Available online at www.sciencedirect.com

ScienceDirect

Green Energy & Environment 11 (2026) 940–967

www.keaipublishing.com/gee

Review Article

Artificial intelligence-driven discovery and design of solid-state hydrogen storage materials

Yuyi Chu [a,b,1], Yanxin Li [c,1], Yanxi Lyu [a,b,e,1], Di Zhang [d], Hao Li [d,*], Weijie Yang [c,*], Jianqiu Li [b,e], Liang Zhang [a,b,e,*]

[a] Center for Combustion Energy, Tsinghua University, Beijing, 100084, China
[b] School of Vehicle and Mobility, State Key Laboratory of Intelligent Green Vehicle and Mobility, Tsinghua University, Beijing, 100084, China
[c] Department of Power Engineering, School of Energy, Power and Mechanical Engineering, North China Electric Power University, Baoding, 071003, China
[d] Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai, 980-8577, Japan
[e] Beijing Huairou Laboratory, Beijing, China

Received 11 September 2025; revised 2 December 2025; accepted 19 January 2026
Available online 13 February 2026

Abstract

Artificial intelligence (AI) is reshaping the discovery and optimization of solid-state hydrogen storage materials, a cornerstone of a scalable hydrogen economy. However, interdependent trade-offs among capacity, operating conditions, and cycling stability still limit progress. This review surveys AI-assisted advances in metallic hydrogen storage through the co-design of features and models. We consolidate descriptor sets that fuse intrinsic crystal, electronic-structure, and thermodynamic properties with extrinsic experimental conditions. We also systematically summarize machine learning approaches for performance prediction, physics-informed simulation, and materials and process optimization. We additionally describe AI-driven platforms that integrate curated datasets, forward–inverse modeling, workflow orchestration, and user-facing tools for high-throughput screening and synthesis-aware decision-making. Looking ahead, interpretability, cross-scale modeling, and large language model driven closed-loop discovery will accelerate the practical deployment of solid-state hydrogen storage.

Keywords: Solid-State Hydrogen Storage; Databases; Feature Engineering; Machine Learning; AI Platforms

1. Introduction

The contemporary world faces the urgent challenge of mitigating extreme climate change, which demands a rapid transition from fossil fuels to cleaner, sustainable energy systems [1,2]. Among the available alternatives, hydrogen energy—characterized by its zero-carbon emissions, high energy density, and potential for large-scale deployment—has emerged as a central focus of global energy strategies [3–5]. The hydrogen energy industry chain comprises three key segments: production, storage and transportation, and utilization [6,7]. Hydrogen storage, as the critical link between upstream production and downstream applications, not only underpins the efficient use of hydrogen but also remains a primary bottleneck to its large-scale commercialization [8].

Compared with high-pressure gaseous and cryogenic liquid storage, solid-state hydrogen storage offers advantages such as higher volumetric density, lower operating pressure, enhanced safety, and longer cycle life [9]. However, according to the U.S. Department of Energy (DOE) targets, solid-state hydrogen storage materials must simultaneously achieve a

* Corresponding authors:.
E-mail addresses: zhangbright@mail.tsinghua.edu.cn (L. Zhang), yangwj@ncepu.edu.cn (W. Yang), li.hao.b8@tohoku.ac.jp (H. Li).
Peer review under the responsibility of Editorial Board of Green Energy & Environment
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.gee.2026.01.007

gravimetric capacity of ≥5.5 wt.%, volumetric capacity ≥40 g L^{-1}, operating temperatures between room temperature and 100 °C, pressures between 5 and 12 bar (absolute), and stability over at least 500 charge–discharge cycles. Currently, no known metallic hydrogen storage material fully satisfies all these criteria. For instance, MgH_2 exceeds the gravimetric capacity requirement but requires desorption temperatures above 300 °C, far beyond DOE specifications [10–12]. This performance gap underscores the urgent need for advanced strategies to design and optimize solid-state hydrogen storage materials.

Conventional material development approaches—such as catalysis [13], doping [14], alloying [15], and nanosizing [16]—have yielded notable improvements but are inherently time-consuming, costly, and limited by the vast compositional and structural design space. Recent advances in artificial intelligence (AI) have opened a new paradigm for materials discovery, enabling rapid screening, predictive modeling, and multi-objective optimization by learning complex, nonlinear correlations between material descriptors and performance metrics [17,18]. In hydrogen storage research, AI can complement experimental and computational methods by accelerating the identification of promising candidates, guiding targeted composition modifications, and uncovering previously inaccessible structure–property relationships.

Data are the rate-limiting resource for AI in solid-state hydrogen storage. Without a large, high-quality, machine-readable database, even sophisticated models risk learning anecdotal correlations rather than transferable structure–property relationships. A database-first workflow consolidates heterogeneous literature and calculations into a unified schema with rigorous unit normalization, provenance tracking, deduplication, and quality flags. This foundation enables fair train/validation/test splits, unbiased benchmarking, uncertainty quantification, and ultimately multi-objective inverse design (Table 1). To the best of our knowledge, the Digital Hydrogen Platform (*DigHyd*: http://www.dighyd.org) is currently the only large-scale, cross-family database dedicated to solid-state hydrogen storage [19]. By turning scattered textual and graphical evidence into structured records, curated databases transform hydrogen-storage research from case-by-case studies into reproducible, scalable discovery pipelines.

In this review, we examine recent progress in AI-assisted research on metallic hydrogen storage materials, with emphasis on three interconnected aspects: (i) hydrogen storage feature engineering, encompassing intrinsic material properties and experimental parameters, (ii) machine learning (ML) applications, including performance prediction, simulation, and optimization, and (iii) AI-driven platforms that integrate datasets, forward–inverse models, and user-facing tools for screening and synthesis guidance. We further discuss prospective directions such as interpretability, cross-scale modeling, and integration of large language models (LLMs), aiming to provide a comprehensive roadmap for the accelerated discovery of solid-state hydrogen storage materials that meet DOE targets.

Table 1
Key DOE performance targets for onboard solid-state hydrogen storage system [20].

Parameter	2025 Target	Ultimate Target
Gravimetric Capacity (wt.%)	5.5	6.5
Volumetric Capacity (g H_2/L)	40	50
Operating Temperature (°C)	−40 to 60	−40 to 60
Minimum/Maximum Delivery Temperature (°C)	−40 to 85	−40 to 85
Cycle Life (1/4 tank to full)	1500	1500
Delivery Pressure (bar)	5 to 12	5 to 12

2. Existing Databases for Hydrogen Storage Research

In AI-assisted hydrogen-storage research, datasets serve two complementary functions: broad inorganic repositories provide machine-tractable structures and computed/experimental properties for feature construction and transfer learning, whereas domain-specific collections supply the hydrogen-storage labels (thermodynamic, kinetic, and adsorption) required for supervised modeling.

The first category includes the Materials Project (https://materialsproject.org), which offers DFT-derived structures, phase stability information, and a mature REST (Representational State Transfer) API; the Crystallography Open Database, COD (https://www.crystallography.net/cod/) and the Inorganic Crystal Structure Database, ICSD (https://icsd.products.fiz-karlsruhe.de/) as primary sources of experimentally determined crystal structures; the Open Quantum Materials Database, OQMD (https://oqmd.org) with large-scale formation energies and phase-diagram queries; the NOMAD Repository (https://nomad-lab.eu) enabling FAIR (Findable, Accessible, Interoperable, Reusable) archiving of raw and processed first-principles data; AFLOW (https://aflowlib.org) with extensive computed properties and a documented API; and the Pauling File/MPDS (https://paulingfile.com) that aggregates crystallography, phase diagrams, and selected physical properties. While indispensable for constructing graph/topological descriptors, coordination environments, and composition-based features, these resources rarely report hydrogen-storage performance labels such as $\Delta H/\Delta S$, van't Hoff/PCT parameters, H/M ratios, kinetics, or cycling durability, and they are most powerful when linked to hydrogen-specific datasets.

The second category comprises the cross-family, ML-ready platform DigHyd (https://www.dighyd.org), which integrates large-scale literature mining with provenance and uncertainty flags, explicit notes on activation/cycling where available, and task-oriented data splits that facilitate fair comparison across algorithms and smooth deployment; and the *Digital Hydrogen-S platform* (http://digital-hydrogen.com/storage/) [21], which compiles multimodal, experimentally grounded hydrogen-storage data from over 1,000 publications, covering more than 3,000 materials and 250,000+ structured records—including compositions, thermodynamics, PCT and kinetic curves, and synthesis parameters across major material families. For physisorption-dominated storage, labels are isotherm-centric and require careful handling of absolute versus excess uptake

and of test conditions; the NIST/ARPA-E Adsorption Isotherm Database (ISODB) (https://adsorption.nist.gov/isodb/index.php) is the principal open repository for single- and multi-component isotherms with rich metadata, and it pairs naturally with computation-ready framework corpora such as CoRE MOF (2014: https://doi.org/10.1021/cm502594j; 2019 update: https://doi.org/10.1021/acs.jced.9b00835) and CoRE-COF (https://core-cof.github.io/CoRE-COF-Database/), as well as simulated adsorption benchmarks like CRAFTED (https://arxiv.org/abs/2210.09456), all of which enable structure-informed surrogate models for H_2 uptake and deliverable capacity under application-relevant windows.

Taken together, these resources form a coherent data ecosystem in which general repositories provide the structural backbone and identifiers, hydride-focused databases contribute performance labels and ML-ready curation, and sorbent/isotherm platforms connect raw adsorption measurements to computation-ready frameworks. Before any selected or newly constructed dataset is adopted for subsequent ML, data cleaning and preprocessing are indispensable—one should first build a preliminary understanding of distributions, trends, missing values, and outliers through exploratory data analysis, then conduct necessary experimental validation and supplementation or, where appropriate, remove anomalous entries based on professional expertise. Afterwards, the dataset should be reasonably partitioned into training, validation, and testing subsets so that the resulting models maintain robust generalization and deliver considerable predictive ability even on unseen data.

3. Hydrogen Storage Feature Engineering

In hydrogen storage materials research, ML enables the discovery of complex, nonlinear relationships within multi-dimensional datasets by constructing adaptive, data-driven predictive models. However, the predictive power of these models is fundamentally constrained by the completeness and physical relevance of the selected features. Feature engineering—the process of designing and selecting appropriate descriptors—plays a pivotal role in mitigating the curse of dimensionality while enhancing the fidelity of structure–property relationship modeling [17]. In metallic hydrogen storage systems, well-chosen descriptors not only improve prediction accuracy but also provide mechanistic insights into hydrogen absorption and desorption. Broadly, features for performance prediction can be classified into two categories: intrinsic material properties (e.g., crystal structure parameters, electronic descriptors, thermodynamic constants) and extrinsic experimental parameters (e.g., temperature, pressure, catalyst composition). The following section outlines these categories with representative examples.

3.1. Material Properties

The features used for predicting hydrogen storage performance differ markedly across material systems. Here, we focus on two representative classes of metallic hydrogen storage materials: metal hydrides and metal–organic frameworks (MOFs).

3.1.1. Metal Hydrides

Metal hydrides are crystalline compounds formed by bonding metallic elements with hydrogen, in which hydrogen atoms typically occupy interstitial lattice sites as hydride anions (H^-) or participate in covalent/ionic bonding. These materials have attracted considerable attention for solid-state hydrogen storage owing to their high volumetric hydrogen density and reversible hydrogenation–dehydrogenation kinetics.

Traditional research on metal hydrides has relied heavily on experimental trial-and-error and theoretical approaches such as first-principles calculations [22] and molecular dynamics simulations [23]. While these methods have enabled important advances, they face two key challenges:

(1) Experimental inefficiency: Exploring the vast compositional space of multi-component alloys through direct experimentation is costly and time-intensive, rendering exhaustive validation impractical.
(2) Computational limitations: First-principles and molecular dynamics simulations can be computationally expensive, particularly when modeling multivariate, non-linear interactions.

These challenges highlight the value of ML approaches, which combine feature engineering with adaptive algorithms to identify the most informative descriptors, construct predictive models, and efficiently screen promising compositions [24,25–27]—thereby reducing the need for exhaustive experimental campaigns. A notable example is the Magpie feature framework proposed by Ward et al. [28], which provides a standardized set of compositional descriptors for inorganic materials and has been widely adopted in ML-based hydrogen storage research. Building on such frameworks, researchers have tailored descriptor sets for specific alloy systems by integrating compositional, structural, and thermodynamic parameters, each linked to hydride formation energetics, lattice stability, and absorption–desorption kinetics. The commonly used intrinsic properties of metal hydrides are summarized in Table 2.

For V–Cr–Ti–Fe alloys [30], selected descriptors include chemical composition, electronegativity, lattice constants, and atomic radii. La–Y–Ni-based alloys [29] are characterized by stoichiometric ratio (B/A), melting process (e.g., ingot melting, rapid solidification), rare earth content (La, Sm, Nd, Gd, Ce, Y), total rare earth fraction (TRE), and transition metal components (Ni, Co, Mn, Al). High-entropy alloys (HEAs) [31] employ parameters such as mixing enthalpy (ΔH_{mix}), mixing entropy (ΔS_{mix}), atomic size difference (δ), electronegativity difference ($\Delta\chi$), valence electron concentration (VEC), elemental composition, and phase type to capture the configurational complexity of multi-principal-element systems (Fig. 1a). In AB_2-type alloys [32], temperature, hydrogen absorption/desorption pressure, and concentrations of key alloying elements (Ti, Cr, Mn, Fe, V, Zr, Ni, Cu, Co) enable comprehensive mapping of pressure–composition–temperature (PCT)

Table 2
Material properties commonly found as features in metal hydrides.

Feature category	Feature description	Ref.
Magpie feature (145)	Stoichiometric attributes	[24]
	Elemental property statistics	
	Electronic structure attributes	
	Ionic compound attributes	
Fundamental properties of elements	Atomic number	[17]
	Atomic radius	
	Atomic mass	
	Atomic density	
	Elemental composition	
	Electron affinity energy	
	Ionization potential	
	Bulk modulus	
	Shear modulus	
	Poisson's ratio	
Thermodynamic parameters	Entropy of fusion	[29]
	Enthalpy of mixing	
	Gibbs free energy of mixing	
	Entropy of mixing	
Hydrogenation properties	Metal-Metal bond-energy	[17]
	Metal-Hydrogen bond-energy	
	Metal-Metal bond-length	
	Metal-Hydrogen bond-length	
	Volume alteration during hydrogenation	
	Energy alteration during hydrogenation	
	Enthalpy of formation of binary hydrides	
	Pauling electronegativity	
	Proportion of hydrogen affinity elements	
	D-band center	

relationships. For C14-Laves-type alloys [33], mean ionic character and Fe content emerge as the most influential capacity-governing factors, while elemental descriptors are restricted to frequently occurring species (e.g., Ti, Mn, Zr, Cr, Fe, V) to mitigate data sparsity and overfitting. Large-scale datasets of general metal hydrides incorporate descriptors (Fig. 1b) such as lattice parameters, VEC, ΔH_{mix}, and other structural–thermodynamic indicators, with redundant variables eliminated to preserve model interpretability [34] (see Section 4.3.1 for model details).

3.1.2. MOFs

MOFs are crystalline coordination polymers composed of metal nodes and organic linkers, exhibiting exceptionally high surface areas and tunable pore structures, making them highly attractive for hydrogen physisorption [36]. Their structural diversity—arising from variations in metal nodes (e.g., Zn^{2+}, Cu^{2+}, Cr^{3+}) and organic linkers (e.g., carboxylates, imidazolates)—creates an immense design space encompassing nearly infinite topological configurations (such as cubic, tetrahedral, hexagonal prism) [37–39]. This diversity poses a significant challenge for experimental screening, highlighting the need for effective descriptor construction to capture structure–property relationships. Based on literature and data analyses, this work distills key factors governing MOFs hydrogen storage performance (Table 3), offering a concise reference for data-driven design. For example, Ahmed et al. [40] compiled

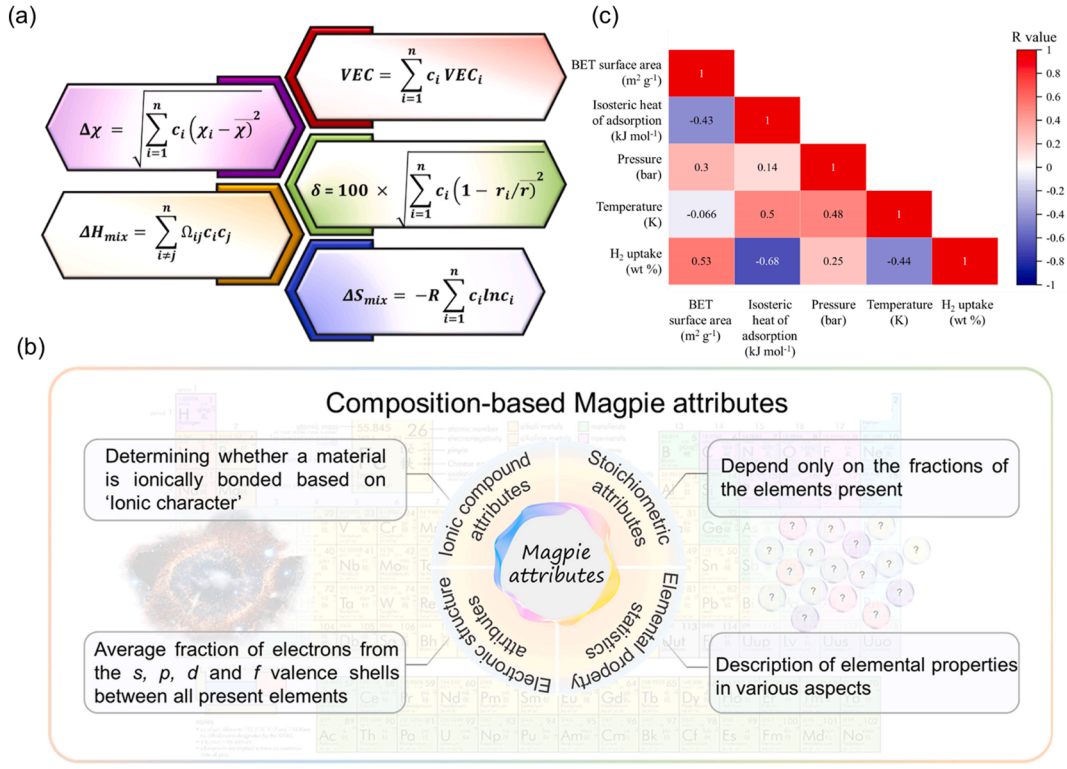

Fig. 1. (a) Formulation of thermodynamic parameter relationships [31]; (b) Overview of Magpie descriptors derived from alloy composition [34]; (c) Heatmap of the Pearson correlation coefficients among the selected material and experimental parameters [35].

Table 3
Material properties commonly found in MOF materials as features.

Feature category	Feature description	Ref.
Structural properties	Crystal density	[44]
	Pore volume	
	Void fraction	
	Pore limiting diameter	
	Largest cavity diameter	
	BET surface area	
	Pore size distribution	
	Gravimetric surface areas	
	Volumetric surface areas	
Chemical properties	Metal type	[38]
	Linker type	
	Functional groups	
	Topology	
Thermodynamic parameters	Isosteric heat of adsorption	[39]
	Adsorption enthalpy	
	Adsorption entropy	

Table 4
Common experimental conditions used as features.

Feature category	Feature description	Ref.
Processing conditions	Ball-to-powder ratio	[29]
	Ball-milling rotating speed	
	Ball-milling time	
	Melt spinning rate	
	Cold rolling rotation times	
	Annealing temperature	
De-/hydrogenation conditions	Atmosphere	[45–47]
	hydrogenation temperature	
	hydrogenation pressure	
	Dehydrogenation temperature	
	Dehydrogenation pressure	
	Rate of temperature rise	
	Number of cycles	

918,734 MOFs from 19 databases and distilled the input space to seven key crystallographic descriptors—single-crystal density, pore volume, gravimetric surface area, volumetric surface area, void fraction, largest cavity diameter, and pore limiting diameter. Borja et al. [41] selected five geometric descriptors—crystal density, pore volume, void fraction, pore-limiting diameter, and maximum cavity diameter—to predict both gravimetric and volumetric adsorption capacities, and further considered mechanical stability parameters (e.g., bulk modulus) derived from crystal density, void fraction, maximum cavity diameter, and gravimetric surface area. Giappa et al. [42] examined ligand functionalization effects by computing hydrogen binding energies for 58 functionalized benzene rings, thus incorporating chemical functional group descriptors alongside structural ones. Shekhar et al. [43] emphasized pore-related metrics such as pore volume, density, void fraction, pore-limiting diameter, and largest cavity diameter, with subsequent analysis revealing pore volume and void fraction as the most influential factors for adsorption performance.

Taken together, these studies demonstrate that feature engineering for MOFs in hydrogen storage applications centers on a combination of pore geometry parameters and chemical functionality descriptors (ligand types, functional groups). Such descriptors effectively bridge microscopic structure and macroscopic adsorption performance, providing a robust foundation for data-driven screening and design of high-capacity MOFs.

3.2. Experimental condition

The performance of metal hydrogen storage materials is governed not only by their intrinsic physicochemical properties but also by external experimental parameters such as hydrogenation/dehydrogenation temperature, pressure, and ball-milling time. These factors critically affect hydrogen diffusion kinetics and thermodynamic equilibrium. Over decades of experimental studies, a wealth of data has been accumulated that links operating conditions with performance metrics (e.g., PCT curves, hydrogen absorption/desorption rates), thereby providing rich, multidimensional feature sets for machine learning analysis. Based on a review of relevant literature, this work systematically summarizes common experimental parameters influencing hydrogen storage performance, as detailed in Table 4.

Kim et al. [32] investigated AB_2-type hydrogen storage alloys, selecting the alloy's chemical composition and hydrogen absorption pressure as key descriptors to predict their PCT behavior across varying temperatures. Similarly, Çolak et al. [24] focused on identifying critical operational parameters, defining temperature, pressure, and orientation values as inputs to model adsorption volume and gravimetric hydrogen storage capacity. Extending the scope to porous materials, Meduri et al. [35] examined MOFs and identified temperature, pressure, isosteric heat of adsorption, and Brunauer-Emmett-Teller (BET) surface area as the most influential factors governing hydrogen uptake (Fig. 1c). Notably, their dataset revealed strong multicollinearity among most features, with pressure consistently acting as the central interacting variable. These studies collectively highlight that judicious selection of physically meaningful parameters—spanning compositional, structural, and thermodynamic properties—is essential for accurately capturing hydrogen storage behavior across diverse material systems.

4. Machine Learning Applications in Hydrogen Storage

Built on the descriptor and experimental-condition space defined in Section 3, this section surveys the mainstream method families used in hydrogen-storage research. In practice, supervised learning is the workhorse: trained on labeled data, these models learn quantitative mappings from composition/structure and processing variables to properties such as capacity, ΔH, PCT behavior, and T_{des}. They provide calibratable accuracy, support uncertainty estimation, and integrate naturally with physics-based constraints. Unsupervised learning operates without labels, emphasizing similarity metrics, dimensionality reduction, anomaly detection, and prototype discovery; it is indispensable for data curation and

motif mining but does not directly yield calibrated property predictions. Accordingly, Section 4 is organized by method families rather than by learning paradigm: Linear models (4.1) offer interpretable baselines; Kernel methods (4.2) capture nonlinearity efficiently in small-to-moderate datasets; Neural networks (4.3) include feedforward/backprop models, graph neural networks, radial-basis networks, and a dedicated subsection on generative neural architectures; Ensemble models (4.4) improve robustness through bagging, boosting, and advanced stacking architectures; Clustering (4.5) supports unsupervised grouping and pattern discovery; and Neural Network Potentials (4.6) enable near-DFT-fidelity atomistic simulations for finite-temperature properties and kinetics. In this organization, 4.1–4.4 and 4.6 are predominantly supervised (with the exception of 4.3.4 on generative models), while 4.5 is unsupervised. For a consolidated view of these methods, please refer to Table 5 (Summary of representative AI models and material systems).

4.1. Linear Models

Linear models, one of the most fundamental machine learning algorithms, aim to capture the optimal linear relationship between feature values and target values. These models are particularly valued for their simplicity, interpretability, and ease of result comprehension. In their basic form, linear models express the relationship as:

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_p x_{i,p} + c \tag{1}$$

Where y_i represents the observed value for the i-th sample, $x_{i,1}$, $x_{i,2}, \ldots, x_{i,p}$ are the feature values, and β_0 is the intercept. The coefficients $\beta_1, \beta_2, \ldots, \beta_p$ reflect the influence of each feature on the target, while c denotes the error term. These models offer clear advantages, such as ease of interpretation and the ability to understand the relationships between variables directly.

However, linear models are highly sensitive to outliers and can exhibit instability in coefficient estimates, particularly when independent variables are highly correlated. To mitigate these issues, regularization techniques are often employed, enhancing model performance by reducing overfitting and selecting the most relevant features. A typical regularized model is expressed as:

$$\text{loss} = \min_{\beta_0,\beta_1,\ldots,\beta_p} \left\{ \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 + l \right\} \tag{2}$$

where $\widehat{y}_i$ is the predicted value and l represents the penalty term. Regularization techniques can be broadly categorized into three types: l_1, l_2, and a combination of both, known as $l_1 + l_2$.

$$l_1 = \lambda \sum_{j=1}^{p} |\beta_j| \tag{3}$$

where λ controls the penalty intensity. In contrast, Ridge regression introduces l_2 regularization, where the penalty term sums the squared values of the coefficients. Ridge regression reduces overfitting by shrinking coefficients but does not shrink them entirely to zero, which makes it less effective in selecting features. The Ridge penalty term is given by:

$$l_2 = \lambda \sum_{j=1}^{p} \beta_j^2 \tag{4}$$

Elastic Net combines the benefits of both Least absolute shrinkage and selection operator (LASSO) and Ridge regularization. By using both l_1 and l_2 penalties, Elastic Net produces sparse solutions while reducing overfitting, addressing some of the limitations of using either LASSO or Ridge alone. The Elastic Net penalty term is:

$$l_{\text{elastic-net}} = \lambda_1 \sum_{i=1}^{p} |\beta_i| + \lambda_2 \sum_{i=1}^{p} \beta_i^2 \tag{5}$$

Despite their advantages, linear models have significant limitations when it comes to solving complex, nonlinear problems. While they are invaluable for simpler tasks, their performance diminishes when the relationships between variables become more intricate and non-linear, such as in more complex materials research scenarios. Gao et al. [72] constructed a predictive framework for hydrogen absorption behavior in Ti–Mn based AB_2-type solid-state hydrogen storage devices using a combination of empirical modeling and machine learning techniques. By fitting experimental isothermal absorption curves with the Richards growth model, they parametrized the hydrogen uptake kinetics through four descriptors (K, B, Q, and ν), with the BFGS algorithm applied for initial parameter optimization. Three algorithms—linear regression, gradient descent, and ANN—were trained to predict the key Richards parameters (K and B) based on operating conditions (temperature and flow rate). Among them, the linear regression model achieved the most stable performance across varying temperatures and rates ($R^2 \geq 0.96$), offering a physically interpretable expression that captured the parabolic trend of storage capacity with flow rate. Chemically, the analysis revealed that higher temperatures and faster hydrogen injection rates led to reduced storage capacity, primarily due to impaired reaction kinetics and pressure build-up. The study underscores the feasibility of integrating data-driven strategies with physical growth models to forecast absorption profiles in solid-state systems under real-world dynamic conditions.

Li et al. [48] propose a physics-grounded, data-driven linear descriptor model for MgH_2 dehydrogenation barriers, bypassing repeated transition-state searches (Fig. 2a). Starting from CI-NEB/AIMD pathways on MgH_2(110), they construct $M = f_H \cdot f_e \cdot (d_{\text{H-H}} - d_{\text{H}_2}) \cdot S_{max}$ (Fig. 2b), where S_{max} is the maximum −ICOHP of Mg–H bonds around the two reacting H atoms, $f_H$captures local H-vacancy topology, and f_e encodes neighborhood electronegativity (including transition-metal dopants). A single linear map $D_{max} = 1.83\,M - 1.66$, reproduces 28 barriers with $R^2 \approx 0.95$ and out-of-sample errors of 0.04–0.11 eV; the predictions correlate with experimental onset temperatures and reveal actionable trends, e.g., Mn/Ni lowering barriers toward high-temperature proton

Table 5
Summary of representative AI models and material systems.

Model	Material system	Main outcomes	Key findings	Year	Ref.
Linear descriptor model	Mg-based	$R^2 \approx 0.95$ (reproduced 28 barriers); out-of-sample error 0.04−0.11 eV	Mn/Ni doping lowers barriers; electronegativity is a key lever	2024	[48]
SVM	Laves-type	R^2>0.96 for cell volume/equilibrium pressure prediction	−	2023	[33]
SVR	Porous carbon	Successfully forecast hydrogen adsorption performance	−	2021	[49]
GPR	AB_2	Exponential kernel performs best ($R^2 = 0.969$, MRE = 2.29%)	Ti and Zr have the most significant influence on ΔH	2022	[50]
GPR	AB_2/HEA	ΔH predictions validated by van't Hoff/DFT; reversible hydrogen uptake ≈1.6 wt.%	$Ti_{0.5}Zr_{1.5}CrMnFeNi$ ($\Delta H \approx -36.5$ kJ mol^{-1}), $TiZrCrMnFeNi$ ($\Delta H \approx -28.8$ kJ mol^{-1})	2024	[51]
GPR + Bayesian optimization	HEA	Screened 2314 HEAs; optimized ΔE_{mono}/ΔE_{di}; correlated mechanical/electronic structures	VNbCrMoMn (predicted capacity 2.83 wt.% H_2)	2024	[52]
MLP	Mg-based	Test set $R^2 = 0.922$; SHAP identified key descriptors	99 Mg−1La, 95 Mg−1Ni−4La (Demax >5.6 wt.% @275 °C)	2023	[34]
BPNN + GA	Hydride reactor	R^2>0.98 (lifespan prediction); 8% improvement in design lifespan	Cooling channel diameter and spacing are key geometric parameters	2024	[53]
DNN	AB_2	$R^2 \approx 0.93$ (PCT behavior prediction); van't Hoff extrapolation for cross-temperature prediction	Zr/V-rich alloys enhance hydrogen uptake and lower equilibrium pressure	2023	[32]
CGCNN	Pd-based	Formation energy prediction with MAE<1 meV atom^{-1}; constructed PCT curves	Rh doping increases plateau pressure; Ag doping decreases plateau pressure	2024	[54]
CGCNN	Borohydride	Screened 350,000 structures; recovered known phases; predicted new polymorphs	31 La−B−H candidates; stable $La(BH)_{12}$ polymorph	2024	[55]
MEGNet (GNN)	Mg-based	Global MAE = 9.1 kJ mol^{-1} H_2; MAE = 5.5 kJ mol^{-1} H_2 for Mg−Ni−M systems	Materials with $\Delta H \approx -39.2$ kJ mol^{-1} H_2 are promising candidates	2022	[56]
GNN	Mg-based	MAE = 5.5 kJ mol^{-1} H_2 for Mg systems; thermodynamic screening	32 candidates including $MgBe_{13}$, $YMgNi_4$; atomic volume correlates with ΔH	2023	[57]
GAN	Mg-based	Generated 500 Mg−H nanosheets; DFT screening identified new polymorphs	Li-decorated 2D MgH_2 (theoretical capacity ≈6 wt.%)	2022	[58]
RBNN	$LaNi_5$-based	Prediction error <2.5%; outperforms empirical correlations	Thermal conductivity has a dominant influence over geometric parameters	2021	[59]
DeePMD	Mg-based	RMSE as low as 0.43 meV atom^{-1} (Mg−Na); outperforms KRR	Na/K doping achieves highest accuracy; transition metals like Cr/Mn are more complex	2023	[60]
MLP-MD	Ca-based / Polyhydride	Observed surface melting-assisted hydrogenation; discovered liquid CaH_4 intermediate	Synthesis criteria for polyhydrides: product melting temperature, pressure-dependent hydrogenation enthalpy	2025	[60]
RF	AB_2	Capacity prediction $R^2 = 0.688$; identified elemental effects	Ni destabilizes hydride; Cr stabilizes C14 phase; Mn enhances hydrogen uptake	2022	[25]
RF	Metal hydride	Screened hydrides	alkali metal/boron/transition metal compounds are promising candidates	2023	[61]
ETR	Alloy	Screened 6480 candidates (H_2 wt.% >2.5 @300 K; ΔH<60 kJ mol^{-1})	Mg/V/Ti-based alloys (Sc/Cr/Mn doping)	2024	[62]
Ensemble + GA	$LaNi_5$-based	$R^2 = 1.00$; optimized fin structure → 12.8% increase in hydrogen storage at 400 s	Optimal fin volume shows non-monotonic dependence on adsorption time	2024	[63]
XGBoost	Multicomponent alloy	MAE<9 kJ mol^{-1}; Pearson correlation coefficient>0.6	−	2024	[64]
GBDT	Metal hydride	$R^2 = 0.83$ for hydrogen content prediction; complex hydrides up to 10.6 wt.%	Complex hydrides and Mg-based alloys are top candidates	2019	[65]
GBR	Mg-based	Identified Td-controlling factors; screened 132 compositions	$Mg_{0.875}Li_{0.125}H_2$ (T_d~400 K), $Mg_{0.875}Be_{0.125}H_2$ (T_d~500 K)	2024	[66]
GBDT	Intermetallic	Identified vpa as key descriptor; predicted UNi_5 behavior	UNi_5 (low-stability hydride, suitable for high-pressure scenarios)	2020	[67]
Stacked ensemble	Alloy	MSE = 0.187, $R^2 = 0.783$; experimental validation error<1.5%	Alloys with predicted capacity >3.45 wt.%	2022	[30]
GBT	HEA	ΔH decreases linearly by ~10 kJ mol^{-1} H_2 with increasing Fe content; validated by DFT/experiment	−	2024	[68]
XGBoost, RF	$LaNi_5$-based	R^2>0.95 for hydrogen uptake/bed temperature prediction	Increased supply pressure, decreased HTF temperature, and increased HTF flow enhance adsorption	2023	[69]

(*continued on next page*)

Table 5 (*continued*)

Model	Material system	Main outcomes	Key findings	Year	Ref.
EoE	$LiBH_4$-based	$R^2 = 0.888$ for single-component; $R^2 = 0.834$ for binary component generalization	TiF_3 and AlF_3 are effective dopants; excessive doping impairs reversibility	2020	[70]
GBR	HEA	Screened over 17,000 HEAs; Pareto optimality; experimental validation	MgTiVZrNbHf, MgTiVCrNb, MgAlTiVCr (experimentally validated)	2023	[71]
Explicit White Box ML	All	Created interpretable models for hydride performance	Identified Be-based hydrides as rare high-performance outliers.	2025	[19]

exchange membrane (HT-PEM) operating windows, with electronegativity as a key lever (Fig. 2c-d). This study exemplifies how interpretable, first-principles-derived descriptors coupled with lightweight linear surrogates can deliver mechanism-aware screening and rapid down-selection in solid-state hydrogen-storage hydrides.

Building on such foundational efforts, subsequent studies have increasingly embraced more flexible and expressive machine learning models to capture the complex, nonlinear along with high-dimensional relationships inherent in hydrogen storage materials. The following sections present a categorized overview of representative ML algorithms, detailing their implementation strategies and their roles in advancing hydrogen storage research.

4.2. Kernel Methods

Kernel-based methods, including Support Vector Machine (SVM) and Gaussian Process Regression (GPR), offer powerful capabilities for modeling complex, nonlinear relationships by implicitly mapping input features into high-dimensional spaces through kernel functions such as linear, polynomial, radial basis, sigmoid, and exponential. These approaches have found wide application in hydrogen storage research, where both accuracy and interpretability are critical for navigating vast compositional design spaces.

SVM is one of the most representative kernel methods and encompasses two major variants: support vector classification (SVC) and support vector regression (SVR). SVC aims to identify the optimal hyperplane (or hypersurface in higher dimensions) that separates different classes with the maximum margin from support vectors—the most critical training samples. SVR, on the other hand, fits a function that lies within a specified margin of tolerance from the data points, balancing accuracy and generalization to avoid overfitting. Compared to traditional linear models, SVM offers stronger generalization, robustness to outliers, and the ability to model nonlinear dependencies, making it particularly attractive in materials modeling, especially when working with small or noisy datasets.

SVMs have been successfully applied to predict structural and thermodynamic properties of hydrogen storage alloys. In the comprehensive machine learning framework developed by

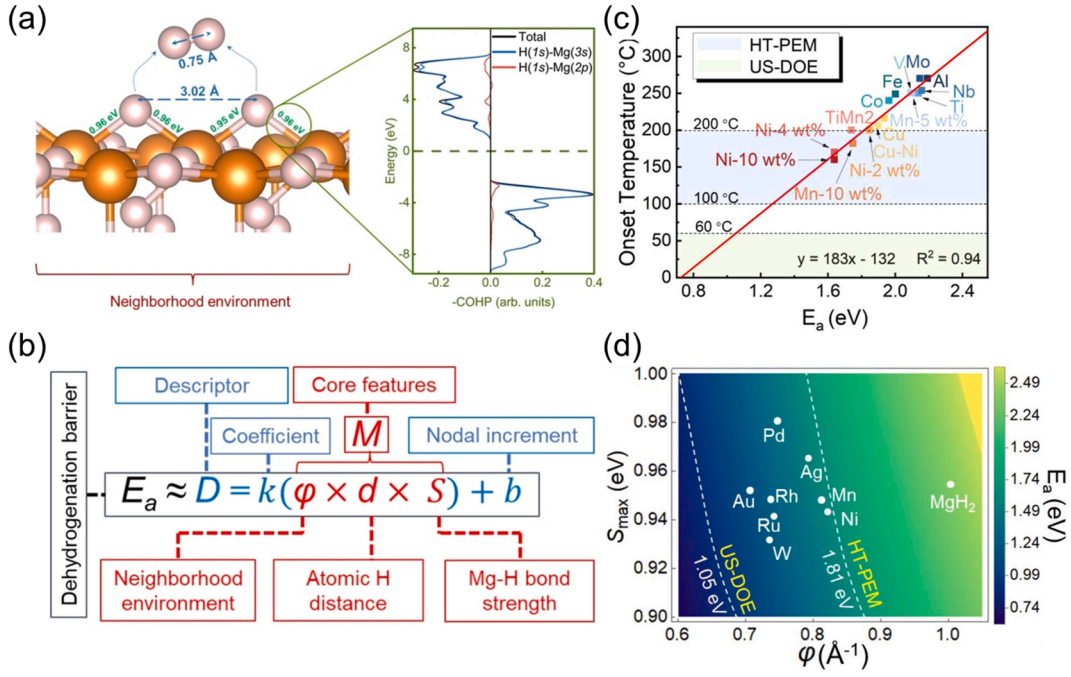

Fig. 2. Kinetic descriptors for MgH_2 dehydrogenation [48]. (a) Representative MgH_2 dehydrogenation with bonding-state map (H: white; Mg: orange). (b) Descriptor scheme: E_a (barrier), D (barrier descriptor); ϕ (local environment), d (H–H distance), S (Mg–H bond strength). M aggregates $\{\phi, d, S\}$; k and b linearly map $M \rightarrow D$. (c) Out-of-sample error histograms for training (blue) and test (yellow). (d) Onset dehydrogenation temperature vs E_a. Predicted barrier for metal-doped MgH_2 as a function of S_{max} and ϕ; white dashed lines mark HT-PEM and US-DOE targets.

Zhou et al. [33], various algorithms were benchmarked to predict the structural and thermodynamic properties of single-phase C14 Laves-type hydrides using Magpie-derived composition features (Fig. 3a-d). Among them, the SVM model consistently delivered superior predictive performance for key targets such as cell volume and hydrogen equilibrium pressures (Pa and Pd), with R^2 values exceeding 0.96 across multiple dataset splits. This outperformance stems from SVM's inherent strength in capturing complex nonlinear relationships within high-dimensional feature spaces—a crucial advantage given the compositional intricacies and subtle structure–property correlations in multi-component Laves-phase alloys. The kernel-based nature of SVM enabled it to efficiently exploit the informative Magpie descriptors, yielding robust and accurate property estimations that make it a particularly effective tool for accelerating compositional screening in hydrogen storage material discovery.

SVR, a regression variant of SVM, constructs an optimal function that approximates the training data within a defined tolerance margin. In hydrogen storage research, SVR has been

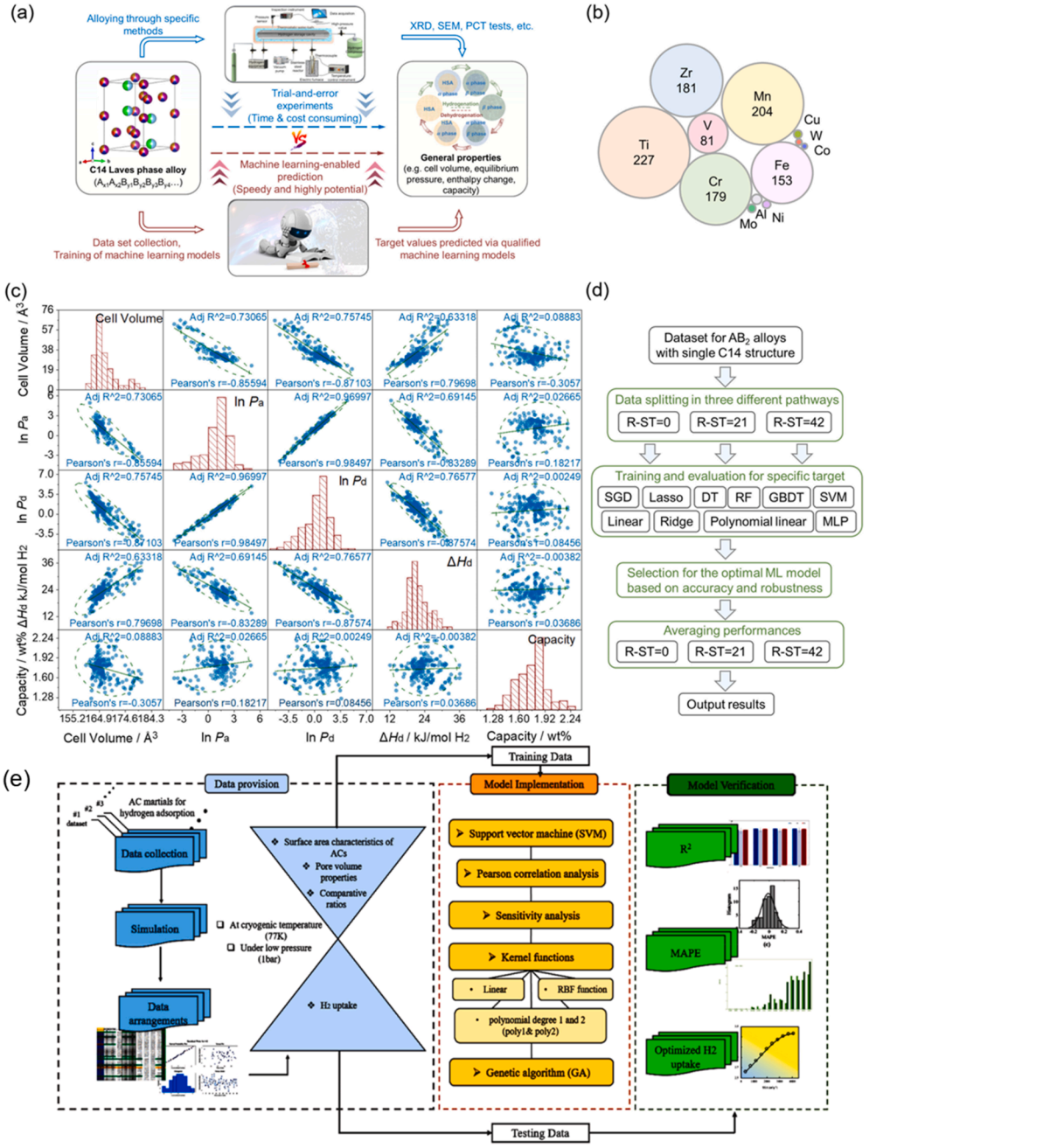

Fig. 3. (a) Contrast between conventional trial-and-error experimental approaches and the composition optimization strategy guided by machine learning; (b) Distribution frequency of elements within the dataset; (c) Pairwise scatter plots illustrating correlations among key target properties; (d) Schematic diagram summarizing the machine learning workflow [33]. (e) A diagrammatic summary of Rahimi et al.'s approach to dataset preparation and machine learning deployment [49].

employed to predict the hydrogen adsorption behavior of porous materials. For example, Rahimi et al. [49] developed an SVR model to successfully forecast the hydrogen uptake performance of activated carbon materials, demonstrating SVM's capability in capturing complex structure–property relationships (Fig. 3e). Despite its strong generalization ability and robustness against outliers, the predictive performance of SVM remains sensitive to kernel selection and hyperparameter tuning, and its computational cost can be high for large datasets.

As a probabilistic counterpart to SVM, GPR also employs kernel functions to capture nonlinear relationships, but instead of producing a single best-fit function, GPR infers a distribution over functions that are consistent with the training data. This Bayesian nature enables GPR to not only deliver accurate predictions, but also quantify the uncertainty associated with them—an important feature when extrapolating to unexplored material spaces. Both SVM and GPR fall under the category of kernel-based learning methods and can operate with similar kernel functions, yet GPR differs in its fully probabilistic framework and nonparametric formulation, which often allows greater flexibility and interpretability. Gheytanzadeh et al. [50] constructed GPR models using four different kernel functions to predict the hydrogen absorption enthalpy (ΔH) of 314 AB_2-type

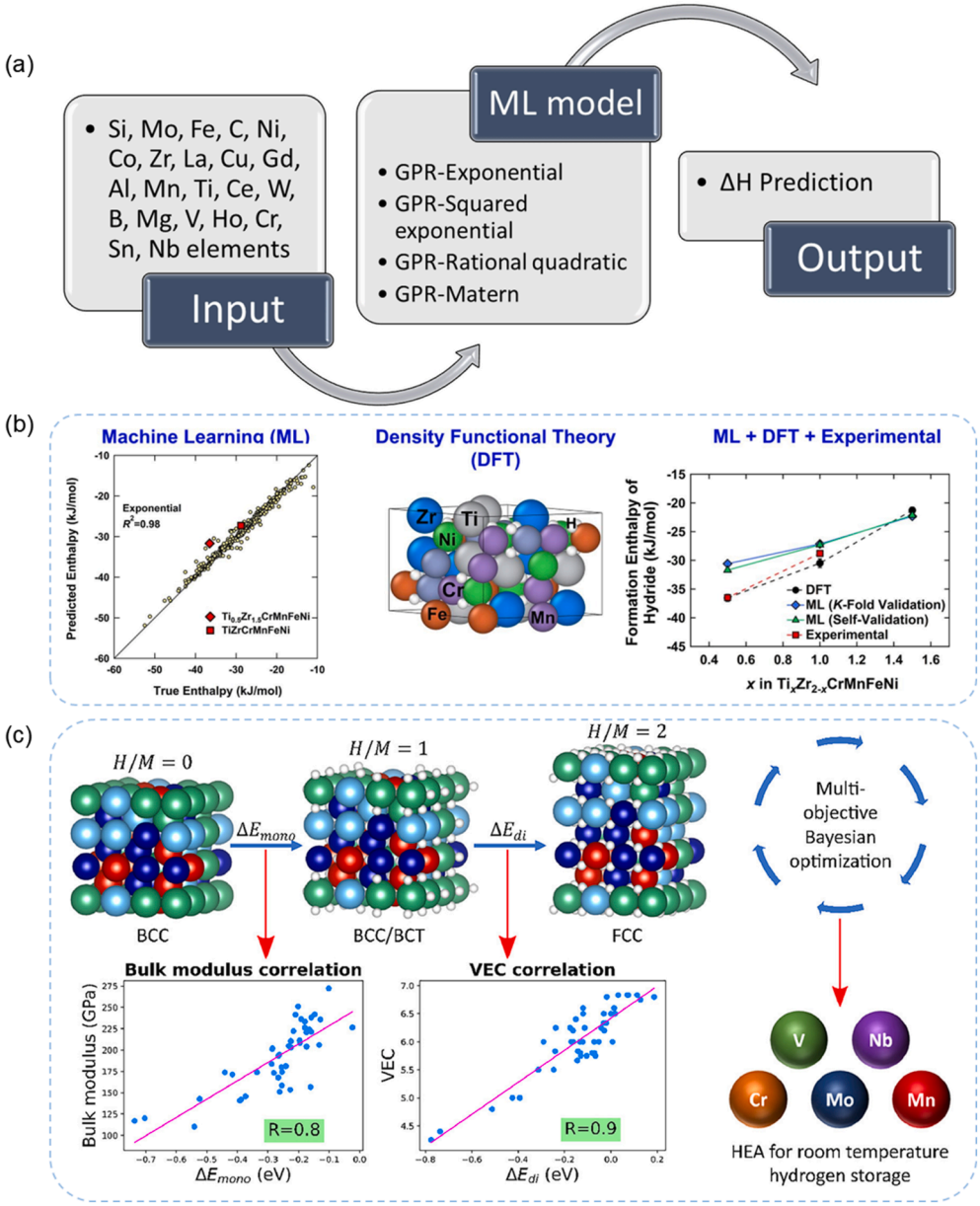

Fig. 4. (a) Schematic workflow of GPR-based prediction of hydride formation enthalpy (ΔH) from compositional inputs across 24 elements using multiple kernel functions [50]. (b) Validation of the GPR model by comparing predicted versus density functional theory (DFT)-calculated enthalpies, along with cross-validation and experimental benchmarking for $Ti_xZr_{2-x}CrMnFeNi$ alloys [51]. (c) Multi-objective Bayesian optimization combining GPR predictions and DFT calculations to identify Body-Centered Cubic (BCC) HEAs with optimal hydrogen binding energies (ΔE_{mono}, ΔE_{di}) and establish their correlations with bulk modulus and VEC [52].

metal hydrides (Fig. 4a). By encoding the chemical composition of A- and B-site elements (e.g., Ti, Zr, Cr, V) as input features, the model learned to capture the complex dependence of ΔH on alloy chemistry. Among the kernels tested, the exponential kernel performed best, yielding an R^2 of 0.969 and a mean relative error of only 2.29%. Further sensitivity analysis revealed that Ti and Zr exerted the most dominant influence on ΔH, consistent with their role as primary hydride-forming elements. The study highlights GPR as a powerful tool for guiding materials selection, particularly when experimental data are scarce and uncertainty quantification is critical. To accelerate the discovery of HEAs with suitable thermodynamics for room-temperature hydrogen storage, Dangwal et al. [51] developed a machine learning framework based on GPR, trained on 420 curated literature data points of hydride formation enthalpy (Fig. 4b). Chemically, the screening focused on AB_2-type alloys, incorporating elemental composition descriptors across 24 elements. The output metric—hydride formation enthalpy (ΔH)—was validated against both experimental van't Hoff analysis and DFT calculations. The model accurately predicted ΔH for $Ti_2Zr_{2-x}CrMnFeNi$ (x = 0.5–1.5), identifying $Ti_{0.5}Zr_{1.5}CrMnFeNi$ ($\Delta H \approx -36.5$ kJ mol^{-1}) and TiZrCrMnFeNi ($\Delta H \approx -28.8$ kJ mol^{-1}) as promising candidates. These alloys exhibited reversible hydrogen uptake (~1.6 wt.%) at room temperature without activation, attributed to their optimal enthalpies within the −25 to −39 kJ mol^{-1} target range. Pushing the scale further, Halpren et al. [52] implemented a multi-objective Bayesian optimization framework, integrating GPR models with DFT calculations, to screen 2314 equimolar BCC HEAs composed of 4–6 transition metals for room-temperature hydrogen storage (Fig. 4c). The model optimized two key thermodynamic descriptors—hydrogen binding energies for monohydride (ΔE_{mono}) and dihydride (ΔE_{di}) formation—targeting values near −0.12 eV, which favor ambient-pressure sorption. Chemically, ΔE_{mono} was found to correlate with the alloy's bulk modulus, while ΔE_{di} depended on d-band characteristics (notably the VEC). This strategy led to the discovery of VNbCrMoMn, exhibiting a theoretical capacity of 2.83 wt.% H_2—substantially surpassing $LaNi_5H_6$ (1.38 wt.%) and $TiFeH_2$ (1.91 wt.%). The work highlights how mechanical rigidity and electronic structure jointly determine HEA hydrogenation thermodynamics, and demonstrates the power of interpretable ML in navigating complex compositional spaces.

Collectively, kernel methods—including SVM and GPR—have emerged as indispensable tools for hydrogen storage materials discovery. SVM and SVR provide strong predictive baselines, particularly for structured and moderately sized datasets, while GPR offers uncertainty quantification and probabilistic insight essential for guiding exploration in vast and data-scarce design spaces.

4.3. Neural Networks

Neural networks (NNs), inspired by the processing mechanism of biological neurons, provide a flexible and data-driven framework for modeling the complex, nonlinear relationships that govern hydrogen storage materials (HSMs). Their multilayer architecture, comprising input, hidden, and output layers, allows for automatic learning from data without predefined descriptors. With the ability to approximate arbitrary functions in high-dimensional spaces, NNs have been widely applied to predict thermodynamic properties, simulate atomic dynamics, and even generate novel hydrogen storage candidates. Depending on their structure and purpose, NNs can be categorized into several types, including feedforward neural networks (FNNs), backpropagation neural networks (BPNNs), neural network potentials (NNPs), graph neural networks (GNNs), and generative models. Each of these architectures targets a specific level of chemical representation, from atomic-scale potentials to crystal-structure-based screening.

4.3.1. Feedforward and Backpropagation Neural Networks

FNNs propagate information in a unidirectional manner from inputs to outputs, often optimized via algorithms like gradient descent. Initial efforts leveraged relatively shallow feedforward architectures to correlate alloy composition with hydrogen-related properties. For instance, Malinova and Guo [73] used a two-layer feedforward ANN (Levenberg–Marquardt with Bayesian regularization) to relate Mg-alloy composition to hydrogen-storage and dehydrogenation behavior. Rahnama et al. [74] trained a multiclass neural-network classifier to screen hydrides, highlighting complex hydrides and Mg-based alloys as leading classes. Tian et al. [75] employed a three-layer back-propagation neural network to forecast cycle-life degradation of Mg-based hydrogen-storage alloys.

In an effort to model the complex interplay between composition, processing, and hydrogen desorption performance of Mg alloys, Dong et al. [34] further adopted a multilayer perceptron (MLP) regression model to predict the maximum hydrogen desorption capacity (De_{max}) based on a curated dataset comprising 113 experimentally characterized alloys. After extensive feature engineering and sequential forward selection, the MLP model achieved a high test-set R^2 of 0.922. SHAP analysis identified key controlling descriptors such as lattice parameter deviation (δLP), dehydrogenation temperature (T_{de}), and ball milling speed. Particularly, higher T_{de} and moderate spinning conditions (e.g., $MS_s < 19$ m s^{-1}) were found to positively impact De_{max}. The model successfully guided the discovery of promising low-temperature Mg-based desorption alloys, such as 99 Mg−1La and 95 Mg−1Ni−4La, both exhibiting predicted De_{max} values >5.6 wt.% at 275 °C, thus showcasing the potential of MLPs for quantitatively capturing nonlinear processing−performance relations in hydrogen storage materials.

Beyond electrochemical applications, Zhao et al. [53] integrated BPNN with finite element simulations and genetic algorithms (GA) to optimize the thermal−mechanical lifespan of hydride-based metal reactors (MHRs), identifying critical geometric factors that govern creep−fatigue degradation (CFL). Structurally, key geometric parameters such as cooling channel diameter and interlayer spacing were identified via Taguchi-designed experiments and sensitivity analysis as the

most influential on thermal stress distribution. BPNN, trained on finite element method (FEM)-derived datasets, demonstrated high accuracy ($R^2 > 0.98$) in predicting MHR lifespan, while GA leveraged this model to locate optimal structural configurations, achieving an 8% improvement in service life. The study reveals a clear inverse correlation between cooling channel size and MHR longevity, emphasizing the dominant role of temperature gradients in degradation processes.

As network depth increases, DNNs have demonstrated superior capability in learning complex thermodynamic patterns and generalizing across temperature regimes. For example, Kim et al. [32] developed a machine learning framework to model the PCT behavior of AB_2-type hydrogen storage alloys. Using curated experimental data and a novel curve-fitting function, they significantly expanded the training set, while van't Hoff-type thermodynamic extrapolations enabled prediction across temperatures. Three algorithms—random forest, k-nearest neighbors, and DNN—were benchmarked, with DNN showing superior accuracy ($R^2 \approx 0.93$). Chemically, alloys rich in Zr and V exhibited enhanced hydrogen uptake and lower equilibrium pressures, attributed to favorable thermodynamics. The approach provides a powerful route to accelerate composition design of solid-state hydrogen storage materials.

4.3.2. Graph Neural Networks

While NNPs primarily focus on reproducing atomic forces and simulating molecular dynamics with near-DFT fidelity, recent developments in GNNs have expanded the modeling capacity of neural networks to large-scale, compositionally complex, and periodic materials systems. By treating atoms as nodes and bonds as edges, GNNs offer a flexible framework for predicting a wide range of material properties, such as formation enthalpy, hydrogen desorption temperature, and phase stability, thereby enabling efficient high-throughput screening of hydrogen storage materials.

(1) SchNet for molecular-scale hydrides and anharmonicity modeling

SchNet is a continuous-filter convolutional GNN that learns from raw atomic coordinates. Pedrielli et al. [76] integrated a SchNet deep neural network with the stochastic self-consistent harmonic approximation (SSCHA) to investigate anharmonic effects in Mg_nH_{2n} nanoparticles ($n \leq 10$), addressing challenges related to undercoordinated surface atoms in HSMs. Their SchNet−SSCHA framework, trained on ab initio forces and energies, accurately reproduced temperature-dependent free energies, enthalpies, and entropies. The model revealed that anharmonicity lowers hydrogen desorption temperatures by up to 10%—e.g., for Mg_6H_{12}, T_d decreased from ~607 K (harmonic) to ~559 K. Furthermore, it captured thermal expansion in Mg–H bonds, with bond lengths exhibiting a nearly linear temperature dependence and variations of ~2% over 300 K. Beyond accurate modeling, this approach enables high-throughput screening of candidate HSMs, where SchNet, trained on DFT data, can rapidly predict key descriptors such as formation enthalpies, equilibrium pressures, and T_d values for large-scale materials discovery.

(2) CGCNN for periodic systems and high-throughput alloy screening

Crystal Graph Convolutional Neural Networks (CGCNNs) have been widely applied to periodic materials due to their crystal-symmetry-aware structure. Witman et al. [54] employed CGCNNs to accelerate the screening of hydrogen storage alloys by predicting DFT-relaxed formation energies from idealized lattice representations with minimal training data. By modeling both metal alloy lattices and hydrogen interstitial occupation, the CGCNN approach accurately reproduced configurational energetics and phase behavior in Pd-based hydrides (e.g., PdH_x, $Pd_{0.91}Rh_{0.09}H_x$), achieving mean absolute errors below 1 meV atom^{-1} (Fig. 5a). This enabled rapid construction of PCT curves, capturing the effects of alloying on plateau pressure and saturation capacity. The model revealed that Rh substitution raises plateau pressure, while Ag lowers it, consistent with experimental trends. The study demonstrates that even low-complexity GNNs can outperform traditional cluster expansion in extrapolative accuracy and data efficiency, offering a scalable route for designing compositionally complex hydride systems with tailored thermodynamics. Cheng and co-workers [55] implemented a CGCNN model to accelerate the discovery of ternary metal borohydrides for hydrogen storage applications. Using La–B–H as a prototype system, they trained a first-generation CGCNN on 59,028 DFT-calculated structures from the Materials Project to predict formation energies, followed by iterative refinement using a second-generation CGCNN trained on La-specific DFT data. This hierarchical learning scheme enabled the high-throughput screening of over 350,000 hypothetical structures, identifying 31 La–B–H compounds with formation energies within 100 meV atom^{-1} of the convex hull. Notably, the model recovered known phases such as $La(BH_4)_3$ and predicted a stable $La(BH)_{12}$ polymorph. By substituting La with other metals (e.g., Mg, Ca, Al, Sn), the study uncovered composition−structure relationships governed by cation size, revealing new trends in the stability of closo-dodecaborates. This work exemplifies how domain-specific neural networks can capture subtle energetic differences in complex hydrides and significantly reduce the search space in multi-element systems.

(3) MEGNet and transfer learning for enthalpy prediction across alloy families

The Materials Graph Network (MEGNet) architecture supports transfer learning and flexible encoding of both elemental and structural features. Batalović et al. [56] utilized a GNN based on the MEGNet architecture to predict the enthalpy of hydride formation (ΔH_{hyd}) directly from the crystal structure of intermetallic compounds (Fig. 5b). Leveraging transfer learning on pre-trained elemental embeddings and a curated dataset of 122 experimental ΔH_{hyd} values, the model encoded atomic and bonding environments via graph representations and achieved an MAE of 9.1 kJ mol^{-1} H_2. Unlike composition-only models, this structure-informed GNN distinguished

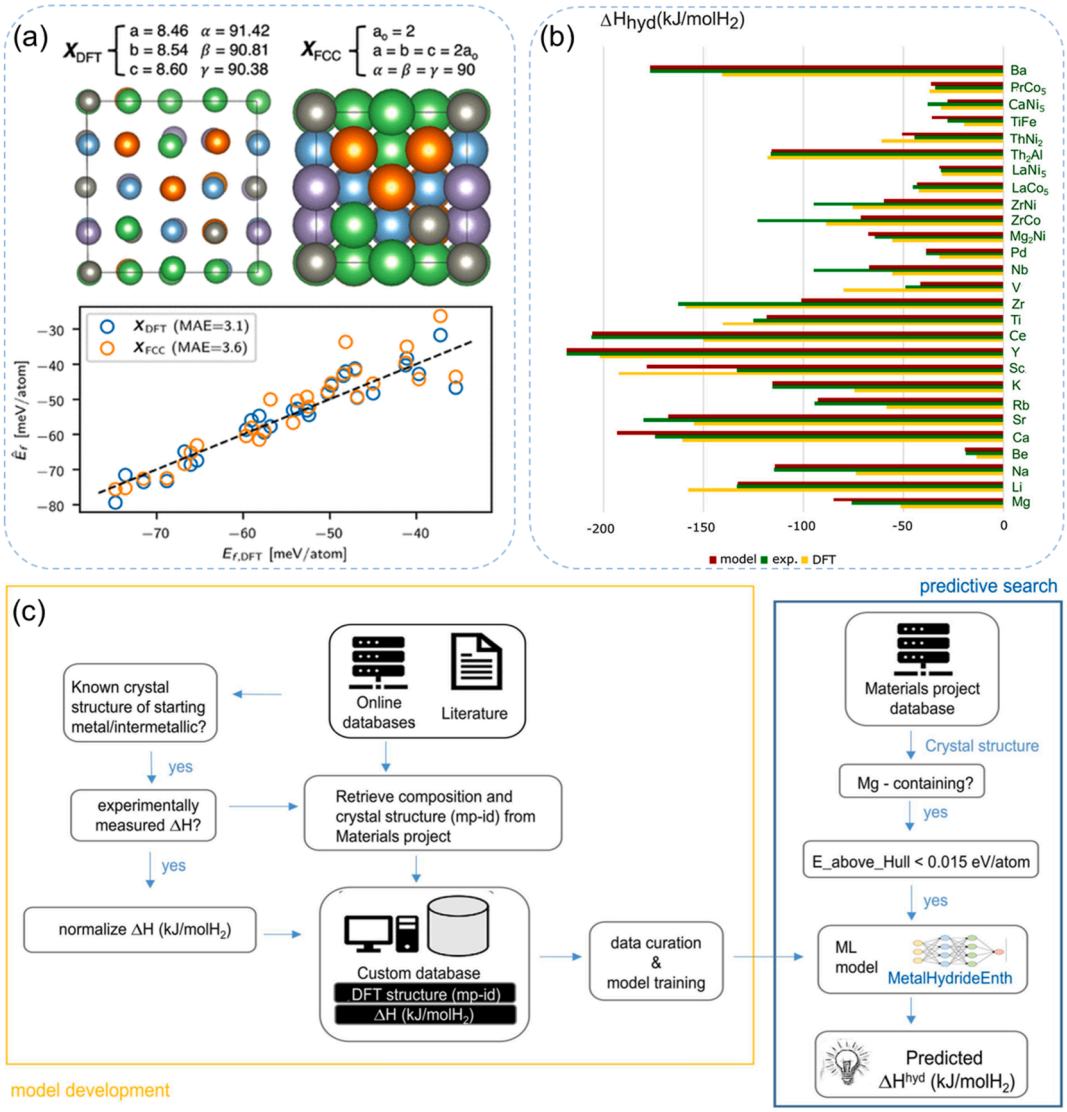

Fig. 5. (a) CGCNN model by Witman et al. [54] predicts DFT-relaxed formation energies of hydride-forming alloys using idealized lattice inputs, achieving low MAEs and capturing key phase behaviors in Pd-based systems. (b) MEGNet-based GNN by Batalović et al. [56] predicts hydride formation enthalpy (ΔH_{hyd}) from crystal structures with good agreement to experimental and DFT values, enabling reliable screening of Mg-containing alloys. (c) Workflow of the MetalHydrideEnth model for predicting ΔH_{hyd} in Mg-based intermetallics using crystal structures from Materials Project, thermodynamic filtering, and GNN-based predictions [57].

polymorphs and accurately captured structure–property relations, as validated by predictions on Mg-containing ternary alloys (e.g., Mg_3MnNi_2, Mg_2Ni). Materials with predicted ΔH_{hyd} close to −39.2 kJ mol^{-1} H_2—corresponding to 1 bar H_2 at 300 K—were identified as promising candidates for ambient-condition hydrogen storage. Notably, the model showed improved performance (MAE = 5.5 kJ mol^{-1} H_2) within Mg–Ni–M systems, supporting its applicability in fast and accurate screening of hydride-forming alloys.

A follow-up study [57] used a refined MetalHydrideEnth model to screen 636 Mg-containing compounds, targeting hydrogen storage and energy conversion applications (Fig. 5c). Adapted from the MEGNet architecture, this graph neural network predicts hydride formation enthalpy (ΔH_{hyd}) directly from DFT-optimized compositions and crystal structures, achieving a mean absolute error of 5.5 kJ mol^{-1} H_2 for Mg-based systems. To identify viable candidates, the study applied multiple filters: thermodynamic stability ($E_{above\text{-}hull} \leq 0.015$ eV $atom^{-1}$), exclusion of toxic or radioactive species, and an enthalpy window appropriate for near-ambient hydrogen storage (−48.3 to −30.1 kJ mol^{-1} H_2). As a result, 32 promising materials were discovered, including $MgBe_{13}$ and $YMgNi_4$. Importantly, the study revealed a strong correlation between ΔH_{hyd} and volume per atom (14.1–15.7 $Å^3$), suggesting volume as a useful design descriptor. Additionally, several Mg-based intermetallics (e.g., Mg_3MnNi_2, $LaMgNi_4$) were proposed as conversion-type anode materials for Li-ion batteries, with favorable equilibrium potentials ranging from 0.1 to 0.5 V vs Li^+/Li^0. This work highlights the capacity of GNN-based frameworks

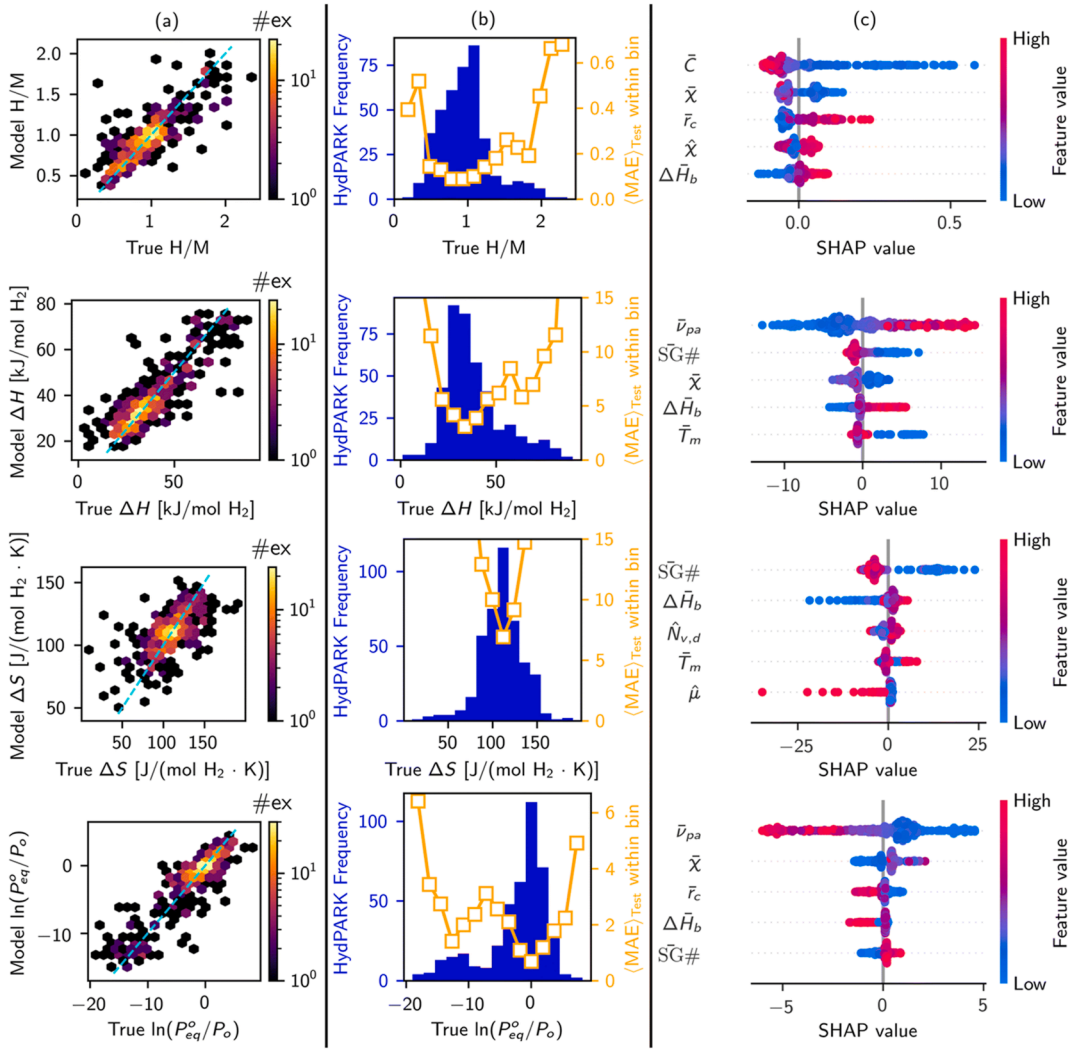

Fig. 6. (a) Parity plots from 10-fold cross-validation for models predicting H/M, ΔH, ΔS, and ln P^{o}_{eq}, with color indicating training data density. (b) Histograms of training data (blue) with test set MAEs overlaid (orange) by property. (c) SHapley Additive exPlanation (SHAP) analysis highlights key features driving predictions, including atomic volume (ν_{pa}), electronegativity (χ), valence electron counts (N_v, $N_{v,d}$), and binary hydride enthalpy (ΔH_b) [71].

to uncover structure–property relationships in chemically diverse systems, offering a scalable strategy for materials discovery across energy domains.

(4) Graph Convolutional Networks (GCNs) for descriptor-free modeling

Graph Convolutional Networks (GCNs) extend traditional GNNs by applying convolution operations over graph-structured data, enabling efficient message passing between atoms. In the context of HSMs, GCNs construct crystal graphs with atoms as nodes and bonds as edges, allowing direct learning of bulk properties such as hydride formation enthalpy and hydrogen storage capacity. GCNs are particularly well-suited for periodic systems like AB_5-type intermetallics, as they can automatically extract relevant chemical features (e.g., bond types, coordination numbers), thereby reducing the need for handcrafted descriptors. However, their applicability to amorphous HSMs is limited due to the lack of well-defined long-range order. Moreover, model performance is highly sensitive to graph construction parameters, including bond cutoff distances and neighbor definitions. Witman et al. [71] employed a GCN-based model to predict the formation enthalpy of transition metal hydrides, leveraging crystal graph descriptors (Fig. 6). The model achieved an MAE of 5.4 kJ mol^{-1} H_2 for Mg-based hydrides, outperforming traditional machine learning models. This capability enabled rapid screening of Mg-Ni-M ternary alloys, identifying compositions with desorption temperatures near 300 °C, suitable for fuel cell applications. Besides, GCNs have been used to model phase transformations in HSMs, such as the BCC to face-centered cubic (FCC) transition in Ti-Zr-Hf-Mo-Nb high-entropy hydrides.

Collectively, these neural network architectures—ranging from graph-based predictors to generative models and engineering surrogates—demonstrate the versatility of deep learning in tackling challenges across scales in hydrogen storage research. Their combined capabilities not only facilitate property prediction and materials discovery but also empower system-level optimization, paving the way for data-driven HSM innovation.

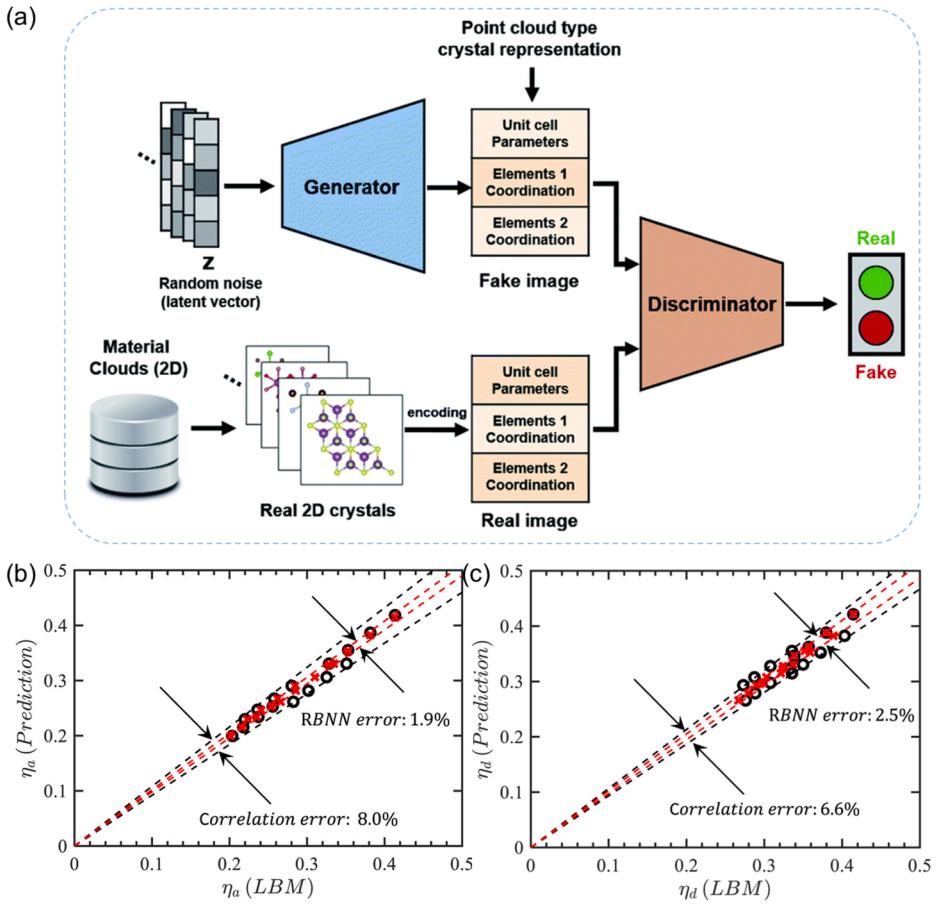

Fig. 7. (a) Schematic of the deep generative model architecture developed for 2D crystal systems [58]. Relative error comparison between empirical models and the R*BNN* approach for predicting hydrogen adsorption (b) and desorption (c) performance [59].

4.3.3. Generative Neural Networks

While GCNs offer a powerful route to predict material properties directly from graph-based crystal structures, they remain constrained to existing structural databases and known topologies. To explore materials beyond conventional design spaces, generative neural networks such as Generative adversarial networks (GANs) have emerged as a compelling tool, enabling the inverse design of entirely new crystal architectures with targeted hydrogen storage properties.

Lee et al. [58] developed a deep learning-guided inverse design framework based on a GAN to explore two-dimensional magnesium hydride (Mg–H) as a hydrogen storage medium (Fig. 7a). Using point cloud–based crystal representations, their GAN model was trained on 134,000 2D crystal structures derived from known monolayers, enabling the generation of 500 new Mg–H nanosheet candidates. After DFT-based screening of formation energies and phonon stabilities, a novel MgH_2 polymorph with $P\,\overline{4}\,m2$ symmetry was discovered, exhibiting superior thermodynamic and dynamic stability. Upon Li-decoration, this 2D phase achieved a theoretical hydrogen gravimetric capacity of 6 wt.% and average adsorption energy of −0.105 eV per H_2, attributed to charge-induced polarization effects. This work highlights the power of generative neural models in uncovering unexplored structural motifs beyond classical substitution-based approaches, providing promising candidates with enhanced storage performance.

4.3.4. Radial Basis Neural Network

Beyond crystal structure discovery, another frontier lies in optimizing the macroscopic performance of hydrogen storage systems. In this regard, surrogate neural models like radial basis neural networks (R*BNN*s) have proven useful for engineering-scale simulations**,** bridging atomistic insights with device-level efficiency optimization. Wang et al. [59] established an RBNN model to predict hydrogen adsorption and desorption efficiency in cylindrical $LaNi_5$-based metal hydride beds, integrating geometric (aspect ratio) and thermal (chamber conductivity) descriptors as inputs. From a chemical engineering perspective, hydrogen uptake and release were governed by heat conduction in the porous hydride medium, where reaction kinetics were sensitive to transient temperature distributions. The efficiency indicator—defined as the ratio of ideal isothermal reaction time to practical time—served as the target output. The RBNN model, trained on data from lattice Boltzmann simulations, achieved prediction errors below 2.5%, significantly outperforming traditional empirical correlations (Fig. 7b-c). Chemically, the study highlighted that increasing chamber thermal conductivity and optimizing bed

geometry enhanced storage efficiency, with thermal properties having a more pronounced effect than geometry. These results underscore the value of ML-based surrogates for engineering design in thermally limited hydride systems.

In summary, neural network models have exhibited remarkable versatility across different stages of hydrogen storage materials research. From graph convolutional frameworks that automate structure–property prediction, to generative models that propose chemically plausible yet novel hydride phases, and finally to surrogate models capable of bridging molecular behavior with device-level performance, these approaches collectively form a powerful ML toolkit. Their success underscores a paradigm shift—from empiricism and trial-and-error experimentation toward data-driven materials discovery and optimization—that is increasingly essential for accelerating hydrogen-based energy solutions.

4.4. Ensemble Models

Ensemble learning methods can be broadly categorized into three main strategies: bagging, boosting, and stacking/ensemble-of-ensemble (EoE) models. These approaches differ in how they aggregate base learners, and each offers distinct advantages depending on the complexity of the dataset, target property, and desired interpretability. In hydrogen storage applications, these strategies have been variously applied—from modeling simple composition–property relationships to capturing intricate interactions in catalyst design and reactor optimization.

4.4.1. Bagging Models: Random Forest and Extra Trees

Bagging-based ensemble learning, especially Random Forest (RF) and Extra Trees Regression (ETR), is widely adopted in hydrogen storage research for its robustness to noisy, sparse, and nonlinear datasets. These models operate by aggregating multiple decision trees trained on random subsets of data and features, thereby reducing overfitting and enhancing predictive stability. Their effectiveness spans from composition–property prediction to large-scale material screening and even device-level optimization.

(1) Composition–Property Modeling

One major application of bagging models lies in predicting thermodynamic and structural properties directly from material composition. Kronberg et al. [77] applied an RF model to investigate hydrogen absorption properties and key influencing factors in nitrogen-doped defective carbon materials. Similarly, Kusdhany et al. [78] extended this to carbon materials with different porous morphologies and surface functionalities, revealing synergistic effects from temperature and structure. By quantitatively linking alloy chemistry to hydrogen sorption behavior, Suwarno et al. [25] applied supervised machine learning models—particularly random forest regressors—to a curated dataset of 314 AB_2-type intermetallic hydrides. Using elemental composition as input features, the models predicted thermodynamic and structural descriptors including hydride formation enthalpy (ΔH), C14 Laves phase abundance, and hydrogen storage capacity (wt.%). The analysis revealed that Ni destabilizes hydride formation, Cr stabilizes the C14 phase via electronic and geometric effects, and Mn enhances hydrogen uptake by increasing interstitial site volume. The optimal performance ($R^2 = 0.688$ for capacity) was achieved for substoichiometric B/A ratios (1.9–2.0), highlighting compositional regimes favorable for reversible hydrogen storage. These studies illustrate the utility of bagging algorithms in capturing nuanced composition–property relationships, particularly when feature interactions are complex or nonlinear.

(2) High-Throughput Screening

When applied to large datasets, bagging models also enable accelerated screening of viable hydrogen storage candidates. Nations et al. [61] employed a RF algorithm to screen metal hydrides for hydrogen storage, using features such as electronegativity and density. The evaluation combined hydrogen weight fraction with a formation energy scaling factor favoring moderately negative energies. Promising candidates included compounds with alkali metals, boron, and transition metals, which balanced stability and hydrogen uptake. Using an RF regression model trained on DOE hydride databases, Hattrick-Simpers et al. [79,80] identified >6000 candidate hydrogen storage alloys, and, by applying techno-economic and structural constraints, ultimately highlighted Fe–Mn–Ti–X quaternaries as cost-effective, high-pressure hydrogen compressor materials with favorable enthalpy (~22 kJ mol^{-1}) and predicted stability. To accelerate the identification of viable solid-state hydrogen storage materials, Verma et al. [62] developed a dual-machine learning framework, HydrogEn storAge propeRty predictor (HEART), integrating two tree-based ETR models—hydrogen storage capacity (HYST) for hydrogen storage capacity (H_2 wt.%) and enthalpy of hydride formation (THOR) for the enthalpy of hydride formation (ΔH). Using features derived from elemental properties, alloy composition, DFT-calculated metal–hydrogen interactions, and absorption temperature, they screened ~6.4 million multicomponent metal alloys. Alloys achieving H_2 wt.% > 2.5 at 300 K and $\Delta H < 60$ kJ mol^{-1} H_2 were prioritized, yielding 6480 promising candidates. Notably, Mg-, V-, and Ti-based alloys doped with elements such as Sc, Cr, and Mn showed favorable thermodynamics and high capacity. Collectively, these works showcase how bagging algorithms scale well with data size and feature dimensionality, offering interpretable, fast, and accurate tools for material down-selection.

(3) Device-Level Optimization

Beyond molecular-scale material prediction, bagging models have also been extended to engineering-scale optimization. Peng et al. [63] systematically explored the influence of dimensionless fin geometries on the performance of adsorption-based hydrogen storage in $LaNi_5$-filled beds, integrating computational fluid dynamics with machine learning

to optimize storage efficiency. By mapping the relationship between fin thickness, height, width, and hydrogen storage amount across different adsorption durations, ensemble-based regression models and decision tree regressors were trained to predict hydrogen uptake. Bayesian hyperparameter optimization was employed to improve prediction accuracy, yielding excellent performance ($R^2 = 1.00$). To determine the optimal fin configurations at various durations, the learned models were coupled with a genetic algorithm. The results revealed a non-monotonic dependence of optimal fin volume on adsorption duration, balancing between enhanced hydrogen fraction and reduced available adsorbent volume. At 400 s, the optimized configuration enhanced the hydrogen storage capacity by 12.8%, underscoring the value of ML-guided geometric design for adsorption systems. This study highlights the versatility of tree-based ensemble models in bridging scales—from material microstructure to macroscopic reactor design—especially when simulation outputs are expensive and highly nonlinear.

In summary, bagging models such as Random Forest and Extra Trees Regression have demonstrated consistent utility in the hydrogen storage domain—offering strong generalization, rapid convergence, and clear feature importance metrics. From predicting composition–property trends to guiding experimental synthesis and optimizing adsorption geometries, their flexibility and resilience to data noise make them well-suited for both early-stage exploration and engineering refinement.

4.4.2. Boosting Models: GBDT, XGBoost

Boosting algorithms construct sequential weak learners—typically decision trees—where each new model focuses on correcting the residuals of the previous one. This process, though computationally more intensive, allows boosting methods such as XGBoost and GBDT to achieve state-of-the-art accuracy in non-linear regression tasks. Their high interpretability (e.g., via SHAP analysis), strong performance on small-to-medium datasets, and flexibility in incorporating domain-specific descriptors make them particularly attractive for hydrogen storage research. Recent studies have applied boosting models to predict composition–property relationships, model thermodynamics, and accelerate high-throughput material screening workflows.

(1) Composition–Property Modeling in Binary/Ternary Alloys

Boosting methods have proven highly effective for modeling hydrogen-related properties from compositional and physicochemical features, especially in relatively simple alloy systems. In a recent study, Lu et al. [64] employed the XGBoost algorithm to predict the dehydrogenation enthalpy of V–Ti–Cr–Fe alloys using a curated dataset of 33 samples with physicochemical descriptors including lattice constants, valence electron concentration, and bulk modulus. Among several models tested—linear regression, SVM, random forest, stacking—the XGBoost model demonstrated superior performance with a mean absolute error below 9 kJ mol^{-1} and Pearson correlation above 0.6. Feature selection further improved generalization ability by eliminating weakly correlated variables. This work highlights the robustness of boosting-based ensemble methods in capturing nonlinear composition–property relationships in complex alloy systems. Rahnama et al. [65] employed GBDT to predict hydrogen storage capacities of metal hydrides using a US Department of Energy database. The study focused on hydrogen weight percent (wt.%) as the key metric, identifying material class, temperature, and heat of formation as critical factors. Complex hydrides and Mg-based alloys emerged as top candidates, with complex hydrides achieving hydrogen weight percents up to 10.6 wt.%. The GBDT model demonstrated superior predictive accuracy ($R^2 = 0.83$) compared to other regression methods. In Zhou et al.'s work [33], GBDT models trained on both implicit and explicit features revealed the dominant influence of the Mean Ionic Character and Fe content on capacity. SHAP analysis further quantified these relationships, enabling interpretable predictions and guiding the successful design of $Ti_{0.9}Zr_{0.12}Mn_{1.2}Cr_{0.55}(VFe)_{0.25}$ with a validated capacity of 1.90 wt.% H_2 (127.3 kg m^{-3}), offering both high performance and cost-effectiveness for proton exchange membrane fuel cell (PEMFC) applications. These studies demonstrate the power of boosting in distilling nonlinear trends and design rules from experimental databases, even when data volume is limited.

(2) Thermodynamic Tuning and Mechanism Insight

Beyond capacity estimation, boosting regressors have enabled detailed thermodynamic modeling and mechanism exploration. Dong et al. [34] employed Gradient Boosting Regression (GBR) to model the maximum hydrogen absorption capacity (Ab_{max}) of Mg-based alloys, leveraging a filtered dataset of 183 experimentally measured samples. After systematic feature reduction via SHAP and Pearson correlation analysis, the GBR model yielded a test-set R^2 of 0.947, outperforming alternative methods such as SVR and Random Forest. Interpretability via SHAP unveiled that low alloy density and elevated hydrogenation temperature (>213 °C) were positively correlated with Ab_{max}, alongside processing features like prolonged ball milling and optimized melt-spinning rates. Guided by the model, high-throughput predictions were performed on 800 binary and 735 ternary Mg alloys, leading to the identification of 96 Mg–4Sm and 95 Mg–1Ni–4Sm as leading candidates with both high absorption and release capacities at sub-300 °C conditions. This study highlights how ensemble models like GBR can integrate compositional, thermodynamic, and kinetic parameters into robust predictors for accelerated alloy discovery. Jiang et al. [66] developed a machine learning-assisted screening strategy to modulate the hydrogen desorption temperatures (T_d) of Mg-based hydrides by alloying MgH_2 with various elements (Me) to form $Mg_{1-x}Me_xH_2$. By integrating DFT calculations with GBR, the study identified Me electronegativity and alloying ratio as key descriptors governing T_d. SHAP analysis revealed that Me elements with higher electronegativity and optimal doping levels weaken Mg–H bonding and facilitate hydrogen

diffusion. Screening 132 compositions led to the selection of $Mg_{0.875}Li_{0.125}H_2$ and $Mg_{0.875}Be_{0.125}H_2$, which offer desorption at 400 K and 500 K, respectively. These were embedded in NLi_4-decorated boron-doped graphene/MgH_2 heterojunctions to achieve stepwise hydrogen release at ambient and elevated temperatures. The study underscores the effectiveness of combining DFT and interpretable ML in tuning thermodynamics of complex hydride systems for practical hydrogen storage applications (Fig. 8). Witman et al. [67] employed a GBDT algorithm to identify promising hydrogen storage materials from a limited dataset. They focused on intermetallic hydrides, using features derived solely from the elemental composition of the intermetallic alloy to predict the equilibrium hydrogen pressure at room temperature, a key thermodynamic property for practical hydrogen storage applications. The GBDT model identified the volume per atom (ν_{pa}) as the most important descriptor, revealing a strong correlation between νpa and the stability of the hydride phase. This simple structure−property relationship allowed them to predict that UNi_5, an intermetallic with a lower ν_{pa} value, would form a low-stability hydride suitable for high-pressure hydrogen storage, a conclusion supported by density functional theory calculations. These studies highlight how boosting models, when coupled with interpretability tools and domain-specific descriptors, can unveil actionable design principles for thermodynamic tuning and material modification.

(3) High-Throughput Design and Multi-Objective Optimization

With access to curated databases and fast predictions, boosting methods also underpin large-scale materials discovery efforts. Witman et al. [71] built an explainable machine learning workflow based on GBR, extending prior models trained on the ML-HydPARK database to predict thermodynamic properties of high-entropy hydrides, including desorption enthalpy (ΔH), entropy (ΔS), hydrogen-to-metal ratio (H/M), and equilibrium pressure. The input features were constructed from Magpie descriptors augmented with binary hydride formation enthalpies to reflect domain-specific chemistry. This enabled high-throughput screening of over 17,000 equimolar HEA compositions. A multi-objective Pareto front was identified to simultaneously optimize hydrogen capacity, thermodynamic favorability, and material cost. Selected candidates such as MgTiVZrNbHf, MgTiVCrNb, and MgAlTiVCr were experimentally synthesized and validated, showing strong consistency with both machine learning predictions and DFT calculations. Such workflows represent a paradigm shift in materials discovery, enabling data-driven identification of high-performing hydrogen storage materials across vast chemical design spaces, with strong consistency between ML predictions and experimental synthesis.

4.4.3. Advanced Architectures: Stacking & EoE Models

While single-level ensemble models like Random Forest and Gradient Boosting have already proven valuable in property prediction tasks, more sophisticated architectures—such as stacking and EoE frameworks—have emerged to further

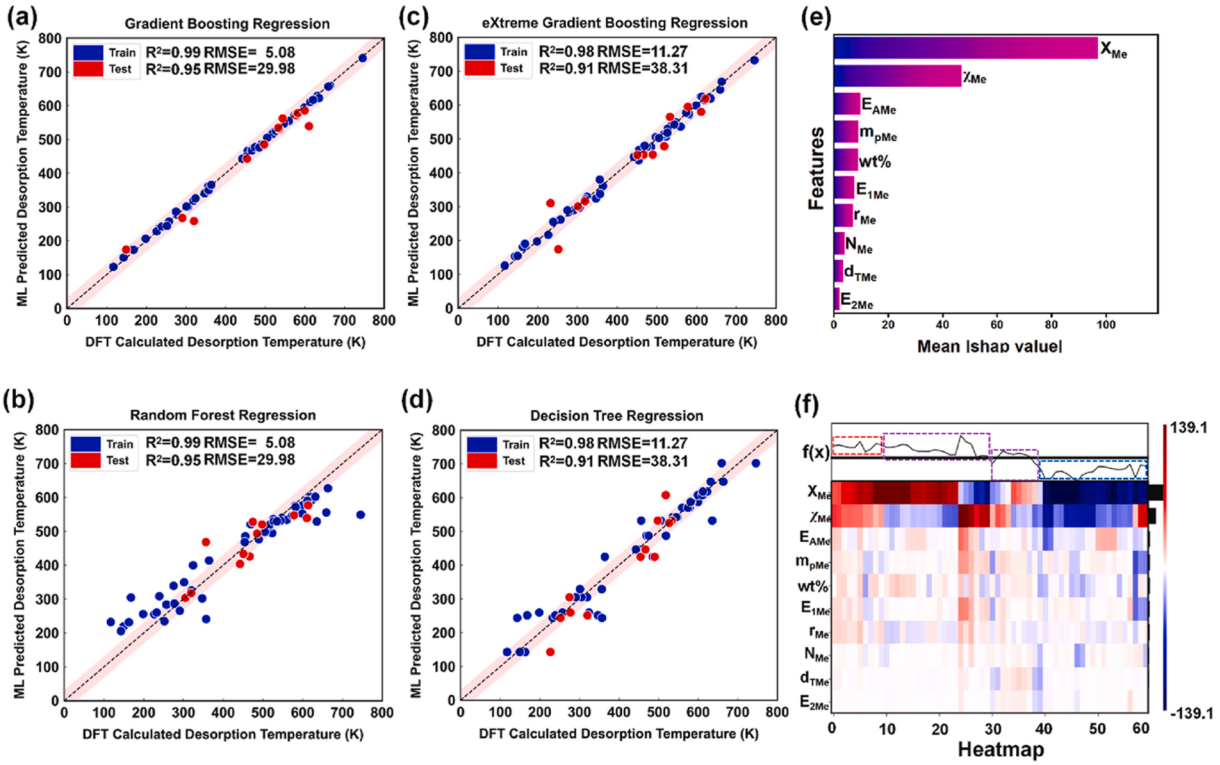

Fig. 8. Comparison between DFT-computed and ML-predicted decomposition temperatures (T_d) using (a) GBR, (b) random forest regression (RFR), (c) XGBoost, and (d) Decision Tree Regressor (DTR) models. (e) Average absolute SHAP values indicating feature importance. (f) SHAP value heatmap showing individual feature contributions; color scale denotes SHAP magnitude [66].

improve model accuracy, robustness, and interpretability. These hierarchical approaches enable the integration of multiple base learners or even entire ensemble models, capturing diverse data representations and complex non-linearities that often arise in hydrogen storage materials.

(1) Stacking-Based Alloy Design

Stacked ensemble learning has been successfully applied to predict the hydrogen storage capacity of multi-component alloys. For instance, a stacking model combining Adaboost-optimized Random Forest, XGBoost, and support vector regression was constructed to estimate the storage capacity of V–Ti–Cr–Fe alloys, based on 19 physicochemical descriptors that capture compositional, structural, and electronic characteristics [30]. The model achieved excellent performance (Mean Square Error, MSE = 0.187, $R^2 = 0.783$), with feature importance analysis highlighting atomic radius, valence electron concentration (e/a), and Z/r^3 as key descriptors governing hydrogen uptake. Furthermore, a genetic algorithm was integrated with the model to guide compositional optimization, yielding candidate alloys with predicted capacities exceeding 3.45 wt.%. Experimental validation showed relative errors below 1.5%, underscoring the potential of stacking models in accelerating data-driven alloy design.

(2) Boosting Models for HEA Screening

Gradient boosting methods have also found wide application in screening HEA hydrides, particularly for mapping compositional trends to thermodynamic outputs such as desorption enthalpy (ΔH). One study constructed a Gradient Boosting Trees (GBT) model using weighted elemental properties (e.g., electronegativity, binary hydride enthalpies) to predict the hydrogen desorption characteristics of Fe-substituted $(TiVNb)_{75}Cr_{25-x}Fe_x$ alloys [68]. Results revealed a linear decrease in ΔH with increasing Fe content—up to 10 kJ mol^{-1} H_2—due to the higher electronegativity and smaller atomic radius of Fe, which weakens H–metal bonding and reduces interstitial volume. This trade-off between thermodynamic destabilization and capacity was further confirmed via DFT and experimental analysis, illustrating how boosting models can uncover chemically intuitive design rules.

(3) Reactor-Level Modeling via Tree Ensembles

Beyond material-level predictions, ensemble methods have also been applied to system-level behavior modeling. Tiwari et al. [69] integrated experimental insights with supervised machine learning to predict the hydrogen absorption and desorption behavior of a $LaNi_5$-based metal hydride reactor embedded with a helical cooling tube. By systematically varying supply pressure, coolant flow rate, and inlet temperature, they compiled over 1700 data points which were used to train and evaluate multiple regression models. Tree-based models—particularly XGBoost and Random Forest—demonstrated superior predictive accuracy ($R^2 > 0.95$) compared to linear and SVM regressors in estimating both hydrogen uptake and bed temperature. Their analysis revealed that higher hydrogen supply pressure, lower heat transfer fluid (HTF) temperature, and increased HTF flow significantly enhance absorption rates, while elevated HTF temperatures favor desorption. The study underscores the efficacy of experimental-data-driven ML models for reactor behavior forecasting and design optimization, culminating in a brute-force algorithm that tailored operating parameters for three practical scenarios: storage, heating, and fuel-cell integration.

(4) Ensemble-of-Ensemble for Catalyst–Hydride Systems

To tackle the multi-variable complexity in chemical processing systems, Ding et al. [70] proposed a two-layer EoE framework to predict hydrogen release capacity in $LiBH_4$-based catalyst mixtures. Using a dataset comprising 2,071 samples of uni-component catalyst systems, the model established correlations between 14 experimental variables—including catalyst identity, thermal parameters, and processing conditions—and hydrogen desorption amounts (wt.% H_2). In the first layer, multiple ensemble regressors (Random Forest, Extra Trees, Gradient Boosting, AdaBoost, Bagging, and Histogram-based Gradient Boosting) were independently trained to capture distinct non-linear mappings. Their outputs were then fed into a second-layer meta-learner, which performed optimized weighted averaging based on cross-validation, enabling the model to consolidate complementary strengths of each base learner while minimizing systematic bias. This two-layer hierarchy markedly improved robustness and predictive power, outperforming both linear regressors and single-level ensemble models ($R^2 = 0.888$, MSE = 1.144), and also generalizing effectively to unseen bi-component catalysts ($R^2 = 0.834$). Through comprehensive feature attribution analysis, this work quantitatively ranked the influence of experimental parameters on hydrogen desorption from $LiBH_4$, highlighting dehydrogenation temperature, catalyst identity, and gas atmosphere as the most critical factors. Chemically, synergistic dopants such as TiF_3 and AlF_3 were identified as effective in lowering desorption temperatures, whereas excessive catalyst loading impaired reversibility due to side reactions. These insights underscore the power of layered ensemble learning in unraveling complex structure–processing–property interdependencies in solid-state hydrogen storage.

More broadly, advanced ensemble architectures—particularly stacking and EoE—significantly extend the capabilities of conventional ensemble models by integrating heterogeneous learners across multiple layers. Their demonstrated success in composition optimization, thermodynamic screening, system-level modeling, and catalytic design illustrates their versatility and robustness in addressing the multifactorial challenges of hydrogen storage research. As the field increasingly embraces high-dimensional, multi-fidelity data, such architectures are poised to become central tools for accelerating the rational discovery of next-generation storage materials.

4.5. Clustering

Clustering, as a representative unsupervised learning approach, has been applied to hydrogen storage research to identify hidden patterns and group materials with similar physicochemical or thermodynamic properties. Common methods include k-means, hierarchical clustering, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [81], each suited to different dataset structures. K-means clustering partitions materials into k groups by minimizing intra-cluster variance. It has been used to correlate hydrogen storage performance with descriptors such as hydrogen content, pressure, and enthalpy. For example, Rahnama and Sridhar [82] applied k-means to a database of metal hydrides and identified three distinct clusters—A_2B, complex hydrides, and Mg-based alloys—all showing favorable storage characteristics due to shared compositional and thermodynamic features. Hierarchical clustering organizes materials into nested clusters, enabling multiscale classification without predefining the number of groups. This method is especially useful in revealing relationships among alloy systems or structural motifs, though it may be sensitive to noise in large datasets. DBSCAN identifies clusters based on data density, effectively handling irregularly shaped or noisy datasets common in nanostructured or composite materials. It is advantageous when no prior knowledge of cluster number exists, though parameter tuning can be challenging. These clustering techniques have been successfully deployed to explore diverse hydrogen storage material datasets [83]. In summary, clustering algorithms complement supervised learning by offering structure discovery in unlabeled datasets. Their integration into hydrogen storage studies helps classify material families, identify performance trends, and support data-driven screening strategies.

4.6. Neural Network Potentials

As a supervised-learning paradigm, neural network potentials (NNPs) learn a mapping from local atomic environments to energies, forces, and oftentimes stresses using DFT-quality labels. In this sense, NNPs extend the supervised toolbox from static property regressors to dynamical, force-bearing surrogates that enable long-time, large-scale simulations with near-DFT fidelity.

One of the earliest and most influential NNP frameworks is the Behler–Parrinello (BP) model, which employs symmetry functions combined with MLPs to ensure translational, rotational, and permutational invariance. It has found wide application in HSMs, particularly in modeling complex hydrides. For instance, Wang and Huang [84] constructed a BP-type ANN potential tailored for the Mg–H system, with the aim of overcoming the computational bottlenecks of DFT in large-scale nanocluster simulations. The ANN was trained on DFT data covering diverse Mg/H compositions and atomic

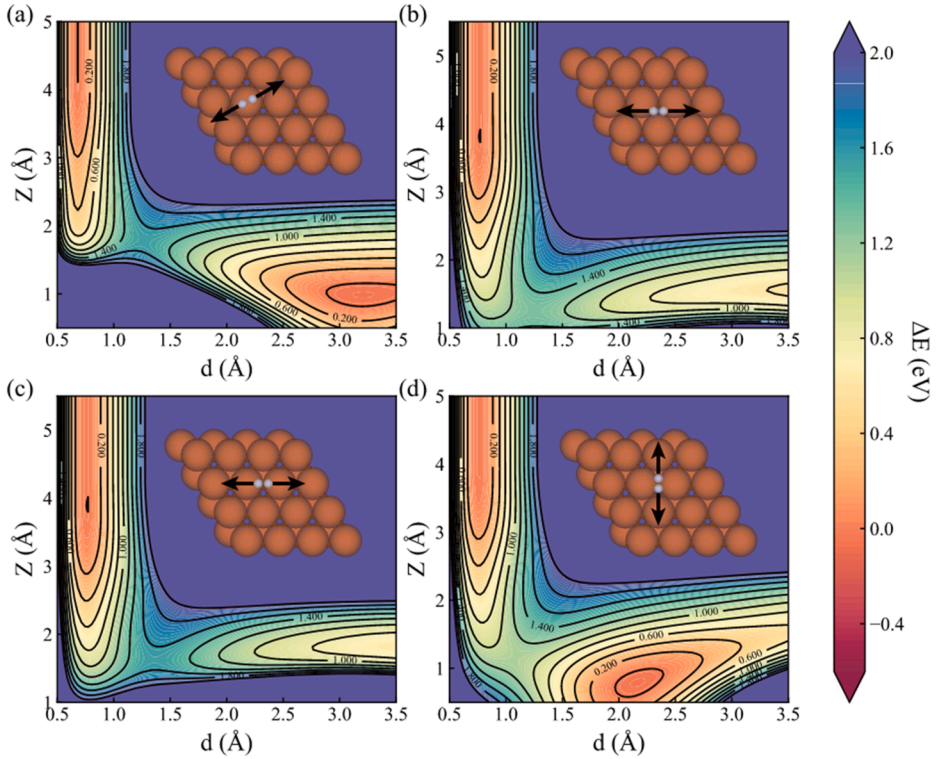

Fig. 9. Contour plots of the ML-derived potential energy surface (PES) for H_2 interacting with Mg(0001), plotted over vertical distance Z and H–H bond length d. (a–d) Illustrate different fixed adsorption sites and dissociation paths. Mg and H atoms are shown in orange and white; arrows indicate H atom movement directions [84].

environments, achieving excellent accuracy in reproducing energies and forces (energy mean absolute error (MAE): 18.23 meV atom^{-1}; force MAE: 95.48 meV Å^{-1}). When applied to molecular dynamics (MD) simulations (Fig. 9), the model revealed that hydrogen-unsaturated clusters (Mg_mH_n, with $m < 2n$) spontaneously undergo Mg/MgH_x phase separation and nanocrystalline ordering at elevated temperatures (600 K), while saturated clusters remain disordered. Furthermore, the diffusion coefficients extracted from ANN-MD trajectories exhibit Arrhenius-type temperature dependence between 400 and 700 K, consistent with experimental trends. This study highlights the capacity of neural network potentials to capture complex chemical bonding and dynamic processes in hydride materials beyond the reach of conventional ab initio methods.

Building upon the BP foundation, the Deep Potential (DP) model—developed by Wang et al. [85]—marked a significant advance in scalability and accuracy. The DP framework uses an end-to-end deep neural network to predict energies and forces from raw atomic coordinates, maintaining symmetry through carefully designed local descriptors. It forms the backbone of the Deep Potential Molecular Dynamics (DeePMD) scheme, which has been increasingly applied to thermodynamically complex materials.

In a notable application to HSMs, He et al. [60] employed DeePMD to model 15 binary Mg-based alloys. The models were trained on high-throughput DFT-derived datasets, using symmetry-preserving descriptors to represent local atomic environments through radial and angular information. The underlying deep neural networks—comprising five hidden layers—were optimized via stochastic gradient descent with adaptive loss weighting. As a result, the models achieved root mean square errors (RMSE) as low as 0.43 meV atom^{-1}, exemplified by the Mg–Na system, thereby demonstrating near-DFT accuracy. From a chemical standpoint, the models successfully captured dopant-specific effects: main group elements such as Na and K yielded highly accurate predictions due to their structural compatibility with Mg, while transition metals like Cr and Mn introduced greater prediction variance, likely due to the involvement of d-orbitals and limited training data coverage. Thanks to the incorporation of local symmetries and many-body correlations, DeePMD exhibited strong generalizability across varied alloying environments. For benchmarking purposes, kernel ridge regression (KRR) models were also developed using global descriptors such as Ewald sums and sine matrices. Although KRR offered significantly lower computational cost, its prediction accuracy—reflected in mean absolute errors (MAEs) ranging from ~4.8 to 127 meV atom^{-1}—was generally inferior to DeePMD, particularly for chemically complex systems. These findings underscore the superior capability of deep neural network potentials in capturing intricate structure–property relationships across diverse material classes.

Sato et al. exemplify how machine-learning interatomic potentials unlock reactive, interface-mediated hydrogenation pathways under extreme conditions [86]. Using large-scale MLP-MD on CaH_2/H_2 at high pressure, they directly observe surface-melting–assisted absorption in which the solid transitions through a liquid CaH_4 intermediate, enabling rapid hydrogen uptake and lattice reorganization. Crucially, pressure lowers not only the $CaH_2(s) + H_2(l) \rightarrow CaH_4(s)$ hydrogenation enthalpy but also the formation enthalpy of the liquid polyhydride, thereby reducing kinetic barriers and making surface melting more favorable than bulk fusion. On this basis, the authors propose practical synthesis criteria—e.g., product melting temperature and pressure-dependent hydrogenation enthalpy—that can be computed alongside structure optimization algorithm to prioritize synthesizable candidates. Beyond high-pressure polyhydrides, this work showcases NNP-driven MD as a route to predictive reaction modeling at solid–fluid interfaces, complementing property prediction with synthesis- and kinetics-aware guidance for hydrogen-rich materials. Collectively, NNPs provide chemically faithful, scalable surrogates that unlock mechanistic exploration and kinetics beyond the reach of routine ab initio simulations.

5. Optimization Algorithms

Optimization algorithms, inspired by natural evolution or collective intelligence, have increasingly been integrated with ML models to tackle both forward prediction and inverse design problems in HSMs. These algorithms—such as genetic algorithms (GA), particle swarm optimization (PSO), and hybrid strategies—enable efficient navigation of high-dimensional search spaces, making them especially suitable for discovering optimal compositions or tuning parameters where brute-force methods are impractical. In the context of HSMs, optimization-enhanced ML frameworks serve two principal purposes: (i) refining prediction accuracy by tuning model parameters and (ii) solving inverse problems where material compositions or physical parameters must be inferred to achieve target performance metrics.

To efficiently predict hydrogen storage capacity in AB_2-type metal hydrides, Maghsoudy et al. [87] developed a hybrid machine learning framework (Fig. 10a) integrating least squares support vector machines (LSSVM) with three optimization algorithms: genetic algorithm (GA), particle swarm optimization (PSO), and a hybrid GA-PSO (HGAPSO). Using a curated dataset of 244 alloy compositions and their experimentally reported hydrogen capacities, the models mapped 22 alloying elements as input features to the target storage capacity. Among the models, the HGAPSO-LSSVM exhibited the highest accuracy (R^2 = 0.980, RMSE = 0.045 wt.%), outperforming standalone GA-LSSVM and PSO-LSSVM approaches. Sensitivity analysis revealed Sn, Co, and Ni as the most influential elements enhancing hydrogen uptake. This strategy offers a powerful route to accelerate the rational design of high-capacity hydrides by correlating elemental composition with hydrogen storage performance. Complementing the above optimization-enhanced SVM frameworks, Gheytanzadeh et al. [88] introduced a Gaussian process regression (GPR) approach to predict hydrogen absorption enthalpy (ΔH) in AB_2-type metal hydrides. Rather than tuning hyperparameters of a single model, they systematically

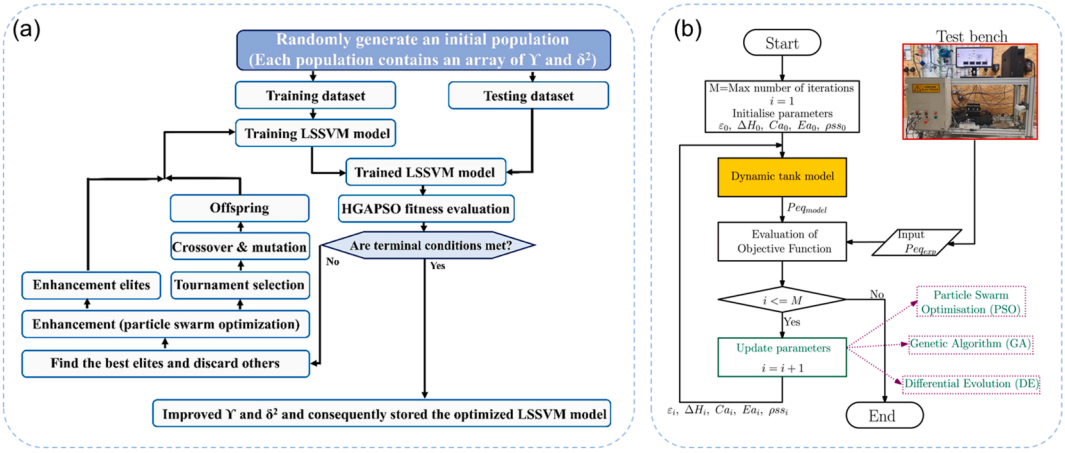

Fig. 10. (a) Workflow illustration of the HGAPSO-LSSVM hybrid prediction model. [87] (b) Schematic representation of the developed identification procedure [89].

evaluated four kernel functions—Exponential, Matern, Squared Exponential, and Rational Quadratic—within a GPR framework trained on 314 experimental data points. The Exponential kernel emerged as the best performer ($R^2 = 0.969$, RMSE = 2.50 kJ mol^{-1} H_2), outperforming the other three in both accuracy and stability. Sensitivity analysis further revealed that Ti, Zr, V, and Cr are the most influential elements in determining ΔH, aligning with known hydride thermodynamics. This study underscores the importance of kernel-level optimization in GPR-based ML models and offers a robust, uncertainty-aware alternative to traditional SVM for predicting thermodynamic properties of hydrogen storage alloys.

While forward prediction aids in evaluating candidate materials, optimization algorithms can also address inverse design problems—identifying input conditions (e.g., alloy compositions) that yield desirable hydrogen storage properties. In a representative study, Lu et al. [64] integrated a genetic algorithm with an XGBoost regression model to design multi-component hydrogen storage alloys targeting a dehydrogenation enthalpy of ~32.5 kJ mol^{-1}. By encoding compositional ratios as binary chromosomes, the GA iteratively selected, crossed, and mutated populations based on predicted enthalpy fitness. The optimized candidates were experimentally validated with < 4.5% error, demonstrating the efficacy of ML-guided GA frameworks in navigating high-dimensional composition spaces for property-driven materials design. Extending beyond intermetallic hydrides, Ghude and Chowdhury [90] demonstrated the utility of Bayesian optimization (BO) for maximizing the volumetric hydrogen deliverable capacity of metal–organic frameworks (MOFs). Starting from a pool of 98,000 real and hypothetical MOFs, the authors treated Grand Canonical Monte Carlo (GCMC)-simulated uptake as a black-box objective over a seven-dimensional structural descriptor space. By coupling a Gaussian-process surrogate with the expected-improvement acquisition function, BO identified the global optimum after evaluating only 27 structures (0.027% of the search space), outperforming random search, Covariance Matrix Adaptation-Evolution Strategy (CMA-ES), one-shot random-forest regression, and two swarm-intelligence variants (PSO/Evolutionary-PSO). This study establishes BO as a sample-efficient optimizer for crystalline porous materials and highlights its transferable value for high-throughput screening of alternative gases governed by dispersion interactions.

Beyond material-level optimization, Suárez and co-workers [89] extended the use of evolutionary algorithms to system-level modeling of hydrogen storage tanks (Fig. 10b). Specifically, they applied GA, PSO, and differential evolution (DE) to calibrate a grey-box dynamic model describing hydrogen absorption/desorption in a $LaNi_5$-based tank. Rather than predicting material properties, the algorithms optimized five key thermophysical parameters—porosity, enthalpy change (ΔH), saturation density, pre-exponential factor, and activation energy—by minimizing the relative error between simulated and experimental equilibrium pressures. Among the methods, GA achieved the lowest error (1.83%) with the fastest convergence (3 iterations), while PSO offered similar accuracy and DE provided robustness with minimal tuning. This study highlights how optimization techniques can compensate for limited experimental data by enabling high-fidelity, model-based parameter extraction in complex hydrogen storage systems.

Taken together, optimization algorithms serve as powerful complements to machine learning models across different length scales and modeling objectives. Whether improving predictive accuracy, guiding inverse design, or calibrating mechanistic models, these techniques expand the capabilities of ML frameworks and contribute to a more rational, efficient exploration of the vast design space of hydrogen storage materials.

6. AI-Driven Platforms for Hydrogen Storage

With the growing availability of curated datasets and advanced ML algorithms, recent years have seen the development of AI-driven platforms that provide integrated environments for data analysis, predictive modeling, and inverse design in hydrogen storage materials (HSMs). In contrast to

stand-alone models, these platforms place greater emphasis on systematic functionality, interoperability, and user accessibility, enabling researchers to perform high-throughput screening, interpret results more effectively, and identify promising design candidates through intuitive and user-oriented interfaces.

To accelerate alloy discovery, Lu et al. developed the FIND (Forward–Inverse Navigation and Discovery) platform (https://find4hyalloy.zicp.fun/) for solid-state hydrogen storage alloys [91]. Built on a knowledge base of more than 1,000 alloy systems and 6,000 experimental records, FIND integrates Magpie-based feature engineering with ensemble regression models to predict plateau pressure, enthalpy, entropy, and hydrogen capacity. In the forward navigation framework, the trained models extend beyond single-property prediction and enable simultaneous evaluation of multiple key thermodynamic and capacity indicators, thereby demonstrating a robust capability for multi-objective modeling. (Fig. 11a)

Beyond alloy-focused platforms, Yao et al. [92] proposed Cat-Advisor, which represents the agent-based platform for solid-state hydrogen storage and is specifically designed to address the sluggish dehydrogenation kinetics of MgH_2. As shown in Fig. 11b, leveraging LLM-driven literature mining, the authors extracted structured information from 759 publications, thereby constructing a knowledge base of 809 catalysts and 6,555 entries with an efficiency nearly forty times higher than manual curation. ML models were then trained to predict onset dehydrogenation temperature and activation energy, with XGBoost selected as the most reliable predictor for subsequent inverse-design tasks. (Fig. 11b)

To the best of current knowledge, DIVE (Descriptive Interpretation of Visual Expression) is the first literature-scale AI system comprising two trained agents tailored to hydrogen-storage research [93]: (i) a DIVE extraction agent for literature-scale data mining and (ii) a *DigHyd* materials-design agent trained on the resulting real database (Digital Hydrogen Platform). Built on a large, heterogeneous corpus, the extraction agent uses task-specific prompt templates to transform plots (e.g., PCT, TPD, discharge curves) into standardized textual descriptions; downstream models then convert these into structured records, and an embedding-based matcher with unit normalization assigns precision and coverage scores against human annotations (Fig. 11c). In head-to-head evaluations, the extraction agent achieves 10–15% higher figure-derived extraction accuracy than leading commercial multimodal models and >30% over open-source alternatives. Applied to ~4,000 experimental papers, the extraction workflow yields a machine-readable database of

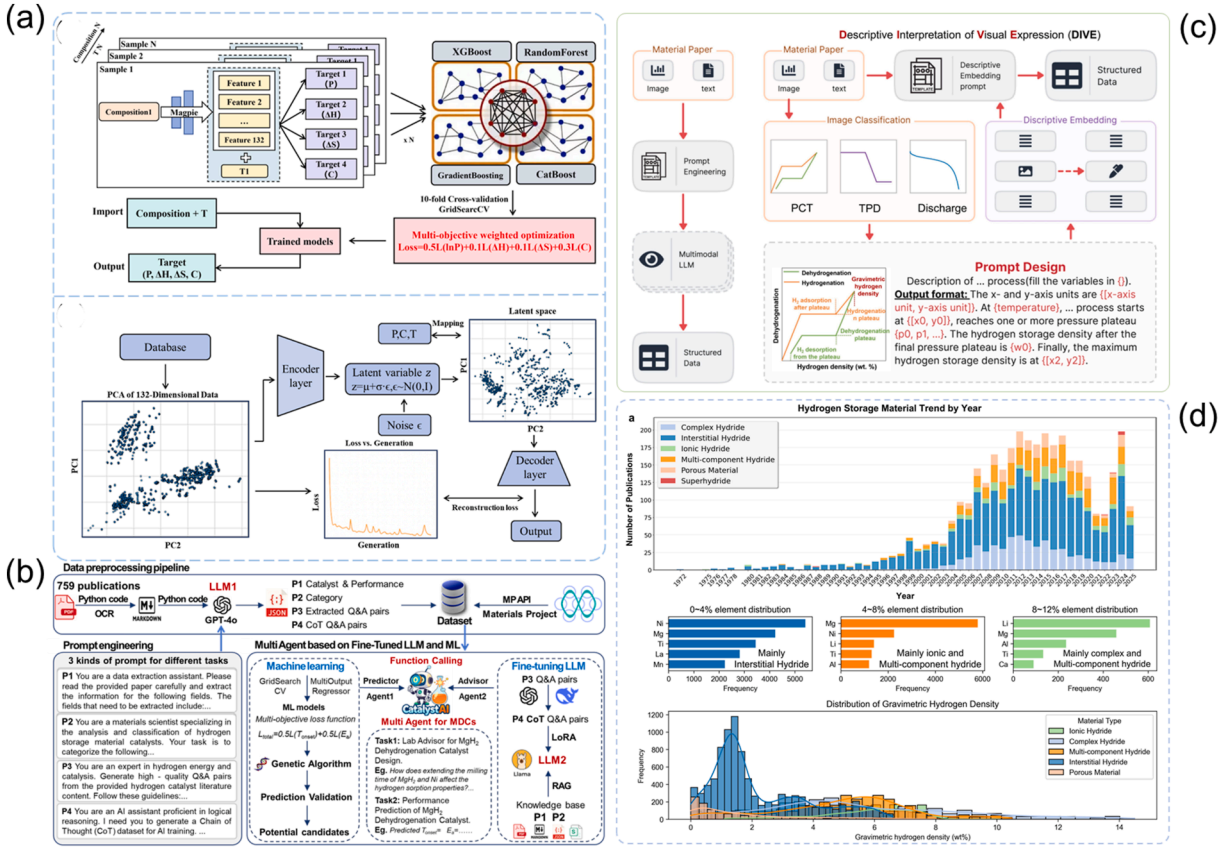

Fig. 11. (a) Forward navigation framework for hydrogen storage alloy prediction on the FIND platform. Inverse discovery framework for hydrogen storage alloy design on the FIND platform. [91] (b) Schematic of a Large Language Model (LLM)-Driven ML Framework for MgH_2 Dehydrogenation Catalyst Design. [92] (c) Overview of the DIVE workflow and its assessment. [93] (d) Corpus-level summary from >4,000 publications on hydrogen-storage materials. [93].

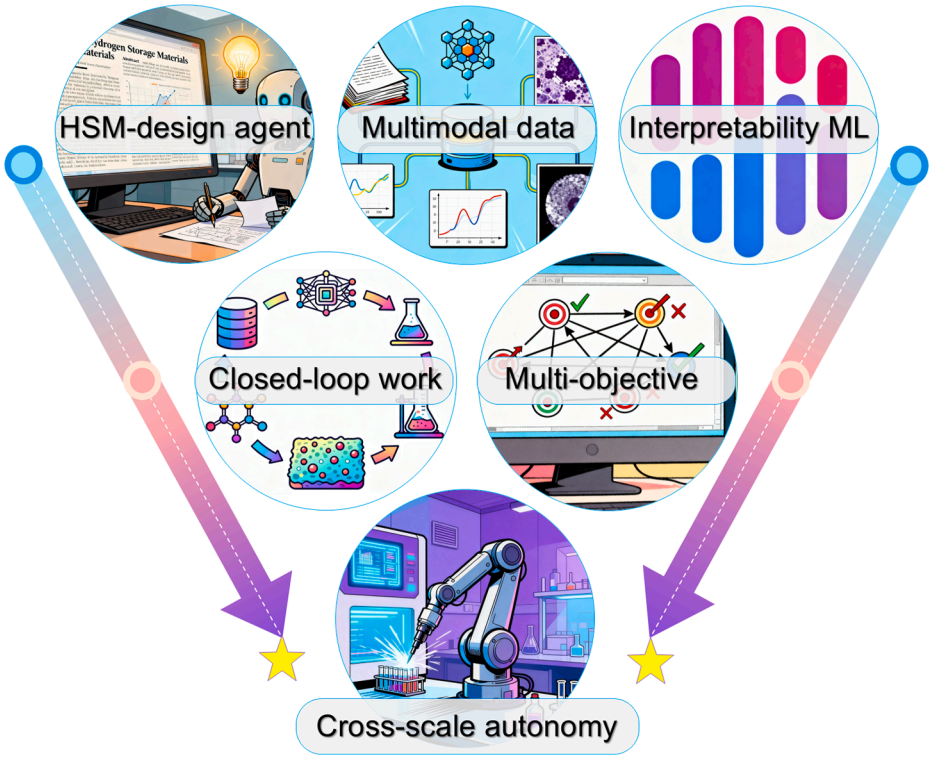

Fig. 12. Roadmap of AI-driven hydrogen storage research, highlighting key developments from present capabilities toward cross-scale autonomous discovery.

>30,000 entries that directly trains and powers the *DigHyd* design agent for rapid inverse design. (Fig. 11c)

By integrating the database, machine learning models trained on this database, and LLMs, it becomes straightforward to construct materials-focused AI agents using simple instructions and schema interface functions. *DigHyd* [93] demonstrated an iterative design–prediction–optimization capability. Researchers can guide the AI agent to propose novel materials by specifying the material class, potential elements, and target properties such as gravimetric hydrogen density, pressure, and temperature. The latter was predicted to achieve 4.19 wt.% hydrogen storage capacity, with Mn (or alternatively Al) contributing to hydride stabilization and plateau pressure optimization. The results highlight the ability of the *DigHyd* agent to rapidly design, predict, and iteratively refine candidate materials in line with researcher-defined goals within minutes. (Fig. 11d)

7. Conclusions and perspectives

7.1. Summary

Over the past decade, the convergence of AI and hydrogen storage research has progressed from feasibility tests to a mature, end-to-end paradigm that accelerates discovery, optimization, and deployment. This Review provides a systematic examination of this integrated workflow. We begin by highlighting existing databases (Section 2), which serve as the foundational data infrastructure for subsequent computational modeling. Proceeding to feature engineering (Section 3), we emphasize how physically grounded descriptors—spanning intrinsic structural, thermodynamic, and electronic attributes together with extrinsic processing and testing conditions—anchor predictive models that connect experimental observables with computational insights. Moving to machine-learning applications (Section 4), we outline how complementary algorithm families now cover the spectrum of tasks in hydrogen storage: linear and kernel methods provide calibrated, interpretable baselines; deep neural networks—including graph-based models—capture high-dimensional, nonlinear structure–property relations; ensembles deliver robust accuracy and transparent design rules; clustering elucidates latent chemistries and regimes; and neural-network potentials extend AI to atomistic dynamics and diffusion. Furthermore, optimization algorithms (Section 5) close the loop by enabling inverse design, multi-objective trade-offs, and experiment planning across vast composition–process spaces. Building on these foundations, AI-driven platforms (Section 6) operationalize the workflow by integrating curated datasets, forward–inverse modeling, and user-facing interfaces for high-throughput screening and synthesis-aware decision-making—for example, FIND for alloy forward/inverse navigation and Cat-Advisor as an agent-based catalyst recommender. Together, these components are shifting the field from intuition-driven trial-and-error toward data-driven, physics-aware, closed-loop discovery, setting the stage for the challenges and prospective directions that follow.

7.2. Challenges

Despite substantial progress, several interlocking barriers must be overcome before AI can reliably guide deployable, industry-scale hydrogen-storage solutions.

First, a discontinuous data lifecycle undermines comparability and thus model generalization. Nominally identical properties often drift across laboratories because synthesis routes, impurity levels, microstructure, and testing protocols differ; crucial metadata (e.g., sample history, cycling protocol, gas purity) are frequently absent, rendering results difficult to reconcile. On the reporting side, literature extraction is labor-intensive and variable in quality; values for the same property can conflict; and bias toward "positive" outcomes distorts the evidence base. Aggregation then compounds these issues: public repositories may include duplicated, missing, or erroneous records; entries arrive in heterogeneous formats and taxonomies with uneven quality control; database ingestion typically lags the literature; parallel, institution-specific efforts fragment coverage; and intellectual-property constraints limit sharing. Without standardized curation, validation, and provenance capture, the field is left with partially interoperable resources that are costly to harmonize and that weaken reliability and reproducibility.

Second, mechanistic interpretability remains challenging. High-accuracy predictors—particularly deep architectures—often act as black boxes, offering limited causal insight into how composition, crystal chemistry, and microstructure govern capacity, (de)hydrogenation thermodynamics and kinetics, cycling stability, and safety. Managing the accuracy–interpretability trade-off therefore becomes central: pipelines built from automated featurizers tend to be high-dimensional and redundant, so stabilizing the structure–property map requires domain-informed feature engineering (systematic cleaning, scaling, selection, and dimensionality reduction) and physically motivated descriptors that translate predictions into actionable design rules.

Third, research coverage and methodological integration remain uneven. Many efforts still optimize a single metric (e.g., gravimetric capacity) rather than the genuinely multi-objective criteria that govern deployability (e.g., usable capacity within realistic temperature, pressure windows, sorption kinetics, cycling durability, safety margins, and cost). At the same time, the deep fusion of high-throughput forward screening with inverse design and process-aware optimization is only beginning, and links to manufacturing variability are rarely explicit. As a result, computational optima do not always survive fabrication tolerances. Closing this simulation–experiment gap will require closed-loop, multi-fidelity workflows that unify computation, synthesis, and operando testing under industrially relevant conditions, supported by user-friendly, provenance-aware platforms accessible to non-specialists.

7.3. Perspectives

Looking forward, meaningful progress in AI-enabled hydrogen storage will hinge on rebuilding the data and modeling stack around physical fidelity, standardization, and rigorous validation. A first priority is data infrastructure and governance: community-endorsed, machine-readable ontologies, metadata schemas, and provenance routines should unify literature and in-house records into traceable, quality-rated repositories. Harmonized file formats and taxonomies—together with classification standards, quality assessment, and IP management—will reduce fragmentation and enable reproducibility. Treating "data management" as first-class research—covering acquisition, annotation, augmentation, curation, ablation, and iterative improvement—will support high-throughput synthesis and characterization while mitigating publication bias and improving cross-study comparability.

From literature-fed to data-grounded design agents. In recent years, AI for materials has progressed along a recognizable trajectory: empirical fitting → early ML (e.g., ANN) → advanced ML (e.g., XGBoost, symbolic regression) → deep learning with interatomic potentials → use of LLMs → literature-fed AI agents → data-grounded design agents trained on curated, real datasets. Hydrogen-storage research follows the same arc. Literature-fed agents—powered by LLMs and automated extraction—are invaluable for scaling evidence collection, but they inherit reporting biases and heterogeneous standards. The next step, and the focus of our forward-looking view, is data-grounded design agents trained on curated, physically consistent, and quality-rated databases (e.g., *DigHyd* [93]), with explicit uncertainty quantification, applicability-domain control, and continuous retraining as new measurements arrive. To avoid a regression in perspective, our summary and outlook should therefore emphasize not only literature-fed agents, but DigHyd-trained design agents that enable mechanism-aware inverse design and deliver experimentally verifiable candidates. AI agents trained solely on unstructured literature are prone to distraction by large volumes of tangential or irrelevant information, which dilutes their design capability. By contrast, structuring past publications into curated databases and knowledge graphs provides the essential foundation for reliable reasoning and targeted exploration—a strategy already prototyped in *DigHyd*.

Beyond hydride-focused resources, the Dynamic Database of Solid-State Electrolytes (DDSE) aggregates >3500 experimentally reported SSEs spanning hydride, halide, sulfide, and oxide chemistries and seven carrier species (Li^+, Na^+, K^+, Ag^+, Ca^{2+}, Mg^{2+}, Zn^{2+}). Crucially, DDSE includes numerous hydride-type entries (e.g., borohydride/closo-borane families) that intersect with classes also investigated for solid-state hydrogen storage, making DDSE an ML-ready resource that facilitates methodological transfer between electrolyte-transport descriptors and hydrogen-storage screening [94].

Architectures for multimodal, multi-fidelity fusion. Building on this foundation, next-generation AI architectures offer a path to unify heterogeneous information and accelerate transfer. Transformer-based and chemistry-specific foundation models can fuse crystal graphs, spectroscopic signatures, microstructure images, and synthesis protocols, capturing long-range dependencies and cross-modal interactions that

elude hand-crafted descriptors. Pretraining on multi-fidelity corpora, with principled weighting of high- and low-accuracy labels, followed by rapid task adaptation, will make these models practical for new chemistries and operating conditions with minimal additional data.

Related advances in hydride solid-state electrolytes illustrate a modular, data-centric AI framework—built on an LLM-curated evidence base, interpretable linear surrogates, genetic-algorithm structure search, and ab initio metadynamics—that elucidates migration mechanisms and yields activation-barrier predictions in close agreement with experiment. We view such literature-fed pipelines as stepping stones: when coupled to curated, real-data repositories, they become data-grounded design agents that are transferable to hydrogen-storage hydrides for mechanism-aware screening and rapid down-selection.

Reliability, interpretability, and validation. Accuracy must advance in tandem with interpretability and reliability. Automated machine learning can reduce manual bias in model selection and hyperparameter tuning, while hybrid paradigms—pairing high-accuracy black-box learners with transparent surrogates—distill actionable design rules. Post-hoc and intrinsic interpretability (e.g., SHAP, rule extraction, tree-based importance) should be embedded to translate patterns into chemistry-aware heuristics for inverse design. Robustness requires disciplined out-of-distribution management: independent test sets, Bayesian regularization and uncertainty quantification, explicit applicability-domain visualization, and outlier/novelty diagnostics that flag when new measurements fall outside the training manifold. In genuinely small-sample regimes, meta-learning (including stacking and transformational ML), reinforcement learning for experimental decision-making, and carefully designed transfer protocols can further improve sample efficiency and portability.

Systematic coverage and closed-loop operation. Meaningful progress also demands systematic coverage of the design space. Coordinated benchmarking across Pd-based alloys, TiFe-type intermetallics, C15 Laves phases, rare-earth systems, ZrCo and Zr_2M families, complex hydrides, and engineered multiphase composites will clarify where specific algorithms succeed or fail and what descriptors truly transport across chemistries. Method synergies should be executed in closed-loop workflows in which experiments, electronic-structure calculations, databases, and active learning co-evolve: DFT and experimental "backfilling" expand the training domain, and *DigHyd*-trained design agents are retrained continuously as evidence accrues. With *DigHyd*, such workflows can be executed within minutes by leveraging past large-scale databases, machine learning models, and the intrinsic reasoning capabilities of LLMs, forming a self-consistent closed-loop design framework that points toward the true "intelligent brain" of future high-throughput experimental platforms.

At the atomistic level, machine-learned interatomic potentials can extend molecular dynamics to nanosecond–microsecond timescales at near-DFT accuracy, exposing rare but performance-critical events—hydrogen diffusion through complex interfaces, phase transformations under cycling, and long-term degradation [95].

Multi-objective, goal-directed optimization. Translational impact will be maximized by reframing discovery as multi-objective inverse design. Coupling surrogate models with evolutionary search and Bayesian optimization enables simultaneous balancing of gravimetric and volumetric capacity, equilibrium pressure/thermodynamics, (de)hydrogenation kinetics, cycling stability, safety, and cost. Embedding life-cycle assessment and recyclability directly into objective functions will align materials choices with industrial and environmental constraints from the outset, yielding candidates that are not only high-performing but also manufacturable and sustainable.

Cross-scale integration and autonomy. Physics-aware digital twins that bridge atomistic predictions to bed/reactor-scale heat- and mass-transfer models will enable real-time performance forecasting, operating-window optimization, and predictive maintenance. In parallel, autonomous laboratories—linking generative design, algorithmic screening, robotic synthesis, automated characterization, and continuous model retraining—can implement a closed loop of "generate → screen → synthesize → test," with demonstrated orders-of-magnitude throughput gains, and support system-level co-optimization (e.g., coupling materials selection with reactor geometry and thermal management).

In conclusion, AI-enabled hydrogen storage should advance from literature-fed workflows to data-grounded, mechanism-aware design agents operating in closed-loop with computation and experiment. In parallel, the convergence of robust data and evidence standards, multimodal foundation models, physics-informed modeling with high-fidelity simulation, multi-objective and mechanism-aware optimization, cross-scale integration, uncertainty-aware validation, and autonomous experimentation will transform discovery. By closing current gaps in standardization, interpretability, and sustainability—and by running end-to-end, self-updating pipelines—the community can deliver explainable, efficient, and industry-ready hydrogen storage solutions, accelerating the transition to a sustainable hydrogen economy. (Fig. 12)

CRediT authorship contribution statement

Yuyi Chu: Writing − original draft, Data curation, Conceptualization. **Yanxin Li:** Writing − original draft, Data curation. **Yanxi Lyu:** Writing − original draft, Data curation. **Di Zhang:** Writing − review & editing. **Hao Li:** Writing − review & editing, Supervision, Funding acquisition. **Weijie Yang:** Writing − review & editing, Supervision, Funding acquisition. **Jianqiu Li:** Supervision, Funding acquisition. **Liang Zhang:** Writing − review & editing, Supervision, Resources, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

L.Z. acknowledges the financially supported by the National Natural Science Foundation of China (22521202, 22373055, 52250710681), the Program of Beijing Huairou Laboratory (ZD2022006A), the Tsinghua University Dushi Research Program. W.Y. acknowledges financial support for this work from the Natural Science Foundation of Hebei (Grant No. E2023502006), the Fundamental Research Fund for the Central Universities (2025MS131) and the Fundamental Research Funds for the Central Universities (Grant No. 2025JC008). D.Z. and H.L. are grateful for the financial support from JSPS KAKENHI (grant nos. JP25K01737, JP25K17991, JP24K23068, and JP25H01508), the Hirose Foundation.

References

[1] W. Ge, N.S. Prime, S.J. Smith, J. Liu, Y. Ren, Y. Zhou, H. Wu, X. Wang, G. Chen, Environ. Pollut. 367 (2025) 125669.
[2] M. Zeng, S. Xue, M. Ma, X. Zhu, Renew. Sustain. Energy Rev. 20 (2013) 169–185.
[3] Y. Huang, Y. Zhou, R. Zhong, C. Wei, B. Zhu, Int. J. Hydrog. Energy 49 (2024) 659–669.
[4] V.M. Avargani, S. Zendehboudi, N.M.C. Saady, M.B. Dusseault, Energy Convers. Manag. 269 (2022) 115927.
[5] G. Aziz, S. Sarwar, R. Waheed, M.S. Khan, Resour. Policy 80 (2023) 103240.
[6] K. Shu, B. Guan, Z. Zhuang, J. Chen, L. Zhu, Z. Ma, X. Hu, C. Zhu, S. Zhao, H. Dang, T. Zhu, Z. Huang, Int. J. Hydrogen Energy 97 (2025) 160–213.
[7] P. Yin, X. Niu, S.-B. Li, K. Chen, X. Zhang, M. Zuo, L. Zhang, H.-W. Liang, Nat. Commun. 15 (2024) 415.
[8] S. Dutta, J. Ind. Eng. Chem. 20 (2014) 1148–1156.
[9] D. Tang, G.-L. Tan, G.-W. Li, J.-G. Liang, S.M. Ahmad, A. Bahadur, M. Humayun, H. Ullah, A. Khan, M. Bououdina, J. Energy Storage 64 (2023) 107196.
[10] M.S. Yahya, N.N. Sulaiman, N.S. Mustafa, F.A.H. Yap, M. Ismail, Int. J. Hydrogen Energy 43 (2018) 14532–14540.
[11] J.H. Shin, G.-J. Lee, Y.W. Cho, K.S. Lee, Catal. Today 146 (2009) 209–215.
[12] F.A.H. Yap, N.S. Mustafa, M. Ismail, RSC Adv. 5 (2015) 9255–9260.
[13] M. Ismail, N.S. Mustafa, N. Juahir, F.A.H. Yap, Mater. Chem. Phys. 170 (2016) 77–82.
[14] A. Kose, M.F. Fellah, Int. J. Hydrogen Energy 79 (2024) 1086–1099.
[15] Y. Chen, J. Dai, L. Wang, Y. Song, Int. J. Hydrogen Energy 44 (2019) 23216–23224.
[16] V.V. Berezovets, R.V. Denys, I.Y. Zavaliy, Y.V. Kosarchyn, Int. J. Hydrogen Energy 47 (2022) 7289–7298.
[17] P. Zhou, Q. Zhou, X. Xiao, X. Fan, Y. Zou, L. Sun, J. Jiang, D. Song, L. Chen, Adv. Mater. 37 (2025) 2413430.
[18] X. Niu, S. Li, Z. Zhang, H. Duan, R. Zhang, J. Li, L. Zhang, ACS Catal. 15 (2025) 4374–4383.
[19] S.-H. Jang, D. Zhang, H.B. Tran, X. Jia, K. Konno, R. Sato, S. Orimo, H. Li, Chem. Sci. 16 (2025) 23111–23120.
[20] Energy.Gov (n.d.). https://www.energy.gov/eere/fuelcells/doe-technical-targets-onboard-hydrogen-storage-light-duty-vehicles
[21] W. Fu, Y. Li, Z. Dong, J. Li, R. Liu, Y. Yang, J. Li, L. Zhang, C. Wang, T. Yao, J. Cai, Z. Xu, H. Tian, Y. Gao, Z. Fu, J. Lin, H. Huang, J. Xu, Y. Du, X. Yang, W. Yang, Int. J. Hydrogen 207 (2026) 153434.
[22] Y. Xu, Y. Zhou, C. Li, S. Dong, H. Liu, W. Yang, Y. Li, H. Jiang, Z. Ding, H. Li, L.L. Shaw, Fuel 373 (2024) 132340.
[23] Y. Wang, J. Liu, Z. Wang, F. Du, Int. J. Hydrogen Energy 88 (2024) 626–637.
[24] A.B. Çolak, Int. J. Hydrogen Energy 144 (2025) 912–923.
[25] S. Suwarno, G. Dicky, A. Suyuthi, M. Effendi, W. Witantyo, L. Noerochim, M. Ismail, Int. J. Hydrogen Energy 47 (2022) 11938–11947.
[26] X.-Y. Zhou, J.-H. Zhu, Y. Wu, X.-S. Yang, T. Lookman, H.-H. Wu, Acta Mater. 224 (2022) 117535.
[27] V. Korostelev, J. Wagner, K. Klyukin, J. Mater. Chem. A 11 (2023) 23576–23588.
[28] L. Ward, A. Agrawal, A. Choudhary, C. Wolverton, Npj Comput. Mater. 2 (2016) 16028.
[29] Y. Zhang, Y. Bai, J. Xu, R. Wei, Int. J. Hydrogen Energy 94 (2024) 687–696.
[30] Z. Lu, J. Wang, Y. Wu, X. Guo, W. Xiao, Int. J. Hydrogen Energy 47 (2022) 34583–34593.
[31] N. Radhika, M.S. Niketh, U.V. Akhil, A.A. Adediran, T.-C. Jen, Results Eng. 23 (2024) 102780.
[32] J. Kim, T. Ha, J. Lee, Y. Lee, J. Shim, Met. Mater. Int. 29 (2023) 861–869.
[33] P. Zhou, X. Xiao, X. Zhu, Y. Chen, W. Lu, M. Piao, Z. Cao, M. Lu, F. Fang, Z. Li, L. Jiang, L. Chen, Energy Storage Mater. 63 (2023) 102964.
[34] S. Dong, Y. Wang, J. Li, Y. Li, L. Wang, J. Zhang, Int. J. Hydrogen Energy 48 (2023) 38412–38424.
[35] S. Meduri, J. Nandanavanam, Energy AI 12 (2023) 100230.
[36] J. Annamalai, P. Murugan, D. Ganapathy, D. Nallaswamy, R. Atchudan, S. Arya, A. Khosla, S. Barathi, A.K. Sundramoorthy, Chemosphere 298 (2022) 134184.
[37] J. Kim, W. Choe, Chem 8 (2022) 617–631.
[38] F.M.A. Noa, M. Abrahamsson, E. Ahlberg, O. Cheung, C.R. Göb, C.J. McKenzie, L. Öhrström, Chem 7 (2021) 2491–2512.
[39] N.S. Bobbitt, A.S. Rosen, R.Q. Snurr, Fluid Phase Equilib. 519 (2020) 112642.
[40] A. Ahmed, D.J. Siegel, Patterns 2 (2021) 100291.
[41] N.K. Borja, C.J.E. Fabros, B.T. Doma, Energies 17 (2024) 927.
[42] R.M. Giappa, E. Tylianakis, M.D. Gennaro, K. Gkagkas, G.E. Froudakis, Int. J. Hydrogen Energy 46 (2021) 27612–27621.
[43] S. Shekhar, C. Chowdhury, Results Surf. Interfaces 14 (2024) 100166.
[44] K. Salehi, M. Rahmani, S. Atashrouz, Int. J. Hydrogen Energy 48 (2023) 33260–33275.
[45] M.S. El-Eskandarany, RSC Adv. 9 (2019) 1036–1046.
[46] S. Ren, Y. Fu, L. Zhang, L. Cong, Y. Xie, H. Yu, W. Wang, Y. Li, L. Jian, Y. Wang, S. Han, J. Alloys Compd. 892 (2022) 162048.
[47] X. Yang, J. Zhang, Q. Hou, X. Guo, Nanomaterials 12 (2022) 2468.
[48] C. Li, W. Yang, H. Liu, X. Liu, X. Xing, Z. Gao, S. Dong, H. Li, Angew. Chem. Int. Ed. 63 (2024) e202320151.
[49] M. Rahimi, M. Abbaspour-Fard, A. Rohani, J. Clean. Prod. 329 (2021) 129714.
[50] M. Gheytanzadeh, F. Rajabhasani, A. Baghban, S. Habibzadeh, O. Abida, A. Esmaeili, M. Munir, Sci. Rep. 12 (2022) 21902.
[51] S. Dangwal, Y. Ikeda, B. Grabowski, K. Edalati, Chem. Eng. J. 493 (2024) 152606.
[52] E. Halpren, X. Yao, Z.W. Chen, C.V. Singh, Acta Mater. 270 (2024) 119841.
[53] P. Zhao, X. Zeng, H. Kou, H. Chen, Int. J. Hydrogen Energy 54 (2024) 1251–1263.
[54] M.D. Witman, N.C. Bartelt, S. Ling, P.-W. Guan, L. Way, M.D. Allendorf, V. Stavila, J. Phys. Chem. Lett. 15 (2024) 1500–1506.
[55] R. Cheng, X. Xue, C. Wang, J. Phys. Chem. C 128 (2024) 7742–7749.
[56] K. Batalovic, J. Radakovic, B. Mamula, B. Kuzmanovic, M. Ilic, Adv. Theory Simul. 5 (2022) 2200293.
[57] K. Batalović, J. Radaković, B. Kuzmanović, M.M. Ilić, B. Paskaš Mamula, J. Energy Storage 68 (2023) 107720.
[58] J. Lee, D. Sung, Y. Chung, S.B. Song, J. Huh, Nanoscale Adv. 4 (2022) 2332–2338.
[59] C. Wang, J. Brinkerhoff, Int. J. Hydrogen Energy 46 (2021) 24256–24270.

[60] X. He, J. Liu, C. Yang, G. Jiang, Comput. Mater. Sci. 223 (2023) 112111.
[61] S. Nations, T. Nandi, A. Ramazani, S. Wang, Y. Duan, J. Energy Storage 70 (2023) 107980.
[62] A. Verma, N. Wilson, K. Joshi, Int. J. Hydrogen Energy 50 (2024) 1518–1528.
[63] C. Peng, X. Liu, R. Long, Z. Liu, W. Liu, Chem. Eng. Res. Des. 207 (2024) 100–109.
[64] Z. Lu, J. Wang, Y. Wu, X. Guo, T. Ma, W. Xiao, Int. J. Hydrogen Energy 50 (2024) 379–389.
[65] A. Rahnama, G. Zepon, S. Sridhar, Int. J. Hydrogen Energy 44 (2019) 7337–7344.
[66] Q. Jiang, Z. Jia, S. Lu, P. Song, Z. Gao, Z. Wang, T. Peng, X. Bai, H. Cui, W. Tian, R. Feng, Q. Kang, Z. Liang, H. Yuan, Ceram. Int. 50 (2024) 23058–23069.
[67] M. Witman, S. Ling, D.M. Grant, G.S. Walker, S. Agarwal, V. Stavila, M.D. Allendorf, J. Phys. Chem. Lett. 11 (2020) 40–47.
[68] A. Agafonov, N. Pineda-Romero, M. Witman, V. Nassif, G.B.M. Vaughan, L. Lei, S. Ling, D.M. Grant, M. Dornheim, M. Allendorf, V. Stavila, C. Zlotea, Acta Mater. 276 (2024) 120086.
[69] S. Tiwari, N. Gupta, S. Kumar, A. Kumar, P. Sharma, J. Energy Storage 72 (2023) 108522.
[70] Z. Ding, Z. Chen, T. Ma, C. Lu, W. Ma, L. Shaw, Energy Storage Mater. 27 (2020) 466–477.
[71] M.D. Witman, S. Ling, M. Wadge, A. Bouzidi, N. Pineda-Romero, R. Clulow, G. Ek, J.M. Chames, E.J. Allendorf, S. Agarwal, M.D. Allendorf, G.S. Walker, D.M. Grant, M. Sahlberg, C. Zlotea, V. Stavila, J. Mater. Chem. A 11 (2023) 15878–15888.
[72] J. Gao, X. Guo, Y. Wu, W. Xiao, L. Hao, Int. J. Hydrogen Energy 58 (2024) 657–667.
[73] T. Malinova, Z. Guo, Mater. Sci. Eng. -Struct. Mater. Prop. Microstruct. Process. 365 (2004) 219–227.
[74] A. Rahnama, G. Zepon, S. Sridhar, Int. J. Hydrogen Energy 44 (2019) 7345–7353.
[75] Q. Tian, Y. Zhang, Y. Wu, Z. Tan, Int. J. Hydrogen Energy 34 (2009) 1931–1936.
[76] A. Pedrielli, P. Trevisanutto, L. Monacelli, G. Garberoglio, N. Pugno, S. Taioli, Nanoscale 14 (2022) 5589–5599.
[77] R. Kronberg, H. Lappalainen, K. Laasonen, J. Phys. Chem. C 125 (2021) 15918–15933.
[78] M. Kusdhany, S. Lyth, Carbon 179 (2021) 190–201.
[79] J. Hattrick-Simpers, K. Choudhary, C. Corgnale, Mol. Syst. Des. Eng. 3 (2018) 509–517.
[80] J. Hattrick-Simpers, B. DeCost, Mol. Syst. Des. Eng. 5 (2020) 589–591.
[81] Y. Chen, L. Zhou, S. Pei, Z. Yu, Y. Chen, X. Liu, J. Du, N. Xiong, IEEE Trans. Syst. Man Cybern. Syst. 51 (2021) 3939–3953.
[82] A. Rahnama, S. Sridhar, Materialia 7 (2019) 100366.
[83] C. Wen, H. Shen, Y. Tian, G. Lou, N. Wang, Y. Su, Scr. Mater. 252 (2024) 116240.
[84] N. Wang, S. Huang, Phys. Rev. B 102 (2020) 094111.
[85] H. Wang, L. Zhang, J. Han, W. E, Comput. Phys. Commun. 228 (2018) 178–184.
[86] R. Sato, L.J. Conway, D. Zhang, C.J. Pickard, K. Akagi, K. Sau, H. Li, S. Orimo, Proc. Natl. Acad. Sci. 122 (2025) e2413480122.
[87] S. Maghsoudy, P. Zakerabbasi, A. Baghban, A. Esmaeili, S. Habibzadeh, Sci. Rep. 14 (2024) 1503.
[88] M. Gheytanzadeh, F. Rajabhasani, A. Baghban, S. Habibzadeh, O. Abida, A. Esmaeili, M.T. Munir, Sci. Rep. 12 (2022) 21902.
[89] S.H. Suárez, D. Chabane, A. N'Diaye, Y. Ait-Amirat, A. Djerdir, Int. J. Hydrogen Energy 50 (2024) 178–189.
[90] S. Ghude, C. Chowdhury, Chem. – Eur. J. 29 (2023) e202301840.
[91] X. Lu, S. Luo, J. Li, M. Chen, T. Yao, Z. Xu, Y. Yan, J. Li, X. Shao, Z. Gao, W. Yang, J. Mater. Inf. 5 (2025) 48.
[92] T. Yao, Y. Yang, J. Cai, R. Liu, Z. Dong, X. Tang, X. Shao, Z. Gao, G. An, W. Yang, J. Magnes. Alloys 14 (2026) 101858.
[93] D. Zhang, X. Jia, T.B. Hung, S.H. Jang, L. Zhang, R. Sato, Y. Hashimoto, T. Sato, K. Konno, S. Orimo, H. Li, Chem. Sci. 17 (2026) 3031–3042.
[94] F. Yang, E. Campos Dos Santos, X. Jia, R. Sato, K. Kisu, Y. Hashimoto, S. Orimo, H. Li, Nano Mater. Sci. 6 (2024) 256–262.
[95] X. Niu, S. Zhen, R. Zhang, J. Li, L. Zhang, Chem. Sci. 16 (2025) 14884–14893.