Artificial Intelligence in Molecular Physics

[Vol-01, Issue-02, July 2026]
DOI: XX
ISSN: XXXX-XXXX

Machine Learning Interatomic Potentials for Molecular Simulations: Methods, Applications, Accuracy and Transferability

Jürgen Bajorath

Department of Life Science Informatics and Data Science, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, University of Bonn, Friedrich-Hirzebruch-Allee 5/6, D-53115 Bonn, Germany.

Received: 06/06/2026 **Revised:** 12/07/2026
Accepted: 20/07/2026 **Published:** 27/07/2026

Abstract—*The combination of the accuracy of quantum mechanical methods with the speed of classical force fields has made machine learning interatomic potentials (MLIPs) a transformative method in molecular and materials simulations. This review gives a complete picture of the basic principles, methodological advancements, applications, accuracy and transferability of the MLIPs. It first considers the theoretical background of potential energy surfaces, atomic interactions and necessary physical restrictions for a reliable interatomic modelling. The full development workflow is then summarized, from the generation of reference data to the generation of atomic descriptors, the training and validation of the model, hyperparameter optimization and the formulation of active learning strategies. The predictive power and computational complexity of major MLIP approaches, such as neural network potentials, Gaussian approximation potentials, moment tensor potentials, atomic cluster expansion models, graph neural networks, and equivariant potentials, are critically assessed. In addition, it considers approaches for the numerical evaluation of energies, forces, stresses, and structural property evaluations, and elucidates the need for careful validation, uncertainty quantification, and computational scalability. Transferability of MLIPs to various chemical compositions, phases, temperatures, defects, interfaces, and chemically reacting systems is also discussed, as is recent progress in foundation models and pretrained atomistic learning frameworks. Lastly, some challenges in the current state, such as limited reference datasets, long-range interactions, model interpretability, and benchmarking, are listed, and future research opportunities are discussed.*

Keywords— ***Machine learning interatomic potentials; Molecular simulations; Potential energy surfaces; Graph neural networks; Atomistic modelling***

1. INTRODUCTION

Machine learning interatomic potentials are now a tool for going beyond the practical limits of conventional quantum-mechanical methods for extending atomistic and molecular simulations. Molecular simulation is the description of the movement, interaction, reorganization and energy exchange of atoms under various thermodynamic and chemical conditions. It is reliable only if based on the potential energy surface used to calculate the energies and forces of atoms. Physically grounded approaches like density functional theory (DFT) give reliable predictions but are, at the same time, very costly in terms of computation for large systems and long simulation times. Classical force fields have the advantage of being faster, but the fixed form and limited parameterization can limit accuracy in a wide variety of bonding contexts (Noé et al., 2020).

The emergence of machine learning interatomic potentials marks a shift from analytically designed interatomic potentials to data-driven, flexible representations of atomic interactions. These models are developed using accurate electronic-structure calculations for reference configurations, energies, forces, and, in some cases, stresses. After training, they can test and validate new configurations in a much shorter time, and support simulations that take thousands or millions of atoms over longer time periods. They have the benefit of being able to learn complex, nonlinear relationships without needing a pre-specified form for each interaction. This ability has broadened the range of molecular dynamics, materials modelling, reaction analysis, phase-transition and defect evolution studies. Their increasing maturity has made them an essential tool for next-generation matter simulations (Friederich et al., 2021).

There have been numerous machine learning architectures developed to build interatomic potentials. The total energy in high-dimensional neural network potentials is expressed as the sum of atomic contributions, and similarity functions in kernel-based potentials compare the structural environments. Recent graph-based and equivariant models capture geometric relationships and rotational symmetries, enabling the learning of energies and forces to be more data-efficient. Important families are atomic cluster expansion, moment tensor potentials, Gaussian approximation potentials, and deep potential models. Although they differ, these methods should meet physical criteria: they should be translationally, rotationally and permutationally invariant when the atoms in the molecule are identical (Unke et al., 2021).

One of the most important criteria to evaluate machine learning potentials for simulation is their accuracy. It is important that a low number of errors are made on energies and forces, but not too many. A model can have good predictive ability when applied to randomly generated test data, but have poorly stable trajectories, wrong phase behavior, unrealistic defect energies, or wrong transport properties. There should thus be a statistical measure and a physical measure to validate. Assessments on lattice constants, elastic properties, phonon spectra, radial distribution functions, diffusion coefficients, reaction barriers, surface energies, thermal expansion and melting behavior may be relevant. Computational efficiency is also crucial as it is if many simulations can't be performed because of evaluation costs, then accurate models are of little value. The development of MLIP thus involves a delicate equilibrium between the model complexity, the speed of prediction, the numerical stability and the accuracy of the target (Mishin, 2021).

Transferability is more difficult in the context of atomistic simulations as the simulations may explore configurations that are not represented in the training set. The potential may not develop for surfaces, defects, liquids, high-pressure phases, fracture processes or for reactive surroundings. Such models may not be generalizable to alloys, doped systems, interfaces, and new coordination environments. This is due to the nature of machine learning potentials: they are fundamentally interpolative in the space of configurations and chemicals they were trained on. To ensure reliable deployment, learning is necessary, uncertainty estimation is required, out-of-distribution detection is needed, and reference sets need to be expanded. The transferability is also linked to long-range electrostatics, charge transfer, dispersion, and magnetic interactions, and short-range local descriptors might not capture these interactions effectively. A focus on scalable, widely applicable potentials that integrate atomistic description with device-scale simulation has gradually emerged in the research field (Wang et al., 2024).

The applications of the MLIPs are now expanded to various fields, including chemistry, physics, materials science, nanotechnology, studies on energy resources, and molecular engineering. They are useful for probing crystalline and amorphous systems, catalytic reactions, interfaces in batteries, ionic conduction, deformation in mechanical properties, thermal conductivity, phase transitions and biomolecular conformations. Their fast speed and very high accuracy, near quantum, is particularly useful in situations where rare events, complex structural evolution or extensive configurational sampling are to be considered.

This study focuses on the main strategies developed to build machine learning interatomic potentials and their use in molecular and materials simulations. It evaluates the predictive accuracy, validation strategies, computational efficiency, transferability and generalization under the chemical and thermodynamic conditions. The review also highlights the methodological challenges and directions toward reliable, scalable and universally applicable atomistic models for the future.

2. FUNDAMENTALS OF MACHINE LEARNING INTERATOMIC POTENTIALS

2.1 Potential Energy Surfaces and Atomic Interactions

The potential energy surface (PES) is an important and critical model used in molecular simulations that is a mapping of atomic configurations to energies that dictates forces, stability, reactivity and thermodynamics. Classical methods use empirical forms to approximate the PES, and the quantum methods use electronic structure calculations to derive the PES. Instead, machine learning interatomic potentials (MLIPs) learn the PES from first-principles data, based on flexible nonlinear mapping between local atomic environments and energies, at lower cost, with near-quantum accuracy (Behler, 2021; Kocer et al., 2022). The typical approach to the calculation of the total energy in an MLIP is to add up the energy contributions to the atom in the local environment. In an MLIP, the total energy is usually calculated as a sum of the energies of the atoms in the local environment. Advanced descriptors are used for the encoding of geometric and chemical environment, which increases the prediction accuracy in various systems (Musil et al., 2021).

2.2 Core Principles of ML-Based Interatomic Modelling

Supervised learning is employed in MLIPs to relate atomic structures to energies, forces and stresses. The quality of the training data, the effectiveness of the descriptors and the learning algorithms will affect model performance. A descriptor encodes a local environment into numbers for a regression or deep learning model for efficient prediction on large-scale simulations (Behler, 2021). Popular architectures are neural network potentials, Gaussian approximation potentials, moment tensor potentials, atomic cluster expansion and graph neural networks. While the differences are present, all strive to mimic quantum interactions in a scalable manner in a way that is accurate, efficient and robust (Kocer et al., 2022).

2.3 Essential Properties: Symmetry, Invariance, Locality, and Differentiability

Physical requirements for MLIPs: Energy is invariant under permutation of identical atoms, translation, and rotation, so that it is consistent and stable. These symmetries are preserved by the modern descriptors and graph-based methods (Musil et al., 2021). Locality: assume that the energy of an atom is primarily determined by neighboring atoms in a cutoff distance, so that scalability is achieved. Forces are calculated as energy gradients for stable molecular dynamics, the function must be differentiable. Rigorous, efficient and systematically improvable representations that satisfy these properties exist as frameworks such as atomic cluster expansion (Drautz, 2020).

2.4 Comparison with Classical and Quantum-Mechanical Approaches

Classical force fields are simple but are not transferable because they have fixed functional forms. Quantum methods are highly accurate but costly about big systems. It is this gap that is filled by MLIPs, which are near-quantum accurate and efficient. They are data-driven and do not assume a form of many-body interactions and are therefore suitable for studying the structural evolution, reactions, and structural properties of materials at scale which cannot be studied by first-principles methods. They have become more important for the modelling of atoms today (Deringer et al., 2019; Behler, 2021).

3. REPRESENTATIONS AND DEVELOPMENT WORKFLOW

3.1 Generation and Selection of Reference Configurations

The first step in the development of a reliable machine learning interatomic potential is to obtain representative atomic configurations from first-principles calculations. These should include distorted structures, defects, phase transitions, under- and overcooled states, and various thermodynamic conditions to properly sample the configurational space of interest. Balanced and diverse reference datasets boost model robustness and accuracy (Podryabinkin et al., 2019).

3.2 Atomic Descriptors and Structural Representations

To use machine learning algorithms, atomic descriptors are used to represent the local chemical environment numerically. Descriptors should be distinct to the atomic neighborhood, but be rotation, translation and permutation invariant. High-quality structural representations lead to good model generalization and can make accurate predictions for a variety of molecular and material systems (Qi et al., 2024).

3.3 Model Training, Validation, and Hyperparameter Optimisation

The training of MLIPs by supervised learning algorithms is performed with reference energies, forces, and stress tensors. Optimization of the model includes choosing the right architecture, adjusting hyperparameters, and testing it on test data sets to ensure that no overfitting occurs and that the accuracy of prediction is maximized. The stable and transferrable models are achieved by the continuous evaluation based on several validation metrics (Novikov et al., 2021).

3.4 Active Learning and Adaptive Sampling Strategies

In active learning, the training set is continually updated by adopting the configuration of the problem for which the uncertainty in the prediction is the greatest, achieving a more economical solution in terms of the number of reference calculations needed. Finally, adaptive sampling can be used to provide even more efficient data and generate more powerful and transferable interatomic potentials for large-scale molecular simulation (Vandermause et al., 2020; Yang et al., 2021).

4. MAJOR METHODS OF MACHINE LEARNING INTERATOMIC POTENTIALS

4.1 Neural Network and Deep Learning Potentials

One of the first MLIPs that could accurately reproduce potential energy surfaces from a quantum-mechanical calculation was neural network potentials (NNPs). They break down energy into atomic contributions built from the local structural environments, allowing highly accurate and efficient computations. More recently, the scalability, robustness, and prediction accuracy for complex molecular and materials systems have been further enhanced by recent advances in deep learning architectures (Behler, 2021; Tokita & Behler, 2023).

4.2 Kernel-Based and Gaussian Approximation Potentials

Kernel-based methods, such as Gaussian Approximation Potentials (GAPs), estimate the energies of atoms based on the similarity of the atomic environment they are in, using kernel functions. These methods are capable of very high accuracy for relatively small training sets and have been used to successfully characterize a wide variety of crystalline and amorphous materials. Descriptor quality and effective kernel construction (Mishin, 2021; Drautz, 2020) play a significant role in their performance.

4.3 Moment Tensor, Spectral Neighbour, and Atomic Cluster Expansion Potentials

Moment Tensor Potentials (MTPs), Spectral Neighbour Analysis Potentials (SNAPs) and Atomic Cluster Expansion (ACE) models use mathematically rigorous descriptors that are systematically accurate in capturing many-body atomic interactions (Novikov et al., 2021). Such techniques are computationally efficient, are more transferable and are scalable to large scale atomistic simulations. Active learning strategies additionally bolster their anticipating ability by continually upgrading the training set (Dusson et al., 2022; Lysogorskiy et al., 2021).

4.4 Graph Neural Network and Equivariant Interatomic Potentials

Graph Neural Networks (GNNs) represent atoms as nodes in a graph and interactions as connections between these nodes, enabling learning of complex relationships in a graph-based manner. Directional message-passing networks and equivariant architectures maintain a rotational symmetry and yield a better prediction of energies, forces, and tensorial properties. Recently, MLIPs such as MACE, NequIP, M3GNet, CHGNet and the related equivariant models have made remarkable progress in terms of data efficiency, transferability, and large-scale molecular simulations (Gasteiger et al., 2020; Gasteiger et al., 2021). Table 1. Comparisons of Interatomic Potential Major Machine Learning Methods.

Table 1. Comparison of Major Machine Learning Interatomic Potential Methods

Method	Principle	Major Advantages	Main Limitations	Typical Applications	References
Neural Network Potentials	Deep neural networks learn atomic energies	High accuracy	Large training datasets	Molecular dynamics, reactions	Behler (2021)
Gaussian Approximation Potentials	Kernel regression	Excellent accuracy	High computational cost	Crystalline materials	Mishin (2021)
Moment Tensor Potentials	Tensor-based descriptors	Fast and transferable	Moderate extrapolation	Metals and alloys	Novikov et al. (2021)
Atomic Cluster Expansion	Polynomial basis expansion	Systematic improvement	Descriptor complexity	Multicomponent systems	Dusson et al. (2022)
Graph Neural Networks	Graph message passing	Excellent transferability	Expensive training	Materials discovery	Chen & Ong (2022)
Equivariant Neural Networks	Rotation-equivariant learning	High physical consistency	Computationally demanding	Universal MLIPs	Batatia et al. (2022)

5. ACCURACY, VALIDATION, AND COMPUTATIONAL PERFORMANCE

5.1 Evaluation of Energies, Forces, Stresses, and Atomic Properties

Prediction accuracy of machine learning interatomic potentials (MLIPs) is evaluated in terms of closely matching reference energies, atomic forces, atomic properties such as stress tensors, and other properties produced by first-principles simulation (Deng et al., 2023). The reliable prediction of these quantities is crucial for a reliable molecular dynamic's simulation, as they directly affect structural evolution, the stability of phases and the behavior of materials under various thermodynamic conditions. The potential developed is validated in extensive range of physical properties, which ensures that the developed potential is both numerically accurate and physically meaningful (Sivaraman et al., 2021; Schütt et al., 2021).

5.2 Performance Metrics and Comparison with Reference Calculations

For the models, the most popular metrics for model performance are the root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). These are the metrics used to quantify the agreement between predicted and reference values as well as to detect systematic prediction errors. Until today, the use of comparisons with density functional theory (DFT) calculations is the benchmark model to evaluate the reliability and generalization ability of the MLIPs for different atomic configurations and material systems (Ghaffari et al., 2024). Figure 1 illustrates different strategies for integrating density functional theory (DFT) and experimental (EXP) data to develop and validate accurate machine learning interatomic potentials.

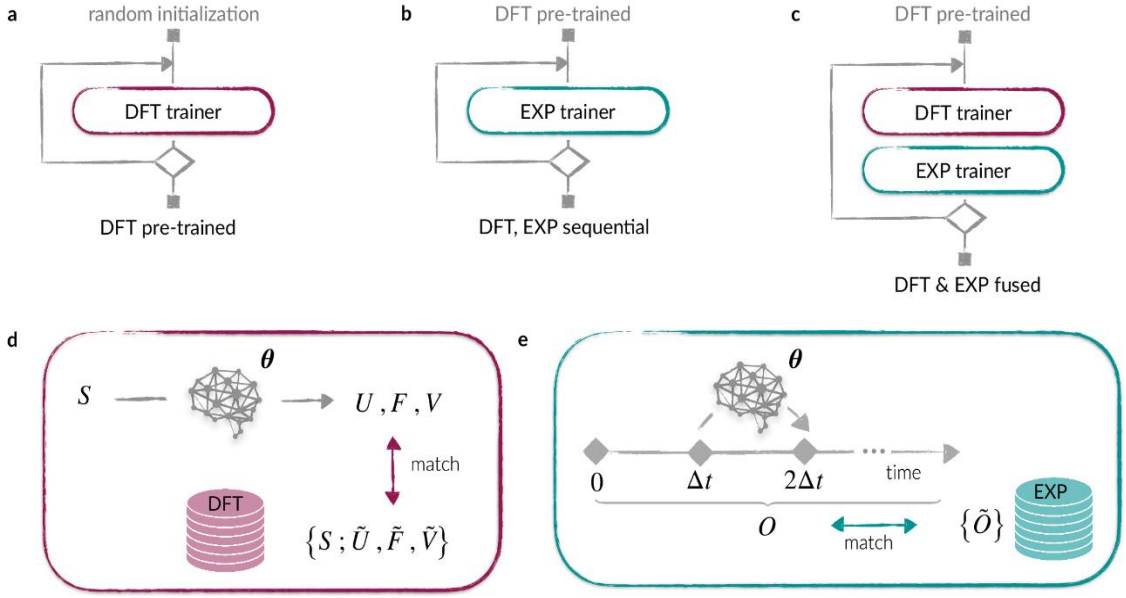

Figure 1. DFT pre-training and experimental data fusion strategies for accurate machine learning interatomic potentials (Röcken & Zavadlav, 2024)

5.3 Structural, Thermodynamic, Mechanical, and Dynamical Validation

In addition to statistical accuracy, the MLIPs need to be able to reproduce experimentally relevant structural and thermodynamic properties. Commonly included in validation are lattice parameters, elastic constants, energy levels for defect formation, radial distribution functions, diffusion coefficients, phonon spectra, and phase transition properties. The assessment of several physical observables guarantees that the learned potential fully captures both the equilibrium and non-equilibrium atomic processes under various simulation conditions (Poul et al., 2023).

5.4 Accuracy-Efficiency Trade-Offs and Computational Scalability

One of the biggest benefits of MLIPs is that they offer a near quantum-mechanical accuracy with computational efficiency of larger scale simulation. The complexity of the model is usually increased to get better predictions, but it can cost more to run the model. Inference time has been dramatically shortened to the point where simulations can be performed on millions of atoms, thanks to efficient parallel algorithms, optimized model architectures and scalable implementations. A key challenge in the future development of next-generation interatomic potentials (IAPs) is therefore to balance predictive accuracy, computational efficiency and scalability (Park et al., 2024; Wang et al., 2024). The most widely used metrics for validating machine learning interatomic potentials for predicting accuracy are summarized in Table 2.

Table 2. Common Validation Metrics for MLIPs

Metric	Purpose	Desired Value	Importance	References
RMSE	Energy prediction error	Lower	Overall accuracy	Ghaffari et al. (2024)
MAE	Average prediction error	Lower	Model robustness	Ghaffari et al. (2024)
Force Error	Accuracy of predicted forces	Lower	Molecular dynamics	Sivaraman et al. (2021)
Stress Error	Mechanical response	Lower	Solid mechanics	Poul et al. (2023)
Computational Time	Simulation efficiency	Lower	Large-scale simulations	Park et al. (2024)
Scalability	Performance with system size	Higher	High-performance computing	Wang et al. (2024)

6. TRANSFERABILITY AND GENERALISATION

6.1 Interpolation, Extrapolation, and Configurational-Space Coverage

Transferability is a feature of a machine learning interatomic potential (MLIP) that allows it to produce accurate predictions for atomic configurations outside of the range of those used to train the model. Interpolation in the training domain can be very accurate, but if the domain is extrapolated into new structures, this is a more difficult task. To enhance model robustness and predictive accuracy, it is crucial to ensure the coverage of a wide range of configurations across various training datasets (Batatia et al., 2025a).

6.2 Transfer Across Compositions, Phases, Temperatures, and Pressures

A good MLIP should be able to describe the materials under different chemical compositions and thermodynamic conditions. The training of models for transferring across phases, temperatures and pressures demands representative atomic environments spanning multiple structural states. Recently, the atomistic models applied at the atomic scale have shown better generalization, while they have used a very large multi-elements set and broad simulation conditions (Yang et al., 2024; Merchant et al., 2023).

6.3 Generalisation to Defects, Surfaces, Interfaces, and Reactions

In practical molecular simulations, defects, grain boundaries, free surfaces, interfaces and reactive processes are commonly found that are not in an equilibrium configuration. Descriptors must be able to represent a variety of atomic neighborhoods in these complex environments, but at the same time retain physical consistency. Efficient representations of the structure and descriptor compression have also improved predictive accuracy in chemically complex systems (Darby et al., 2022).

6.4 Uncertainty Quantification and Out-of-Distribution Detection

Mechanisms are needed to detect predictions outside the training distribution when deploying MLIPs. Uncertainty quantification and out-of-distribution detection can be used to identify unreliable predictions and call for extra reference calculations when needed. In recent years, pretrained frameworks for deep learning have been enhanced with strong training strategies and attention-based architectures that lead to better confidence estimation, reliability, and transferability to a wide range of atomistic simulations (Zhang et al., 2024). Table 3 provides an overview of the key issues found by MLIPs in the space of transferability and the strategies proposed to address these issues and enhance their generalization ability.

Table 3. Transferability Challenges and Solution Strategies

Challenge	Effect	Recent Solution	References
Limited training domain	Poor extrapolation	Foundation models	Batatia et al. (2025a)
Complex chemical compositions	Reduced accuracy	Large multi-element datasets	Merchant et al. (2023)
Unseen structures	Prediction uncertainty	Out-of-distribution detection	Zhang et al. (2024)
Diverse thermodynamic conditions	Weak generalization	MatterSim framework	Yang et al. (2024)
Descriptor limitations	Reduced transferability	Descriptor compression	Darby et al. (2022)

7. APPLICATIONS IN MOLECULAR AND MATERIALS SIMULATIONS

7.1 Molecular Dynamics, Thermodynamics, and Phase Transitions

Recently, the use of machine learning interatomic potentials (MLIPs) has greatly enhanced the range of molecular dynamics simulation applications by allowing the accurate prediction of the trajectory of atoms over long time and length scales. Their near quantum-mechanical accuracy enables accurate investigation of thermodynamic properties, diffusion processes, melting behavior, and solid–liquid phase transitions while significantly reducing computational cost, substantially compared to first-principles simulations (Jacobs et al., 2025; Batatia et al., 2025b). Figure 2 shows how machine learning interatomic potentials enable materials-dynamics simulations across extreme spatial, temporal, and thermodynamic conditions.

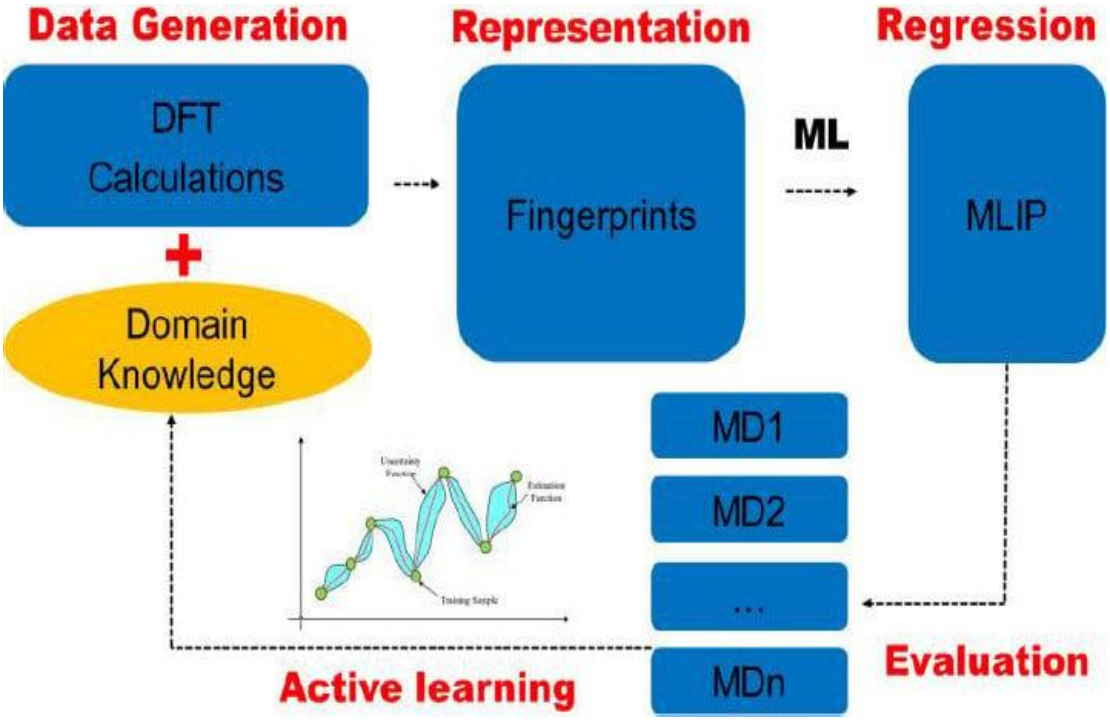

Figure 2. MLIP-enabled materials-dynamics simulations under extreme conditions (Yang et al., 2021)

7.2 Solid-State Materials, Defects, Alloys, and Nanomaterials

MLIPs are widely used on crystalline materials, amorphous solids, alloys, semiconductors and nanostructured systems. They support the understanding of lattice dynamics, defect formation, grain boundaries, dislocations, fracture mechanisms and microstructural evolution of a variety of material classes. More recently, the simulation of chemically complex materials has been enhanced by the ability to transfer atomistic representations from one material to another for a wide range of elements and environments (Kalita et al., 2025).

7.3 Chemical Reactions, Catalysis, and Energy Materials

Because MLIPs can describe the details of bond formation and bond breaking with great accuracy, they have been used in applications ranging from heterogeneous catalysis, reaction pathway analysis, electrochemical interfaces, battery materials, hydrogen storage, and fuel-cell technologies. Advanced functional materials for sustainable energy applications are being discovered and optimized at a rapid pace using large-scale atomistic simulations, with the assistance of foundation models (Yuan et al., 2026).

7.4 Molecular, Biological, and Pharmaceutical Systems

With the success of the MLIPs in biomolecules, organic compounds, pharmaceutical systems and large molecular assemblies simulation is growing. They are efficient in the evaluation of molecular conformation, protein–ligand interactions, solvent effects and protein stability yet have high predictive accuracy. As the models are becoming more data efficient to be constructed, their applicability in the fields of chemistry, biology, and drug discovery will be further enhanced, with the expectation of being less dependent on large amounts of first-principles data (Zhang et al., 2026). In Table 4, we list some representative application fields of ML interatomic potentials and their key computational targets. Figure 3 illustrates the AI workflow in space missions, integrating data collection, AI models, mission applications, and decision-support operations.

Table 4. Representative Applications of MLIPs

Application Area	Primary Objective	Representative Systems	References
Molecular Dynamics	Atomic trajectory prediction	Liquids, gases	Jacobs et al. (2025)
Solid-State Materials	Defects and lattice behavior	Metals, semiconductors	Kalita et al. (2025)
Energy Materials	Battery and catalytic simulations	Electrodes, catalysts	Yuan et al. (2026)
Nanomaterials	Surface and interface studies	2D materials, nanoparticles	Batatia et al. (2025b)
Biomolecular Systems	Molecular interactions	Proteins, drug molecules	Zhang et al. (2026)

Figure 3. AI workflow in space missions

8. CHALLENGES AND FUTURE PERSPECTIVES

Although significant advances have been made, there are still a few challenges that hinder the broader use of machine learning interatomic potentials (MLIPs). One of the key challenges is the availability of a variety of high-quality reference datasets and calculating first principles data for large and chemically complex systems is still expensive. The models may not be generalizable due to the limited amount of training data, and they may not be as reliable when performing simulations on novel atomic environments. Another major problem is the accurate description of long-range electrostatic interactions, dispersion forces, charge transfer and highly reactive systems, especially in the case of multicomponent materials, where it might be necessary to rely on local descriptors. There are no universally accepted benchmark datasets, standardized validation protocols, or reproducible evaluation frameworks to further compound the objective comparisons of different MLIP architectures. Furthermore, the lack of interpretability of many deep learning models decreases the trust in the model predictions, particularly in the context of scientific and engineering problems in which safety is an important consideration. Researchers should work on designing physically informed and data-efficient learning algorithms that can perform strong extrapolation over a variety of chemical compositions, thermodynamic conditions and structural configurations. The availability of universal atom-atom potentials and atom-atom foundation models, trained on large atomistic databases, is an exciting new path towards very transferable and widely applicable simulations. The use of transfer learning, active learning, uncertainty-aware modelling and automated datasets generation will enhance

the robustness of the model and decrease the computational cost. Moreover, coupling MLIPs with multiscale simulation methods and experimental characterization tools will aid seamless investigations spanning the atomic, mesoscale, and macroscopic scales.

9. CONCLUSION

Machine learning interatomic potentials (MLIPs) have revolutionized molecular and materials simulations, providing a viable way to effectively fill the gap between the high computational cost of quantum mechanical methods and the low level of accuracy of classical force fields. Based on sophisticated machine learning algorithms, physically informed atomic representations and high-quality reference datasets, MLIPs allow for near quantum-level accuracy while maintaining the computational efficiency needed for large-scale and long-timescale simulations. The ongoing development of neural network potentials, kernel-based approaches, atomic cluster expansion methods, graph neural networks and equivariant architectures has significantly improved the prediction of atomic energies and atomic forces in various chemical systems and material properties. The versatility and scientific relevance of MLIPs are highlighted by their increasing use in molecular dynamics, thermodynamics, materials in solid-state, catalysis, energy storage, nanotechnology, and biological systems. However, issues remain regarding their transferability, data availability, long-range interactions, uncertainty quantification, and model interpretability. To overcome these challenges, the ability to extend the currently available benchmark datasets, to use physically consistent model architectures, to adopt active learning strategies, and to incorporate uncertainty-aware frameworks will be crucial in the future for the development of more reliable and universally applicable interatomic potentials. Moreover, the next generation of atomistic modelling should receive a boost from the development of foundation models, transfer learning methods and multiscale simulation frameworks.

REFERENCES

1. Batatia, I., Batzner, S., Kovács, D. P., Musaelian, A., Simm, G. N., Drautz, R., ... & Csányi, G. (2025a). The design space of E (3)-equivariant atom-centred interatomic potentials. Nature Machine Intelligence, 7(1), 56-67.
2. Batatia, I., Benner, P., Chiang, Y., Elena, A. M., Kovács, D. P., Riebesell, J., ... & Csányi, G. (2025b). A foundation model for atomistic materials chemistry. The Journal of chemical physics, 163(18).
3. Batatia, I., Kovacs, D. P., Simm, G., Ortner, C., & Csányi, G. (2022). MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. Advances in neural information processing systems, 35, 11423-11436.
4. Batzner, S., Musaelian, A., Sun, L., Geiger, M., Mailoa, J. P., Kornbluth, M., ... & Kozinsky, B. (2022). E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. Nature communications, 13(1), 2453.
5. Behler, J. (2021). Four generations of high-dimensional neural network potentials. Chemical Reviews, 121(16), 10037-10072.
6. Chen, C., & Ong, S. P. (2022). A universal graph deep learning interatomic potential for the periodic table. Nature Computational Science, 2(11), 718-728.
7. Darby, J. P., Kermode, J. R., & Csányi, G. (2022). Compressing local atomic neighbourhood descriptors. npj Computational Materials, 8(1), 166.
8. Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C. J., & Ceder, G. (2023). CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. Nature Machine Intelligence, 5(9), 1031-1041.
9. Deringer, V. L., Caro, M. A., & Csányi, G. (2019). Machine learning interatomic potentials as emerging tools for materials science. Advanced Materials, 31(46), 1902765.
10. Drautz, R. (2020). Atomic cluster expansion of scalar, vectorial, and tensorial properties including magnetism and charge transfer. Physical Review B, 102(2), 024104.
11. Dusson, G., Bachmayr, M., Csányi, G., Drautz, R., Etter, S., van Der Oord, C., & Ortner, C. (2022). Atomic cluster expansion: Completeness, efficiency and stability. Journal of Computational Physics, 454, 110946.
12. Friederich, P., Häse, F., Proppe, J., & Aspuru-Guzik, A. (2021). Machine-learned potentials for next-generation matter simulations. Nature Materials, 20(6), 750-761.
13. Gasteiger, J., Becker, F., & Günnemann, S. (2021). Gemnet: Universal directional graph neural networks for molecules. Advances in neural information processing systems, 34, 6790-6802.
14. Gasteiger, J., Groß, J., & Günnemann, S. (2020). Directional message passing for molecular graphs. arXiv preprint arXiv:2003.03123.
15. Ghaffari, K., Bavdekar, S., Spearot, D. E., & Subhash, G. (2024). Validation workflow for machine learning interatomic potentials for complex ceramics. Computational Materials Science, 239, 112983.

16. Jacobs, R., Morgan, D., Attarian, S., Meng, J., Shen, C., Wu, Z., ... & Wood, B. M. (2025). A practical guide to machine learning interatomic potentials–Status and future. Current Opinion in Solid State and Materials Science, 35, 101214.

17. Kalita, B., Gokcan, H., & Isayev, O. (2025). Machine learning interatomic potentials at the centennial crossroads of quantum mechanics. Nature Computational Science, 5(12), 1120-1132.

18. Kocer, E., Ko, T. W., & Behler, J. (2022). Neural network potentials: A concise overview of methods. Annual review of physical chemistry, 73, 163-186.

19. Lysogorskiy, Y., Oord, C. V. D., Bochkarev, A., Menon, S., Rinaldi, M., Hammerschmidt, T., ... & Drautz, R. (2021). Performant implementation of the atomic cluster expansion (PACE) and application to copper and silicon. npj computational materials, 7(1), 97.

20. Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. Nature, 624(7990), 80-85.

21. Mishin, Y. (2021). Machine-learning interatomic potentials for materials science. Acta Materialia, 214, 116980.

22. Musaelian, A., Batzner, S., Johansson, A., Sun, L., Owen, C. J., Kornbluth, M., & Kozinsky, B. (2023). Learning local equivariant representations for large-scale atomistic dynamics. Nature Communications, 14(1), 579.

23. Musil, F., Grisafi, A., Bartók, A. P., Ortner, C., Csányi, G., & Ceriotti, M. (2021). Physics-inspired structural representations for molecules and materials. Chemical reviews, 121(16), 9759-9815.

24. Noé, F., Tkatchenko, A., Müller, K. R., & Clementi, C. (2020). Machine learning for molecular simulation. Annual review of physical chemistry, 71(1), 361-390.

25. Novikov, I. S., Gubaev, K., Podryabinkin, E. V., & Shapeev, A. V. (2021). The MLIP package: moment tensor potentials with MPI and active learning. Machine Learning: Science and Technology, 2(2), 025002.

26. Park, Y., Kim, J., Hwang, S., & Han, S. (2024). Scalable parallel algorithm for graph neural network interatomic potentials in molecular dynamics simulations. Journal of chemical theory and computation, 20(11), 4857-4868.

27. Podryabinkin, E. V., Tikhonov, E. V., Shapeev, A. V., & Oganov, A. R. (2019). Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. Physical Review B, 99(6), 064114.

28. Poul, M., Huber, L., Bitzek, E., & Neugebauer, J. (2023). Systematic atomic structure datasets for machine learning potentials: Application to defects in magnesium. Physical Review B, 107(10), 104103.

29. Qi, J., Ko, T. W., Wood, B. C., Pham, T. A., & Ong, S. P. (2024). Robust training of machine learning interatomic potentials with dimensionality reduction and stratified sampling. npj Computational Materials, 10(1), 43.

30. Röcken, S., & Zavadlav, J. (2024). Accurate machine learning force fields via experimental and simulation data fusion. Npj Computational Materials, 10(1), 69. https://doi.org/10.1038/s41524-024-01251-4

31. Schütt, K., Unke, O., & Gastegger, M. (2021, July). Equivariant message passing for the prediction of tensorial properties and molecular spectra. In International conference on machine learning (pp. 9377-9388). PMLR.

32. Sivaraman, G., Gallington, L., Krishnamoorthy, A. N., Stan, M., Csányi, G., Vázquez-Mayagoitia, Á., & Benmore, C. J. (2021). Experimentally driven automated machine-learned interatomic potential for a refractory oxide. Physical review letters, 126(15), 156002.

33. Tokita, A. M., & Behler, J. (2023). How to train a neural network potential. The Journal of Chemical Physics, 159(12).

34. Unke, O. T., Chmiela, S., Sauceda, H. E., Gastegger, M., Poltavsky, I., Schutt, K. T., ... & Muller, K. R. (2021). Machine learning force fields. Chemical reviews, 121(16), 10142-10186.

35. Vandermause, J., Torrisi, S. B., Batzner, S., Xie, Y., Sun, L., Kolpak, A. M., & Kozinsky, B. (2020). On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. npj Computational Materials, 6(1), 20.

36. Wang, G., Wang, C., Zhang, X., Li, Z., Zhou, J., & Sun, Z. (2024). Machine learning interatomic potential: Bridge the gap between small-scale models and realistic device-scale simulations. Iscience, 27(5).

37. Yang, H., Hu, C., Zhou, Y., Liu, X., Shi, Y., Li, J., ... & Lu, Z. (2024). Mattersim: A deep learning atomistic model across elements, temperatures and pressures. arXiv preprint arXiv:2405.04967.

38. Yang, H., Zhu, Y., Dong, E., Wu, Y., Yang, J., & Zhang, W. (2021). Dual adaptive sampling and

machine learning interatomic potentials for modeling materials with chemical bond hierarchy. Physical Review B, 104(9), 094310.

39. Yang, Y., Zhao, L., Han, C. X., Ding, X. D., Lookman, T., Sun, J., & Zong, H. X. (2021). Taking materials dynamics to new extremes using machine learning interatomic potentials. Journal of Materials Informatics, 1(2), N-A.

40. Yuan, E. C. Y., Liu, Y., Chen, J., Zhong, P., Raja, S., Kreiman, T., ... & Head-Gordon, T. (2026). Foundation models for atomistic simulation of chemistry and materials. Nature Reviews Chemistry, 10(3), 212-230.

41. Zhang, D., Bi, H., Dai, F. Z., Jiang, W., Liu, X., Zhang, L., & Wang, H. (2024). Pretraining of attention-based deep learning potential model for molecular simulation. npj Computational Materials, 10(1), 94.

42. Zhang, W., Wu, X., Wang, C., Hu, S., Liu, Y., & Wang, L. W. (2026). Constructing machine learning interatomic potentials with minimum amount of ab initio data. npj Computational Materials, 12(1), 174.