defects.²⁴ Compared with other interatomic potentials, ML potentials are highly

flexible, allowing for facile adjustment of parameters and refinement of datasets to improve accuracy and predictive performance.²⁵ This flexibility allows ML models to adapt to new data and evolve continuously, providing a dynamic and scalable approach for studying a wide range of materials and their complex behaviors. In addition, ML models enable faster updates as more data becomes available, streamlining the process of material discovery and optimization.

Among these ML models, the recently introduced Kolmogorov-Arnold Network (KAN), inspired by the Kolmogorov-Arnold Representation theorem, provides a new way to model complex systems.²⁶ Unlike multilayer perceptrons (MLPs) with fixed activation functions on nodes, KAN has a learnable activation function on the edge, parameterized as a spline function. This seemingly simple change enables KAN to outperform MLP in terms of accuracy and interpretability. In data fitting and partial differential equation solving tasks, KAN has demonstrated faster neural scaling laws than MLP, achieving comparable or better accuracy with a smaller network.²⁷ The architectural advantage of KAN makes it particularly suited for materials science applications. Its ability to process complex, non-linear relationships makes it ideal for tackling quantum phenomena, like Anderson localization, where disorder in a system leads to the localization of electronic wave functions.²⁸ KANs excel in extracting mobility edges in various tight-binding models, such as the Mosaic model²⁹ and the Aubry-André model,³⁰ where previous methods may struggle. Material properties often follow complex, non-linear relationships that cannot be well-captured by standard activation functions, making KAN's learnable activation functions especially valuable in this domain. The model's ability to discover and represent underlying physical patterns without requiring explicit physical constraints in the training process represents a significant advancement in materials informatics. By allowing users to incorporate prior knowledge and assumptions, KAN can collaborate with human intuition to simplify complex expressions while maintaining high accuracy. This makes KAN a powerful tool for predicting physical properties in materials and provides a robust framework for handling disorders and defects in computational material studies. Therefore, integrating KAN into the materials informatics workflow can enhance our understanding of complex material phenomena, accelerate the discovery of new materials with desired properties, and provide results that are easier to interpret than traditional neural networks.

Our study addresses the lack of ML models specifically tailored for Ru and introduces the novel application of KAN to predict material properties and construct ML interatomic

potentials (Figure 1). Compared to other established neural network architectures including CalHousNet Feedforward Neural Network (CalHousNet), Accurate Neural Network Engine for Molecular Energies (ANI), Continuous-filter Convolutional Neural Network (SchNet), and traditional graph neural networks (GNNs), our approach simplifies data preprocessing and significantly reduces training time. When benchmarked against experimental data, our model demonstrates superior accuracy in predicting mechanical properties using molecular statics simulations, highlighting the precision of KAN. Furthermore, we successfully generated an interatomic potential capable of accurately capturing key physical phenomena using MD simulations, including phase transitions and the melting point of Ru. This development not only underscores the advantages of KAN in computational materials science but also lays the groundwork for incorporating additional elements into the framework. This advancement enables the development of multi-element interatomic potentials, potentially facilitating accurate simulations of more complex high-entropy material systems.

2. Methods

2.1. Molecular models and simulations

First-principles calculations were performed within DFT through the Quantum Open-Source Package for Research in Electronic Structure, Simulation, and Optimization (Quantum ESPRESSO) package,^{31,32} using the Perdew-Burke-Ernzerhof³³ exchange-correlation functional and projector-augmented plane wave³⁴ potentials. We optimized the wavefunction and charge density parameters to cutoffs of 52.123 and 353.301 Ry, respectively, and used a $4 \times 4 \times 4$ Monkhorst-Pack grid for k-point sampling. With this scheme, the total energy and force converged at 10^{-5} Ry and 10^{-4} Ry/Bohr, respectively. Through these calculations, we generated 699 data points to relate the changes in the volume and energy of the element Ru under various stress conditions, facilitating detailed analysis of its structural properties. From the DFT calculations, we obtained the initial crystal structure parameters, including lattice constants (a , b , c), angles (α , β , γ), unit cell volume, and atomic positions in both fractional (a , b , c) and Cartesian (x , y , z) coordinates. The DFT calculations further provided the corresponding total energies, stress tensors, and atomic forces. This combined dataset was then processed and structured into a comprehensive dataset using the pymatgen (Python Materials Genomics) package.³⁵ This dataset size is comparable to typical single-element training sets in literature; for instance, previous studies have demonstrated that accurate potentials can be developed with 461 and 284 structures for Ni and Mo, respectively.³⁶ While multi-element systems typically

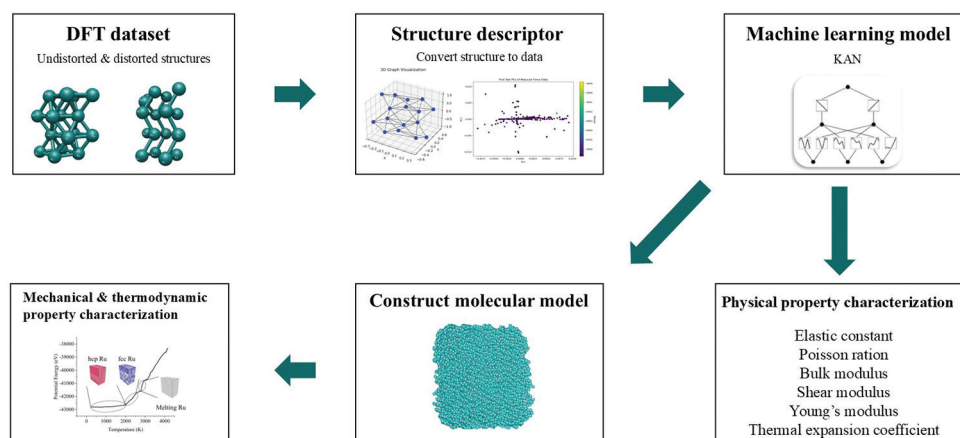


Figure 1. Schematic of the workflow of constructing the dataset using DFT calculations, converting the structures into machine-learnable descriptors, constructing the ML model using KAN, and measurements of material properties either directly or by constructing ML interatomic force fields for MD simulations.

Abbreviations: DFT: Density functional theory; ML: Machine learning; KAN: Kolmogorov-Arnold Network; MD: Molecular dynamics.

require thousands of structures to capture diverse atomic environments and interactions, single-element systems can achieve good accuracy with several hundred carefully selected configurations that comprehensively sample the relevant phase space.

MD simulations were performed using the open-source Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software package to calculate elastic constants and melting point.³⁷ The system was constructed using periodic boundary conditions with an hcp lattice structure in a $10 \times 10 \times 10$ unit cell configuration. Interatomic interactions were described using our tabulated KAN potential combined with a Lennard-Jones potential in a hybrid/overlay scheme. To calculate the elastic constants, the system was constructed with periodic boundary conditions using a lattice parameter of 2.70 Å, which was adopted from crystallographic data for Ru's hcp structure and documented in the crystallographic database maintained by Springer Materials.³⁸ We then performed energy minimization of this initial structure with a convergence criterion of 10^{-6} eV/Å for forces and 10^{-8} eV for energy, followed by anisotropic box relaxation. The elastic constants were then calculated through strain-stress relationships: C_{11} , C_{12} , and C_{13} were determined by applying uniaxial strain, for example, by applying a strain along the x-axis and measuring the resulting stress response in the x-, y-, and z-directions. C_{44} was determined through shear deformation in the xy-plane using a triclinic box transformation. All calculations were performed under quasi-static conditions, with system relaxation achieved through energy minimization after each deformation step. The melting point was determined using a heating method, where the system was first equilibrated at 300 K using a

constant Number, Pressure, and Temperature (NPT) ensemble, followed by gradual heating from 300 to 3000 K. The temperature was controlled using a Nosé-Hoover thermostat with a damping parameter of 1.0, and the pressure was maintained at 0 bar using a Parrinello-Rahman barostat during the NPT phase. Neighbor lists were updated every timestep with a cutoff distance of 0.3 Å and a bin-based approach.

2.2. Training data generation

Developing robust ML models for material science heavily relies on training data that encompasses a wide variety of atomic environments. To this end, we performed structure relaxation on all symmetrically distinct configurations within a 16-atom supercell of Ru arranged in an hcp structure. Our objective was to meticulously explore the potential energy surface by optimizing these configurations to their lowest energy states, ensuring the atomic positions and lattice parameters closely matched experimental and theoretical benchmarks. This optimization would ensure that the material's behavior under varying temperatures and pressures was captured accurately. We focused on three primary structural types in our study:

- Undistorted ground state structure: This represents the element's most stable configuration, free from external strains or forces.
- Distorted structures: By applying strains ranging from -10% to $+10\%$ in six distinct modes – uniaxial tension, uniaxial compression, biaxial tension, biaxial compression, shear, and torsional strain – to the bulk conventional cell, we generated atomic structures to analyze the material's behavior under mechanical stress.³⁹

2.3. Training data processing

To effectively process crystal structures for ML applications, the Python Materials Genomics (pymatgen) Library³⁵ was utilized in conjunction with the Smooth Overlap of Atomic Positions (SOAP) descriptor, implemented through the Dscribe Python package,⁴⁰ to convert structural data into tensor representations. SOAP provides a robust framework for representing atomic environments, as it ensures rotational, translational, and permutational invariance of the structural descriptors. This invariance means that equivalent atomic configurations yield identical representations regardless of rigid body transformations or reordering of atoms, which is essential for reliable structure-property predictions in crystalline systems. Each structure's atomic positions were denoted as $r_i = x_i y_i z_i$, where i denotes the index of the atom. These position vectors were flattened into a one-dimensional tensor for each structure:

$$T_{structure} = x_1 y_1 z_1 \dots x_n y_n z_n \quad (I)$$

where n is the total number of atoms in the structure. The tensors from each structure were then stacked to form a two-dimensional tensor T , representing the entire dataset and making it suitable for ML models:

$$T = \begin{bmatrix} T_{structure 1} \\ \dots \\ T_{structure m} \end{bmatrix} \quad (II)$$

where m is the number of structures.

To address the high dimensionality of our data, we applied principal component analysis (PCA). We centered the data by subtracting the mean of each feature to produce a mean-centered data matrix:

$$X_{centered} = X - \mu \quad (III)$$

where X is the original data matrix and μ is a vector containing the mean values of each feature.

Next, we computed the covariance matrix C from the mean-centered data:

$$C = \frac{1}{m-1} X_{centered}^T X_{centered} \quad (IV)$$

where m is the number of structures.

We then performed eigen decomposition to extract eigenvalues (λ) and eigenvectors (V), satisfying:

$$CV = \Lambda V \quad (V)$$

where V is the matrix of eigenvectors (principal components [PCs]) and Λ is the diagonal matrix of

eigenvalues. The data were then projected onto the PCs to obtain a reduced data matrix X :

$$X = X_{centered} V_{reduced} \quad (VI)$$

where $V_{reduced}$ includes eigenvectors corresponding to the largest eigenvalues that capture 95% of the variance.

In the PCA of crystal structures, Figure 2A elucidates the cumulative explained variance ratio as a function of component number. The first PC accounts for 80.5% of the total variance, demonstrating its dominant role in capturing the dataset's primary features. The second PC contributes an additional 12.2%, bringing the cumulative explained variance to 92.7%. Notably, the first three PCs collectively explain 97.6% of the dataset's variance, surpassing the often-used 95% threshold for dimensionality reduction.⁴¹ This rapid accumulation of explained variance, visualized by the steep rise in the cumulative contribution line, exhibits a characteristic elbow-shaped curve. The shape of this curve, with its sharp initial increase followed by a plateau, indicates highly efficient dimensionality reduction, suggesting that the complex crystal structure data can be effectively represented using just these three PCs.

Figure S1 presents a heatmap of structural PCA loadings, revealing the complex relationships between the original structural features and the three PCs. In this context, "loadings" refer to the coefficients that describe how much each original feature contributes to a given PC, with red indicating positive correlations and blue indicating negative correlations. The first PC displays a nuanced pattern of strong positive and negative loadings across features, suggesting that it encapsulates a multifaceted combination of structural attributes. This component exhibits the most variation in loadings, indicating its capacity to capture complex, opposing relationships among the original features. In contrast, the second PC exhibits predominantly positive loadings, with several features displaying very strong positive correlations. Importantly, the magnitude of loadings across all PCs indicates that each original structural feature is well-represented in the PC space, suggesting that the PCA effectively captures the key variance in the dataset without substantial loss of information from any feature.

Meanwhile, for energy and force data extracted from DFT calculations, we employed different preprocessing techniques. The energy data were directly converted into the tensor format without further modifications. In contrast, the force data, representing a three-dimensional vector (one per atom with components $F_i = (F_{ix}, F_{iy}, F_{iz})$ representing the forces along the x -, y -, and z -directions), required additional processing. We also applied PCA to reduce its dimensionality, as discussed above. Figure 2B

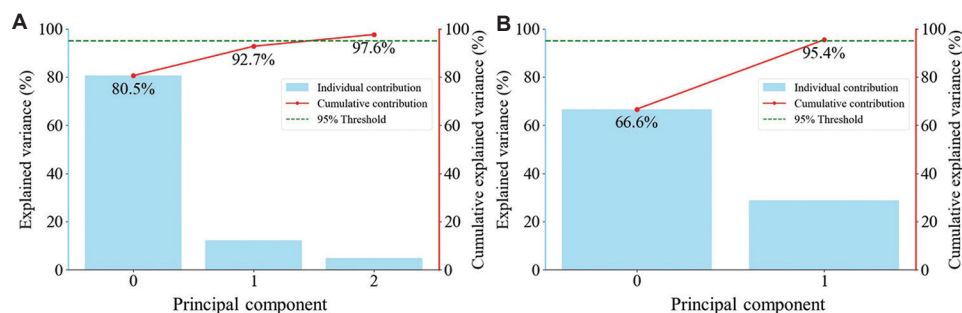


Figure 2. Principal component analysis (PCA) reveals dominant structural features and efficient dimensionality reduction in the structure and force data. (A) Cumulative explained variance ratio vs. structure component number. PC0 accounts for 80.5% of the total variance, with PC1 adding 12.2%, PC2 adding 4.9%, and the first three PCs collectively explaining 97.62%, exceeding the 95% threshold. (B) Cumulative explained variance ratio vs. force component number. PC0 accounts for 66.6% of the total variance, with PC1 adding 28.8%, exceeding the 95% threshold. Abbreviation: PC: Principal component.

elucidates the cumulative explained variance ratio as a function of component number. The first PC accounts for 66.6% of the total variance, demonstrating its dominant role in capturing the dataset's primary features. The second PC contributes an additional 28.8%, bringing the cumulative explained variance to surpass the often-used 95% threshold for dimensionality reduction. Figure S2 presents a heatmap of force PCA loadings, indicating that each original force feature is well-represented in the PC space, suggesting that the PCA effectively captures the key variance in the dataset without substantial loss of information from any feature.

After performing PCA on the structural data, we normalized the target outputs to address the inherent scale differences between energy and force values, such that:

$$Z = \frac{X - \mu}{\sigma} \quad (\text{VII})$$

where Z is the normalized value, X is the original value, μ is the mean of the data, and σ is the standard deviation of the data. This normalization is crucial as energy and force are measured in different units and exhibit distinct ranges of magnitudes, which could potentially bias the model training. By normalizing these quantities, we ensure that both energy and force components contribute equally to the loss function during model training, preventing the larger-scale variable from dominating the optimization process. This standardized representation of target variables helps maintain balanced gradients during training and enables the model to learn the relationships between structural features and both energy and force predictions with equal emphasis. The normalization process preserves the underlying data patterns while adjusting the scale, and the numerical ranges are standardized to comparable scales (Figure 3). This consistency in patterns before and after normalization confirms that our preprocessing approach

effectively standardizes the data without distorting its inherent features.

Subsequently, we optimized the training process by leveraging graphical processing units (GPUs), which allowed for significantly faster computation and more efficient handling of complex tensor operations required by our ML algorithms. We configured our computational framework to dynamically allocate tasks across available GPUs, maximizing resource utilization and reducing processing times. The dataset is split into 60% for training, 20% for validation, and 20% for testing to ensure a comprehensive evaluation of model performance. As displayed in Figure 4, the distribution of samples across the PCs exhibits consistent cross-shaped patterns across all three subsets. This similarity in data distribution confirms that our splitting strategy has effectively maintained the representativeness of the original dataset in each subset. The validation set allows us to monitor model performance during training and prevent overfitting, while the test set provides an unbiased evaluation of the final model performance.

2.4. KAN model

The KAN model was configured with a network width of [3, 4, 3] and utilized a grid size of 8 across three dimensions ($k = 3$).²⁷ To enhance the training process, we employed the Adam optimizer, a stochastic gradient descent method known for its ability to compute adaptive learning rates for each parameter.⁴² This optimization method is particularly effective in managing sparse gradients and dynamically adjusting the learning rate, which significantly improves the speed of convergence and the overall performance of the model.

By adopting the Adam optimizer, we ensured that our model efficiently navigated through the complex loss landscape, which led to faster convergence and improved predictive accuracy. In addition, we implemented L2

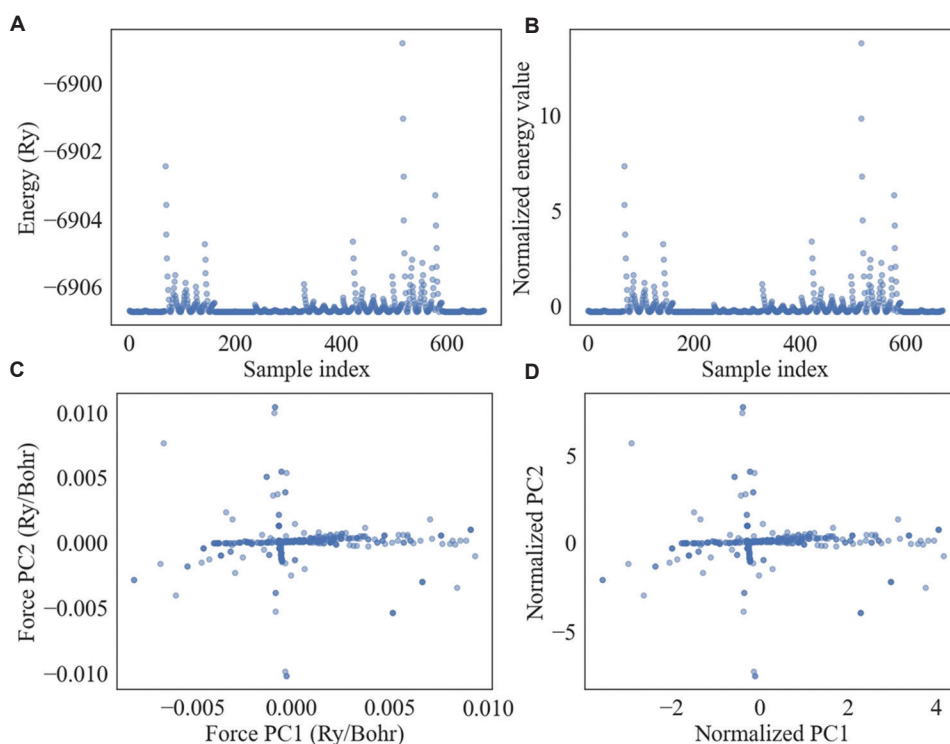


Figure 3. Distribution patterns before and after normalization. (A) Original energy values of the dataset samples; (B) normalized energy values displaying preserved distribution pattern; (C) original force features in the PC1-PC2 space; and (D) normalized force features demonstrating maintained relative relationships. The consistent patterns between pre- and post-normalization data validate the effectiveness of the standardization process.

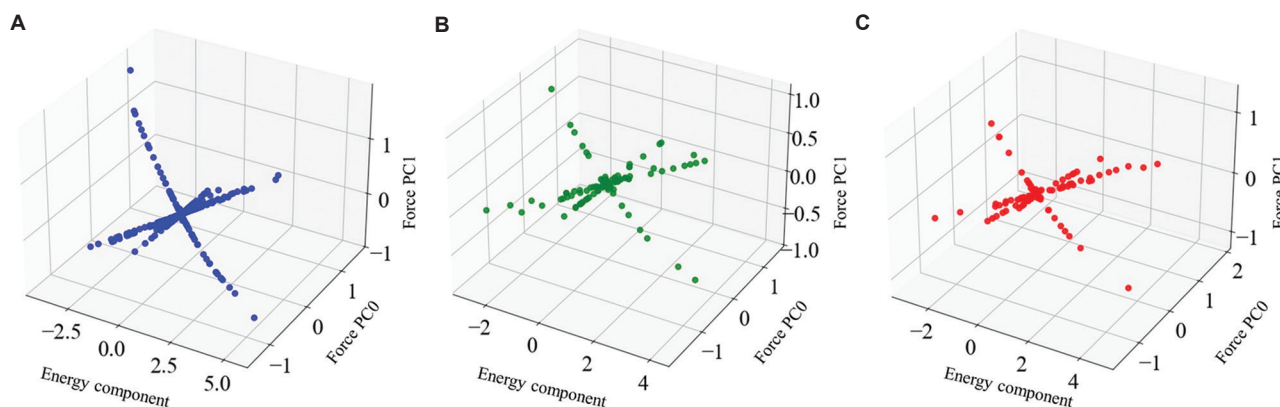


Figure 4. Scattering plots of (A) training dataset, (B) validation dataset, and (C) test dataset, demonstrating consistent cross-shaped patterns across all three subsets.

regularization ($\lambda = 0.02$)⁴³ and entropy regularization ($\lambda_{\text{entropy}} = 4.0$)³² to enhance the model's ability to generalize. This is crucial for preventing overfitting and maintaining robust performance on new, unseen data.

2.5. CalHousNet feedforward neural network model (CalHousNet)

We implemented CalHousNet, a fully connected feedforward neural network using the PyTorch framework.⁴⁴

The model employs a sequential architecture comprising three distinct layers designed for regression analysis. The input layer accepts three features, corresponding to the reduced-dimensionality data obtained through PCA of atomic structures. These inputs are processed through a hidden layer containing three nodes, where each node implements a linear transformation followed by a ReLU activation function.⁴⁵ This activation introduces essential non-linearity into the model, enabling the capture

of complex relationships between input features and target outputs. The output layer consists of a single node performing a final linear transformation, specifically configured for continuous value prediction in regression tasks.

The training process employs the Adam optimizer with a learning rate of 0.001.⁴² This optimizer was selected for its ability to dynamically adjust learning rates for each parameter based on first- and second-moment estimates of the gradients, enhancing both convergence speed and model performance. The training procedure processes data in batches to optimize memory usage and computational efficiency, leveraging stochastic gradient descent for gradient computation. This batch processing approach, combined with multiple training epochs, allows the model to progressively refine its predictions through repeated iterations over the complete dataset.

2.6. Graph neural network model

We developed a GNN model specifically designed to handle the structural complexities of Ru crystals. The model presents atomic structures as graphs, where each node corresponds to an atom and incorporates both static and dynamic properties. The node feature vector combines the constant atomic number ($Z = 44$ for Ru) with the components of the force vector acting on each atom (F_{ix}, F_{iy}, F_{iz}). Thus, each node incorporated both elemental and dynamic properties and is represented as:

$$x_i = [Z, F_{ix}, F_{iy}, F_{iz}] \quad (\text{VIII})$$

The edge structure of the graph is determined by a predefined connectivity matrix characteristic of Ru crystal structures. Each edge carries geometric information, specifically the Euclidean distance between connected atoms, calculated from their respective position vectors. This distance-based edge attribution ensures that spatial relationships between atoms are explicitly encoded in the graph structure, allowing the model to learn from both local atomic environments and their spatial arrangements:

$$d_{ij} = ||r_i - r_j|| \quad (\text{IX})$$

where r_i and r_j are the position vectors of atoms i and j , respectively. The neural network architecture implements a hierarchical processing scheme through multiple graph convolutional layers. The initial layer transforms the four-dimensional input features into a 64-dimensional latent space, initiating the abstract representation of the atomic environment. The network progressively expands this representation through subsequent layers, first to 128 dimensions and then to 256 dimensions. Each convolutional operation is followed by a ReLU activation

function,⁴⁵ introducing non-linearity and sparsity into the representation, thereby focusing the model's attention on the most significant features of the atomic structure.

The feature aggregation process culminates in a global max pooling operation that consolidates information across the entire graph structure. This pooled representation then passes through a series of fully connected layers, systematically reducing the dimensionality from 256 to 128, and finally to the output value representing the predicted energy and force of the crystal structure.

The model optimization utilizes the Adam optimizer with an initial learning rate of 0.001, carefully chosen to balance training efficiency with optimization stability. This configuration allows for effective convergence while mitigating the risk of overshooting the optimal parameters during training. The gradient-based optimization process enables the model to learn the complex relationships between atomic structure and crystal energy while maintaining numerical stability throughout the training procedure.

2.7. Accurate neural network engine for molecular energies model

We implemented a standard ANI model following the original framework specifications.⁴⁶ The model utilizes an atomic environment vector (AEV) computed with radial and angular cutoffs of 5.2 and 3.5 Å, respectively.⁴⁷ The radial basis functions employ EtaR parameters of 16.0, 8.0, 4.0, and 2.0, with corresponding ShfR values of 0.0, 0.9, 1.8, and 2.7. The angular component uses EtaA values of 8.0, 4.0, and 2.0, a zeta parameter of 32.0, and both ShfA and ShfZ values set to 0.0, 0.9, and 1.8.

The neural network architecture consists of three primary blocks processing the AEV inputs. The first block transforms the input to 128 dimensions using a linear layer with SELU activation,⁴⁸ followed by layer normalization and 0.1 dropout. The second block maintains this dimension with identical transformation and regularization schemes. The third block halves the dimensionality while retaining SELU activation and layer normalization. The network concludes with a linear transformation to three output dimensions.

We also employed a batch size of 16 across 70 epochs of training, with an initial learning rate of 0.001 managed by the Adam optimizer. The learning rate was dynamically adjusted using a ReduceLROnPlateau scheduling mechanism, which reduced the rate by a factor of 0.5 when validation performance plateaued for five consecutive epochs.⁴⁴ Training stability was maintained through gradient clipping with a maximum norm of 1.0,

while model selection was guided by an early stopping mechanism monitoring validation loss.

2.8. Continuous-filter convolutional neural network model

We implemented the SchNet model, a specialized deep-learning architecture designed for molecular systems.⁴⁹ The model consists of an embedding layer followed by multiple interaction blocks and separate prediction heads for energy and forces. The input features are first normalized through a batch normalization layer, then transformed into a 128-dimensional hidden representation through a linear embedding layer with subsequent batch normalization. The model comprises three SchNet interaction blocks containing a filter network and two dense layers. The filter network processes distance-based features through linear transformations and ReLU activations, while the dense layers maintain the hidden dimension of 128 channels. Each interaction block employs layer normalization and the SiLU activation function to introduce non-linearity and ensure stable training.⁵⁰ The model implements two parallel output pathways: one for energy prediction and another for forces. Both pathways share a similar architecture, starting with a 128-dimensional representation that is progressively reduced through multiple layers. Each pathway includes two batch normalization layers, ReLU activations, and a dropout rate of 0.1 for regularization.⁴⁵ The energy pathway concludes with a single output unit, while the forces pathway produces a two-dimensional output.

The training process employs a sophisticated optimization strategy with hierarchical learning rates implemented through the AdamW optimizer. The embedding layer uses a learning rate of 10^{-5} ; interaction blocks operate at 2×10^{-5} ; and both output pathways utilize 5×10^{-5} . A weight decay of 10^{-4} is applied across all parameters. The learning rate schedule combines a warm-up period lasting 20% of the total training steps with cosine decay thereafter. The model is trained for 100 epochs with a batch size of 2048 and employs gradient accumulation over 24 steps to stabilize training. A progressive loss weighting scheme is implemented, where force prediction training begins after 50 epochs and gradually increases in importance. The energy weight starts at 1.0 and decreases to 0.75, while the force weight increases from 0.0 to 0.5 over 50 epochs. Training stability is maintained through gradient clipping with a maximum norm of 0.01.

2.9. Model evaluation

The evaluation of model performance is conducted through a comprehensive loss function framework implemented across training, validation, and test datasets.

Specifically, we employed mean absolute error (MAE) and mean squared error (MSE), defined as:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (X)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (XI)$$

where N is the number of data points, y_i represents the true values (actual labels), and \hat{y}_i represents the measured values, including both energies and forces. This tripartite evaluation strategy serves distinct purposes: the training loss guides the optimization process, the validation loss enables model selection and hyperparameter tuning, and the test loss provides an unbiased estimate of model generalization. This systematic computation of loss metrics across training, validation, and test datasets provides a robust mechanism for detecting potential overfitting phenomena, characterized by divergence between these performance metrics. This comprehensive monitoring strategy ensures that the model develops representations that generalize effectively to unseen data rather than merely memorizing training examples.

3. Results and discussion

3.1. Model evaluation

3.1.1. Model comparison

As presented in Table 1, KAN demonstrated exceptional performance and practical advantages. KAN achieved impressive accuracy with MAE values of 0.36, 0.30, and

Table 1. Performance metrics comparison of KAN against other typical ML models (SchNet, ANI, GNN, and CalHousNet) for MAE and MSE across training, validation, and test sets

Model	Evaluation	Training sets	Validation sets	Test sets
KAN	MAE	0.36	0.30	0.31
	MSE	0.79	0.50	0.49
SchNet	MAE	0.46	0.38	0.41
	MSE	0.88	0.55	0.59
ANI	MAE	0.79	0.77	0.79
	MSE	1.01	0.93	0.97
GNN	MAE	0.15	0.13	0.22
	MSE	0.11	0.04	0.50
CalHousNet	MAE	0.34	0.27	0.29
	MSE	0.76	0.34	0.51

Abbreviations: KAN: Kolmogorov-Arnold Network; ML: Machine learning; MAE: Mean absolute error; MSE: Mean squared error.

0.31 across training, validation, and test sets, respectively, alongside stable MSE metrics (0.79, 0.50, and 0.49). While the GNN demonstrated marginally superior numerical performance (MAE: 0.15, 0.13, and 0.22), it demands substantially more complex data preprocessing and feature engineering, significantly limiting its practical applicability. In contrast, KAN's straightforward implementation and minimal preprocessing requirements make it particularly valuable for real-world applications where computational efficiency and ease of deployment are crucial.

As featured in Figures S3-S7, the training dynamics of KAN display remarkable learning efficiency and stability. During the initial 30 epochs, KAN exhibits rapid learning characterized by a sharp decline in MAE, indicating swift comprehension of the underlying data patterns. The learning curve demonstrates consistent convergence, with both training and validation errors stabilizing after 50 epochs, suggesting robust generalization capabilities. This stability in later epochs, coupled with the close alignment between training and validation performance, indicates that KAN effectively avoids overfitting while maintaining strong predictive power.

Other benchmarked models displayed less competitive performance-to-complexity ratios. SchNet demonstrated moderate but inferior accuracy (MAE: 0.46, 0.38, and 0.41) while requiring specialized graph-based data structures. ANI exhibited the highest error metrics (MAE consistently around 0.79), suggesting limitations in its feature extraction capabilities. CalHousNet's performance (MAE: 0.34, 0.27, and 0.29) approached KAN's accuracy but lacked its efficient training dynamics and straightforward implementation. The consistent performance of KAN across diverse datasets, combined with its minimal preprocessing requirements and stable training behavior, establishes it as the most practical and efficient choice for large-scale molecular property prediction tasks.

In contrast, KAN provides a balanced approach, achieving good performance with less complex data preparation and lower computational demands, as demonstrated in Figure S8. Leveraging the principles of the Kolmogorov-Arnold representation theorem, KAN captures complex non-linear relationships efficiently, without the need for extensive preprocessing or high computational overhead. This efficient utilization of resources combined with robust predictive capabilities grants KAN a significant advantage in computational materials science, making it a preferred choice for researchers seeking to balance performance with practical operational demands. By leveraging KAN's predictive accuracy and using molecular statics, we compute

properties such as elastic constants, thermal expansion coefficients, Poisson's ratio, bulk modulus, shear modulus, and Young's modulus, each essential for understanding the material's mechanical behavior.

To facilitate direct comparison in their natural units, we performed denormalization of the model predictions on the test dataset. Figure 5 presents the distribution of prediction errors in their original units, with energy errors presented in Ry (Figure 5A) and force errors in Ry/au (Figure 5B). After denormalization, the box plots reveal that KAN maintains excellent prediction accuracy with median energy errors around 0.1 Ry, while most predictions fall within a narrow range of 0.05 – 0.2 Ry, demonstrating robust performance. For force predictions, the median error is approximately 0.0003 Ry/au, with most predictions showing errors between 0.0001 – 0.0005 Ry/au. The presence of a few outliers (i.e., points above the whiskers) in both plots indicates occasional challenging cases, but these represent a small fraction of the predictions. This analysis in physical units further validates KAN's strong predictive capabilities, as it maintains high accuracy even when evaluated in the original, unnormalized scale of calculations.

3.1.2. Mechanical properties measurement

Various mechanical properties were measured, as follows:

- (i) Elastic constants calculation: Elastic constants, including C_{11} , C_{12} , C_{13} , and C_{44} , were derived by applying the relevant strains to the crystal structure and measuring the resultant stress tensors with KAN. The strain tensor ϵ is applied to the structure's lattice:

$$R' = R(I + \epsilon) \quad (\text{XII})$$

where R is the original lattice matrix, and I is the identity matrix. The relationship between stress (σ) and strain (ϵ) tensors is defined as:

$$\sigma = C\epsilon \quad (\text{XIII})$$

The stress (σ) and strain (ϵ) data were linearly regressed to determine the material constants (C) from the slope of the stress-strain relationship.

- (ii) Thermal expansion coefficient: The thermal expansion coefficient was determined by simulating temperature-induced lattice expansions. The expansion was modeled by scaling lattice vectors according to a temperature-dependent factor (α), and the resulting volumetric changes were used to compute α by fitting these changes over a range of temperatures. The lattice vectors are scaled based on a temperature-dependent expansion factor (α):

$$R' = R(1 + \alpha\Delta T) \quad (\text{XIV})$$

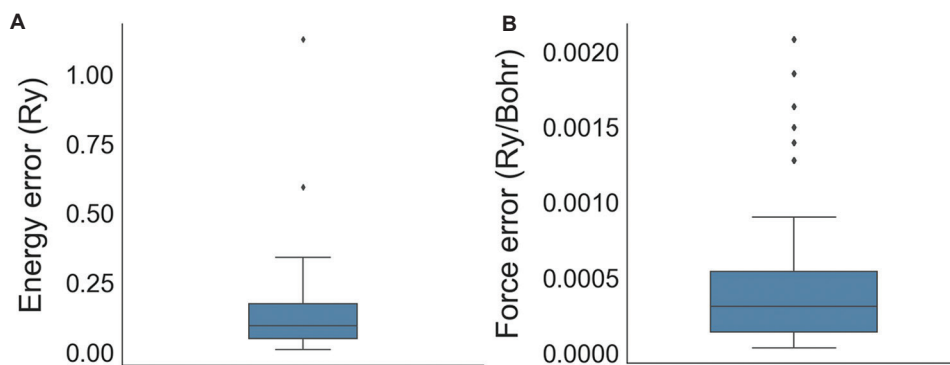


Figure 5. Performance evaluation of KAN through denormalized energy and force predictions on test dataset: box plot distribution of (A) energy errors (Ry) and (B) force errors (Ry/au), demonstrating high predictive accuracy in physical units. Abbreviation: KAN: Kolmogorov-Arnold Network.

The change in volume (V) is computed for different temperatures and the thermal expansion coefficient (α) is determined by fitting the volume change:

$$\alpha = \frac{1}{V_0} \frac{\Delta V}{\Delta T} \quad (\text{XV})$$

where V_0 is the initial volume.

- (ii) Poisson's ratio: This ratio was calculated by applying uniaxial stress and using KAN to predict the resultant strain tensors. Stress (σ) is applied in one direction, and the resulting strain tensor (ϵ) is measured by KAN. Poisson's ratio (ν) is calculated as the negative ratio of transverse strain to longitudinal strain:

$$\nu = \frac{\epsilon_{\perp}}{\epsilon_{\parallel}} \quad (\text{XVI})$$

- (iii) Bulk modulus: By applying hydrostatic pressure and predicting the volumetric changes with KAN, the bulk modulus (K) was calculated from the relationship between pressure and volume change:

$$K = -V_0 \frac{dP}{dV} \quad (\text{XVII})$$

- (iv) Shear modulus: Similarly, the shear modulus (G) was determined by applying shear stress and predicting the corresponding shear strain, with G calculated from the linear relationship between these two measurements.

$$G = \frac{\tau}{\gamma} \quad (\text{XVIII})$$

- (v) Young's modulus: Young's modulus (E) was computed by applying uniaxial stress, predicting the resultant longitudinal strain with KAN, and calculating E from the stress-strain relationship:

$$E = \frac{\sigma}{\epsilon} \quad (\text{XIX})$$

Table 2 highlights the precision of KAN's predictions. Notably, the computed values for properties such as the elastic constants (C_{11} , C_{12} , C_{13} , and C_{44}) demonstrate minimal deviations from established reference data,⁵¹ demonstrating the model's high accuracy. For example, C_{11} was measured at 5.737 GPa against a reference of 5.763 GPa, and C_{44} at 1.956 GPa against 1.84 GPa. In addition, the thermal expansion coefficient, Poisson's ratio, bulk modulus, shear modulus, and Young's modulus all closely align with their respective reference values, with discrepancies within 6%. These results underscore the robustness and reliability of KAN in accurately predicting and understanding material properties, thereby providing a comprehensive tool for materials science research and development.

3.2. MD simulations

To ensure the consistency and reliability of the predictions made by KAN, it is crucial to set the model to evaluation mode during the inference phase. This adjustment is necessary to stabilize the output by deactivating any layers that introduce stochastic behavior, such as Dropout and BatchNorm2d layers. By disabling these layers, we ensure that the model's predictions are stable and reproducible, free from the variability that these layers might otherwise introduce during training. For input data, KAN processes the interatomic distances as a tensor that spans from 1 to 10 Å. This range is intentionally broad to capture a variety of atomic interactions, from extremely close to relatively distant separations, thus providing a detailed dataset for modeling potential energy landscapes. To align with the model's architecture, the distance tensor is expanded to match the required input dimensionality of the KAN.

As displayed in **Figure 6A**, our calculations yield an equilibrium distance r_0 of 2.89 Å for Ru atoms. This result

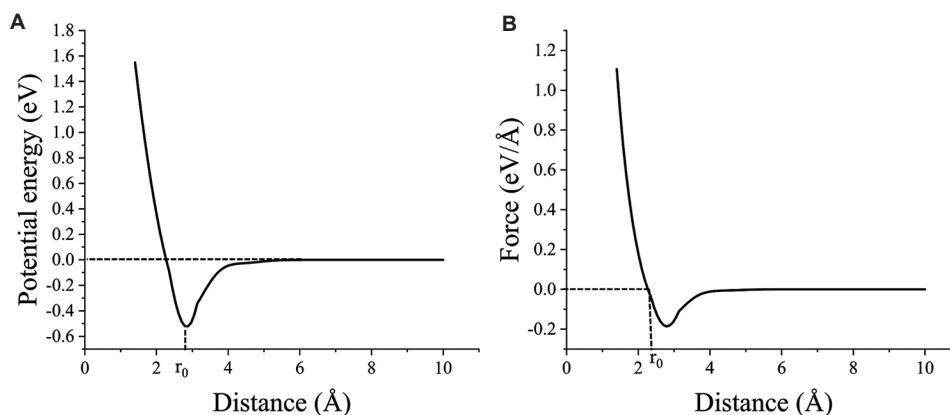


Figure 6. Interatomic potential energy and force curves for Ru atoms. (A) Potential energy curve showing the equilibrium distance, $r_0 \approx 2.89$ Å, at the turning point of the curve. (B) Corresponding force curve derived from the gradient of the potential energy, which crosses zero at r_0 . Both curves demonstrate strong repulsion at short distances (left of r_0) and weaker attraction at longer distances (right of r_0).

Table 2. Comparison of the material properties directly calculated from the KAN model

Property	KAN (%)	Reference ⁵¹
C_{11}	5.726 (0.6)	5.763
C_{12}	1.995 (5.8)	1.872
C_{13}	1.667 (0.3)	1.673
C_{44}	1.966 (6.0)	1.84
Thermal expansion coefficient (10^{-6} K^{-1})	6.44 (0.6)	6.4
Poisson's ratio	0.28 (6.0)	0.3
Bulk modulus (GPa)	217 (1.3)	220
Shear modulus (GPa)	167 (3.4)	173
Young's modulus (GPa)	435 (2.7)	447

Note: The KAN values agree with the reference values (Ref), where parentheses indicate the deviation as a percentage difference from the reference data.⁵¹

Abbreviation: KAN: Kolmogorov-Arnold Network.

closely aligns with the theoretically predicted nearest-neighbor distance of 2.71 Å in the hcp structure of Ru, differing by only about 6.6%.³⁸ This small discrepancy may be attributed to various factors, including the specifics of our computational model, potential stress or defects in real materials, and temperature effects.

The calculation of forces between atoms is another critical component of the model's functionality. Forces are derived by computing the gradient of the potential energy with respect to distance, as defined by the equation:

$$F = -\nabla E \quad (\text{XX})$$

where E represents the potential energy and F represents the forces acting between atoms. This calculation is performed by defining a function that outputs the energy for given distances and using automatic differentiation to

compute the Jacobian matrix of these outputs. The Jacobian matrix provides the necessary gradients, and the net forces are then calculated by summing these derivatives across all relevant distances.

Figure 6B displays the corresponding force curve, which crosses zero at r_0 , confirming the equilibrium position identified in the potential energy plot. To the left of r_0 (at shorter distances), a steep increase in potential energy is observed, corresponding to a strong repulsive force. This reflects the Coulomb repulsion between atomic nuclei at close proximity. To the right of r_0 (at longer distances), the potential energy rises more gradually, with the force becoming negative, indicating a weaker attractive force likely due to electron cloud overlap. These observations validate the consistency of our calculations with known material properties.

This meticulous computation of the Jacobian ensures that the forces derived from the model are accurate reflections of the underlying potential energy predictions. The final step in the workflow involves formatting the computed forces and energies into a LAMMPS potential file. This file, which includes potential energy values for a spectrum of interatomic distances, is compatible with LAMMPS, a widely used MD simulation software. By generating this file, we enable the integration of our model's outputs into broader MD simulations, facilitating detailed and practical analyses of material behaviors under various conditions.

3.2.1. Elastic constant

We used LAMMPS software to calculate the elastic constants using the interatomic potential generated by the KAN model. Elastic constants, such as C_{11} , C_{12} , C_{13} , and C_{44} , are important for describing the mechanical response of

materials to various types of deformation and are integral for analyzing mechanical properties. We optimized the Ru crystal structure to its minimum energy configuration using the potential derived from the KAN model. We then subjected the structure to specific strain conditions – uniaxial, biaxial, and shear – to calculate the stress responses within Ru and the respective elastic constants. For a comparative analysis, we applied identical strain conditions using a well-established potential model known as the Finnis-Sinclair-type interatomic potential (FS potential).⁵² This method is commonly used for simulating metallic systems and serves as a benchmark to validate the predictions made by the KAN interatomic potential.

Our results revealed that the elastic constants C_{11} , C_{12} , C_{13} , and C_{44} measured by our MD simulations with the KAN potential exhibited remarkable alignment with reference experimental values, displaying minimal deviations (Table 3). Our model's predictions were also comparable to those of the FS potential.

3.2.2. Thermodynamic properties

To validate the ability of the interatomic potential generated by KAN to predict thermodynamic properties, we conducted a detailed MD simulation using LAMMPS. The primary objective was to determine the melting point of Ru, leveraging the potentials derived from KAN. The MD simulation was initiated with Ru in its hcp lattice structure with fully periodic boundary conditions. The simulation utilized hybrid/overlay pair styles combining Lennard-Jones (LJ) and table potentials. This combination ensured a smooth transition to zero at larger distances, since the table potential only described atomic interactions in the range of 0.01 – 5 Å, and the LJ potential was incorporated to account for interactions beyond this range. The LJ parameters were carefully chosen to complement the table potential, ensuring continuous and physically meaningful interactions across all relevant distances. This hybrid approach allowed for precise modeling of short-range interactions using the detailed table potential, while also capturing the essential features of longer-range forces through the LJ potential, resulting in a more complete description of the system's behavior.

Ru atoms were assigned their appropriate masses, and the system was configured according to defined lattice parameters. An initial step of energy minimization brought the system to its minimum energy configuration, setting a stable baseline for subsequent dynamic processes. Following stabilization, the system was subjected to a controlled heating process under an NPT ensemble, allowing both temperature and volume to adjust naturally, mirroring real-world thermodynamic conditions. The

temperature was incrementally increased from 300 K to over 4000 K. Throughout this process, we closely monitored the system's thermodynamic properties, such as potential energy and volume.

To obtain accurate relationships between the different phases, we performed common neighbor analysis⁵³ and calculated the radial distribution function (RDF)⁵⁴ on the lattice structure at varying temperatures during the simulation. As depicted in Figure 7, the simulation revealed a phase transition in Ru from its initial hcp structure to a face-centered cubic (fcc) configuration at approximately

Table 3. The elastic constants calculated from the different interatomic potentials in LAMMPS

Elastic constant (Mbar)	Table potential	FS potential ^a	Ref ^b
C_{11}	5.740	5.720	5.763
C_{12}	1.866	1.831	1.872
C_{13}	1.669	1.668	1.673
C_{44}	1.856	1.817	1.84

Note: The values display good agreement with the: ^aFS potential⁵² and ^breference values.⁵¹

Abbreviations: LAMMPS: Large-scale Atomic/Molecular Massively Parallel Simulator; FS: Finnis-Sinclair.

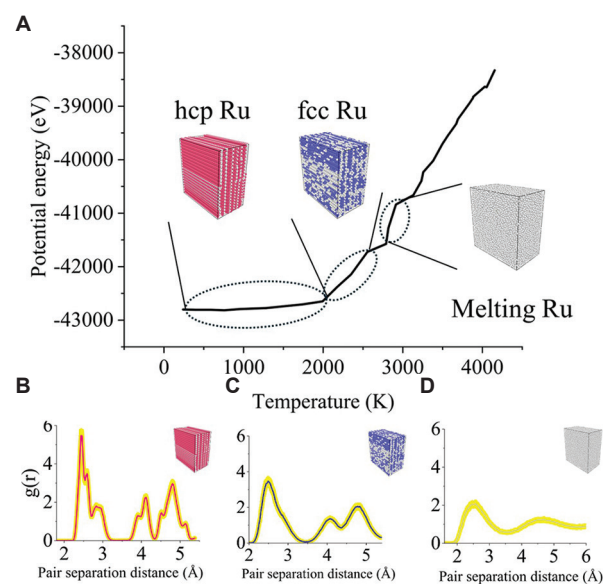


Figure 7. Phase transitions and potential energy evolution of Ru as a function of temperature. (A) Key phase changes (A) from hcp to fcc Ru at ~1882 K, followed by melting at ~2416 K; insets depict the atomic configurations for hcp, fcc, and liquid Ru. (B-D) The time-averaged radial distribution functions (RDF) for the hcp, fcc, and liquid phases, respectively. The well-defined RDF peaks in the (B) hcp and (C) fcc phases indicate short- and medium-range atomic order typical of crystalline structures, while the broadened peaks in (D) the liquid phase signify the loss of long-range order. Shaded regions around the RDF curves represent error bars, capturing statistical variability in atomic positions as the temperature rises. Abbreviations: hcp: Hexagonal close-packed; fcc: Face-centered cubic.

1882 K; the proportion of the hcp phase also decreases from 100% initially to 0%. As displayed in Figure 7B, before 1882 K, Ru exhibits an hcp crystal structure.

The first significant peak appears at about 2.70 Å, representing the distance between the nearest neighboring atoms. The second and third peaks correspond to the distances of the second and third nearest neighbor atoms, respectively, between 4.3 – 4.4 and 5.1 – 5.4 Å. These peaks reflect the short- and medium-range orders of the hcp crystal, and the characteristic peaks are clear and relatively sharp. As observed in Figure 7C, around 1882 K, when the temperature rises, the crystal structure of hcp becomes unstable and gradually transforms, which is manifested by the obvious shift of the peak position of the characteristic peak in RDF. The second and third peaks gradually become wider and begin to shift, which makes the characteristic peaks relatively more symmetrical and regular, reflecting the high symmetry of atomic stacking in the fcc structure. This indicates that the regularity of atomic arrangement has changed during the transformation of hcp to fcc. This phase transition is consistent with the known behavior of Ru under heating and aligns with results from previous studies.⁵⁵

Further heating led to a significant increase in potential energy at around 2416 K, which we identified as the melting point, marking the transition from the fcc solid phase to the liquid phase. As displayed in Figure 7D, the characteristic peaks in the RDF become broader and smoother, reflecting that the long-range order of the crystal structure completely disappears. There is only the first short-range order peak in the RDF, while the subsequent peaks gradually disappear, which is different from the crystal structure of the solid. The melting point observed during the simulation is strikingly close to the experimentally known true melting point of Ru, affirming the potential's accuracy.⁵⁶ The energy continued to rise beyond this temperature, consistent with the behavior of melting materials. Although a slight deviation is observed, this can be attributed to inherent limitations of the simulation model, such as quantum effects,⁵⁷ long-range electrostatic interactions,⁵⁸ and thermal and statistical fluctuations,⁵⁹ rather than a deficiency in KAN's potential.

4. Conclusion

This study demonstrates the efficacy of the KAN model in predicting the mechanical and thermodynamic properties of Ru with high accuracy and computational efficiency. The KAN model not only achieves comparable accuracy to prior literature but also offers significant advantages in terms of reduced computational requirements and descriptor complexity. Compared to traditional neural

networks and GNN, our KAN-based model demonstrated superior performance in terms of learning speed, accuracy, and computational efficiency. While GNNs reported comparable accuracy, they required more complex data preparation and higher computational resources. KAN's ability to capture complex non-linear relationships efficiently, without extensive preprocessing or high computational overhead, makes it particularly suitable for computational materials science applications, particularly in the exploration of RSAs. Our model's predictions of elastic constants, thermal expansion coefficient, Poisson's ratio, bulk modulus, shear modulus, and Young's modulus displayed excellent agreement with experimental reference data, with discrepancies within 6%. This high level of accuracy validates the reliability of our KAN-based approach. Furthermore, the integration of KAN models with MD simulations accurately captured Ru's phase transitions, including the transition from hcp to fcc structure and melting point.

However, it is important to acknowledge the limitations of our approach. The slight deviation observed in the melting point prediction, while small, highlights the inherent challenges in accurately modeling complex material behaviors. This deviation may be attributed to several factors, including quantum effects not accounted for in classical MD simulations, incomplete representation of long-range electrostatic interactions, and thermal and statistical fluctuations inherent in MD simulations. In addition, while our model performs well for Ru, its applicability to a wider range of elements and compounds warrants further investigation. The model's accuracy may vary for materials with significantly different electronic structures or bonding characteristics. Although more efficient than some alternatives, the KAN approach still requires substantial computational resources for training and large-scale simulations, which may limit its applicability in certain research settings.

Despite these limitations, the success of KAN in characterizing Ru properties has significant implications for the broader field of computational materials science. By providing a method that balances high-fidelity predictions with practical computational demands, this approach could accelerate the discovery and design of novel RSAs and other advanced materials. KAN's ability to generate accurate interatomic potentials for MD simulations further enhances its utility in predicting complex material behaviors under various conditions. Looking ahead, the KAN framework demonstrates promise for extension to other elements and more complex alloy systems. For transferring this model to other materials, the approach differs depending on the system's complexity. For single

elements, we can leverage the existing successful KAN structure that was developed for Ru. This transfer simply requires inputting different DFT-calculated datasets for the new target element while maintaining the same architectural framework. However, when considering binary or higher-order materials systems, the KAN structure needs to be appropriately adjusted to accommodate more complex crystal structures and atomic interactions. This may involve modifying the input layer and potentially adjusting the network's depth or width to capture the additional complexity of multi-element systems. In addition, integrating KAN with high-throughput screening methods could further expedite the materials discovery process. Efforts to incorporate quantum mechanical effects and improve the handling of long-range interactions could enhance the model's accuracy, especially for extreme conditions and diverse material types. This research not only advances our understanding of Ru but also provides a valuable tool for the broader materials science community. The successful implementation of KAN in predicting Ru properties represents a significant advancement in materials informatics, establishing a more efficient and accurate pathway for material property prediction that could revolutionize the discovery and design of complex systems like HEAs and refractory superalloys while setting a new standard for ML applications in materials research.

Acknowledgments

This work used SDSC Expanse at the San Diego Supercomputer Center through allocation BIO240093 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support program supported by National Science Foundation (NSF) grants #2138259, #2138286, #2138307, #2137603, and #2138296.

Funding

J.Y. acknowledges support from the US NSF under the award CMMI-2338518.

Conflict of interest

Jingjie Yeo is an Editorial Board Member of this journal but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: All authors

Data curation: Zhiyu An

Formal analysis: Zhiyu An
Funding acquisition: Jingjie Yeo
Investigation: Zhiyu An
Methodology: Zhiyu An
Project administration: Jingjie Yeo
Software: Zhiyu An
Resources: Jingjie Yeo
Supervision: Jingjie Yeo
Validation: Jingjie Yeo
Visualization: Zhiyu An
Writing – original draft: Zhiyu An
Writing – review & editing: Jingjie Yeo

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The code and dataset for our implementation of the KAN model can be found at <https://github.com/ZhiyuAn-byte/KAN-Model-for-Ru-element> (doi: 10.5281/zenodo.13856262)

References

1. Naka S, Khan T. Designing novel multiconstituent intermetallics: Contribution of modern alloy theory in developing engineered materials. *J Phase Equilibria*. 1997;18(6):635-649.
doi: 10.1007/BF02665823.
2. Yeh JW, Chen SK, Lin SJ, *et al.* Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes. *Adv Eng Mater*. 2004;6(5):299-303.
doi: 10.1002/adem.200300567
3. Cantor B, Chang ITH, Knight P, Vincent AJB. Microstructural development in equiatomic multicomponent alloys. *Mater Sci Eng A*. 2004;375-377:213-218.
doi: 10.1016/j.msea.2003.10.257
4. Miracle DB, Senkov ON. A critical review of high entropy alloys and related concepts. *Acta Mater*. 2017;122:448-511.
doi: 10.1016/j.actamat.2016.08.081
5. Soni V, Gwalani B, Senkov O, *et al.* Phase stability as a function of temperature in a refractory high-entropy alloy. *J Mater Res*. 2018;33:1-12.
doi: 10.1557/jmr.2018.223
6. Senkov ON, Crossman B, Rao SI, *et al.* Mechanical properties of an Al10Nb20Ta15Ti30V5Zr20 A2/B2 refractory superalloy and its constituent phases. *Acta Mater*.

- 2023;254:119017.
doi: 10.1016/j.actamat.2023.119017
7. Miracle DB, Tsai MH, Senkov ON, Soni V, Banerjee R. Refractory high entropy superalloys (RSAs). *Scr Mater.* 2020;187:445-452.
doi: 10.1016/j.scriptamat.2020.06.048
 8. Prins SN, Cornish LA, Stumpf WE, Sundman B. Thermodynamic assessment of the alru system. *Calphad.* 2003;27(1):79-90.
doi: 10.1016/S0364-5916(03)00033-6
 9. Haynes WM, editor. *CRC Handbook of Chemistry and Physics.* 95th ed. Boca Raton, FL: CRC Press/Taylor and Francis; 2015.
 10. Hulm JK, Goodman BB. Superconducting properties of rhenium, ruthenium, and osmium. *Phys Rev.* 1957;106(4):659-671.
doi: 10.1103/PhysRev.106.659
 11. Anzellini S, Errandonea D, Cazorla C, *et al.* Thermal equation of state of ruthenium characterized by resistively heated diamond anvil cell. *Sci Rep.* 2019;9(1):14459.
doi: 10.1038/s41598-019-51037-8
 12. Güler E, Uğur Ş, Güler M, Uğur G. Molecular dynamics exploration of the temperature-dependent elastic, mechanical, and anisotropic properties of Hcp ruthenium. *Eur Phys J Plus.* 2024;139(5):372.
doi: 10.1140/epjp/s13360-024-05177-0
 13. Zhi-Peng LU, Wen-Jun Z, Tie-Cheng L, Chuan-Min M, Liang X, Xu-Hai L. Structural phase transition of ru at high pressure and temperature. *Acta Phys Sin.* 2013;62(17):176402-176402.
doi: 10.7498/aps.62.176402
 14. Jain A, Ong SP, Hautier G, *et al.* Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater.* 2013;1(1):011002.
doi: 10.1063/1.4812323
 15. Mobarak MH, Mimona MA, Islam MA, *et al.* Scope of machine learning in materials research-a review. *Appl Surf Sci Adv.* 2023;18:100523.
doi: 10.1016/j.apsadv.2023.100523
 16. Mannodi-Kanakthodi A, Pilania G, Huan TD, Lookman T, Ramprasad R. Machine learning strategy for accelerated design of polymer dielectrics. *Sci Rep.* 2016;6:20952.
doi: 10.1038/srep20952
 17. Stanev V, Oses C, Kusne AG, *et al.* Machine learning modeling of superconducting critical temperature. *NPJ Comput Mater.* 2018;4(1):29.
doi: 10.1038/s41524-018-0085-8
 18. Faber FA, Lindmaa A, Von Lilienfeld O.A, Armiento, R. Machine learning energies of 2 million elpasolite $(AB_2C_4D_6)$ crystals. *Phys Rev Lett.* 2016;117(13):135502.
doi: 10.1103/PhysRevLett.117.135502
 19. Pilania G, Balachandran PV, Kim C, Lookman T. Finding new perovskite halides via machine learning. *Front Mater.* 2016;3:19.
doi: 10.3389/fmats.2016.00019
 20. Malkiel I, Mrejen M, Nagler A, Arieli U, Wolf L, Suchowski H. Plasmonic nanostructure design and characterization via deep learning. *Light Sci Appl.* 2018;7(1):60.
doi: 10.1038/s41377-018-0060-7
 21. De Jong M, Chen W, Notestine R, *et al.* A statistical learning framework for materials science: Application to elastic moduli of k-nary inorganic polycrystalline compounds. *Sci Rep.* 2016;6(1):34256.
doi: 10.1038/srep34256
 22. Song K, Zhao R, Liu J, *et al.* General-purpose machine-learned potential for 16 elemental metals and their alloys. *Nat Commun.* 2024;15(1):10208.
doi: 10.1038/s41467-024-54554-x
 23. Liyanage M, Turlo V, Curtin WA. Machine learning potential for the Cu-W system. *Phys Rev Mater.* 2024;8(11):113804.
doi: 10.1103/physrevmaterials.8.113804
 24. Liu Y, Zhao T, Ju W, Shi S. Materials discovery and design using machine learning. *High Throughput Exp Model Res Adv Batter.* 2017;3(3):159-177.
doi: 10.1016/j.jmat.2017.08.002
 25. Anstine DM, Isayev O. Machine learning interatomic potentials and long-range physics. *J Phys Chem A.* 2023;127(11):2417-2431.
doi: 10.1021/acs.jpca.2c06778
 26. Kolmogorov A. On the representation of continuous functions of several variables as superpositions of continuous functions of a smaller number of variables. *Dokl Akad Nauk.* 1956;108(2):25-46.
 27. Liu Z, Wang Y, Vaidya S, *et al.* KAN: Kolmogorov-Arnold Networks; 2024. Available from: <https://arxiv.org/abs/2404.19756> [Last accessed on 2025 Jan 26].
 28. Anderson PW. Absence of diffusion in certain random lattices. *Phys Rev.* 1958;109(5):1492-1505.
doi: 10.1103/PhysRev.109.1492
 29. Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. *Science.* 1972;175(4023):720-731.
doi: 10.1126/science.175.4023.720
 30. Li Y, Zhang JH, Mei F, Ma J, Xiao L, Jia S. Generalized Aubry-André-Harper models in optical superlattices. *Chin*

- Phys Lett.* 2022;39(6):063701.
doi: 10.1088/0256-307x/39/6/063701
31. Giannozzi P, Andreussi O, Brumme T, *et al.* Advanced capabilities for materials modelling with quantum ESPRESSO. *J Phys Condens Matter.* 2017;29(46):465901.
doi: 10.1088/1361-648X/aa8f79
 32. Giannozzi P, Baroni S, Bonini N, *et al.* QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J Phys Condens Matter.* 2009;21(39):395502.
doi: 10.1088/0953-8984/21/39/395502
 33. Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett.* 1996;77(18):3865-3868.
doi: 10.1103/PhysRevLett.77.3865
 34. Blöchl PE. Projector augmented-wave method. *Phys Rev B.* 1994;50(24):17953-17979.
doi: 10.1103/PhysRevB.50.17953
 35. Ong SP, Richards WD, Jain A, *et al.* Python materials genomics (Pymatgen): A robust, open-source python library for materials analysis. *Comput Mater Sci.* 2013;68:314-319.
doi: 10.1016/j.commatsci.2012.10.028
 36. Chen X, Wang LF, Gao XY, *et al.* Machine learning enhanced empirical potentials for metals and alloys. *Comput Phys Commun.* 2021;269:108132.
doi: 10.1016/j.cpc.2021.108132
 37. Thompson AP, Aktulga HM, Berger R, *et al.* LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput Phys Commun.* 2022;271:108171.
doi: 10.1016/j.cpc.2021.108171
 38. Villars P, Cenzual K. *Ru Hcp (Ru) Crystal Structure: Datasheet from "PAULING FILE Multinaries Edition.* Springer Materials; 2022. Available from: https://materials.springer.com/isp/crystallographic/docs/sd_1244358 [Last accessed on 2025 Jan 26].
 39. De Jong M, Chen W, Angsten T, *et al.* Charting the complete elastic properties of inorganic crystalline compounds. *Sci Data.* 2015;2(1):150009.
doi: 10.1038/sdata.2015.9
 40. Himanen L, Jäger MOJ, Morooka EV, *et al.* DScribe: Library of descriptors for machine learning in materials science. *Comput Phys Commun.* 2020;247:106949.
doi: 10.1016/j.cpc.2019.106949
 41. D'Agostino D, Serani A, Campana EF, Diez M. Nonlinear methods for design-space dimensionality reduction in shape optimization. In: Nicosia G, Pardalos P, Giuffrida G, Umeton R, editors. *Machine Learning, Optimization, and Big Data.* Cham: Springer International Publishing; 2018. p. 121-132.
 42. Kingma DP, Jimmy B. Adam: A Method for Stochastic Optimization. In: *3rd International Conference for Learning Representations, CoRR;* 2015
 43. Cortes C, Mohri M, Rostamizadeh A. L2 Regularization for Learning Kernels. In: *Twenty-Fifth Conference on Uncertainty in Artificial Intelligence;* 2004. p. 109-16.
 44. Paszke A, Gross S, Massa F, *et al.* Pytorch: An imperative style, high-performance deep learning library. *Adv Neural Inf Process Syst.* 2019;32:8024-8035.
 45. Agarap AF. *Deep Learning Using Rectified Linear Units (Relu).* [Preprint]; 2018.
 46. Smith JS, Isayev O, Roitberg AE. ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chem Sci.* 2017;8(4):3192-3203.
doi: 10.1039/C6SC05720A
 47. Najm HN, Yang Y. *AEVmod-Atomic Environment Vector Module Documentation;* United States. California: Sandia National Laboratories; 2021.
doi: 10.2172/1817835.
 48. Zhang G, Haopeng, L. *Effectiveness of Scaled Exponentially-Regularized Linear Units (SERLUs)* [Preprint]; 2018.
 49. Schütt KT, Sauced HE, Kindermans PJ, Tkatchenko A, Müller KR. SchNet—a deep learning architecture for molecules and materials. *J Chem Phys.* 2018;148(24):241722.
doi: 10.1063/1.5019779
 50. Elfving S, Uchibe E, Doya K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural Netw.* 2018;107:3-11.
doi: 10.1016/j.neunet.2017.12.012
 51. Fast L, Wills JM, Johansson B, Eriksson O. Elastic constants of hexagonal transition metals: Theory. *Phys Rev B.* 1995;51(24):17431-17438.
doi: 10.1103/PhysRevB.51.17431
 52. Fortini A, Mendeleev MI, Buldyrev S, Srolovitz D. Asperity contacts at the nanoscale: Comparison of Ru and Au. *J Appl Phys.* 2008;104(7):074320.
doi: 10.1063/1.2991301
 53. Honeycutt JD, Andersen HC. Molecular dynamics study of melting and freezing of small lennard-jones clusters. *J Phys Chem.* 1987;91(19):4950-4963.
doi: 10.1021/j100303a014
 54. Abbaspour M, Jorabchi MN, Akbarzadeh H, Ebrahimnejad A. Investigation of the thermal properties of phase change materials encapsulated in capped carbon nanotubes

- using molecular dynamics simulations. *RSC Adv.* 2021;11(40):24594-24606.
doi: 10.1039/D1RA02033A
55. Lu ZP, Zhu WJ, Lu TC, *et al.* Structural phase transition of Ru at high pressure and temperature. *Wuli Xuebao Acta Phys Sin.* 2013;62(17):176402.
doi: 10.7498/aps.62.176402
56. Savitskii EM. *Noble Metals: A Handbook*; Izdatel'stvo Metallurgii. Moscow: USSR; 1984.
57. Dammak H, Hayoun M, Briec F, Geneste G. Nuclear quantum effects in molecular dynamics simulations. *J Phys Conf Ser.* 2018;1136:012014.
58. Piana S, Lindorff-Larsen K, Dirks RM, Salmon JK, Dror RO, Shaw DE. Evaluating the effects of cutoffs and treatment of long-range electrostatics in protein folding simulations. *PLoS One.* 2012;7(6):e39918.
doi: 10.1371/journal.pone.0039918
59. Hoyt JJ, Trautt ZT, Upmanyu M. Fluctuations in molecular dynamics simulations. *Multiscale Model Mov Interfaces Mater.* 2010;80(7):1382-1392.
doi: 10.1016/j.matcom.2009.03.012