Fingerprinting diverse nanoporous materials for optimal hydrogen storage conditions using meta-learning

Yangzesheng Sun, Robert F. DeJaco, Zhao Li, Dai Tang, Stephan Glante, David S. Sholl, Coray M. Colina, Randall Q. Snurr, Matthias Thommes, Martin Hartmann, J. Ilja Siepmann

Adsorptive hydrogen storage is a desirable technology for fuel cell vehicles, and efficiently identifying the optimal storage temperature requires modeling hydrogen loading as a continuous function of pressure and temperature. Using data obtained from high-throughput Monte Carlo simulations for zeolites, metal-organic frameworks, and hyper-cross-linked polymers, we develop a meta-learning model that jointly predicts the adsorption loading for multiple materials over wide ranges of pressure and temperature. Meta-learning gives higher accuracy and improved generalization compared to fitting a model separately to each material and allows us to identify the optimal hydrogen storage temperature with the highest working capacity for a given pressure difference. Materials with high optimal temperatures are found in close proximity in the fingerprint space and exhibit high isosteric heats of adsorption. Our method and results provide new guidelines toward the design of hydrogen storage materials and a new route to incorporate machine learning into high-throughput materials discovery.

INTRODUCTION
The past decade has witnessed a marked transition in the development of automobile engines. Large naturally aspirated gasoline engines were replaced by smaller turbos, and the market for hybrid and electric vehicles has undergone a rapid growth (1). Enabled by a combination of electrochemical reactions and a replenishable, clean fuel source, hydrogen fuel cells have also emerged as an alternative (2). Several hydrogen fuel cell vehicle models are already available in the market, but these vehicles store hydrogen compressed at pressures of up to 700 bar (3). These high pressures require the fuel tank to be made from a highly pressure-tolerant material into a cylindrical shape with large wall thickness and pose greater threats of hydrogen leakage or even explosion after a severe traffic accident (4).

Adsorptive hydrogen storage has enjoyed growing interest to address these aforementioned problems, with many types of nanoporous materials (NPMs) explored including zeolites (5), carbon-based materials (6), and metal-organic frameworks (MOFs) (7). By (partially) filling the fuel tank with these materials, the goal is to achieve a storage capacity at pressures as low as 100 bar (7, 8) comparable to that of commercial compressed hydrogen tanks. The reduction of the maximum tank pressure not only allows for a more flexible shape of the fuel tank in a vehicle but also improves its safety and reduces compression costs at the filling station. Using adsorptive hydrogen storage not only opens up an enormous NPM space but also leads to an underexplored complexity of the state space (pressure and temperature differences between tank filling and fuel depletion).

High-throughput molecular simulations have been the predominant approach to screen the adsorption properties of large numbers of nanoporous frameworks (9, 10). Typically, a hierarchy of simulations is conducted to calculate the objective property with increasing accuracy and computational cost. These multiple rounds of simulations conclude at identifying a sufficiently small number of best materials that can potentially be experimentally synthesized (10, 11). To reduce the total computational cost of the high-throughput screening process, machine learning algorithms have been used to predict adsorption properties from structural information of NPMs (8, 12).

Most of the simulation and machine learning approaches to NPM discovery focus on predicting the property at the same thermodynamic state for all materials. For instance, the temperature of 77 K is usually chosen to measure hydrogen storage due to the belief that it gives larger storage capacity than higher temperatures (13). However, the working capacity of an NPM, defined as the equilibrium loading difference between two pressures, does not monotonically decrease with increasing temperature (14). At a low temperature, the adsorbate molecule may bind so strongly to the material that it remains near saturation even at the lower depletion pressure resulting in a low working capacity (15). Therefore, increasing the temperature may still retain saturation at the higher pressure while decreasing the loading at the lower pressure, thus increasing the working capacity. Because adsorption of gas molecules is very weak at very high temperatures due to the inherent entropic cost, there may exist an optimal temperature for an NPM at which it obtains the largest working capacity between two pressures (14).

The nonmonotonic relationship makes it insufficient to only consider the adsorption at one temperature when ranking hydrogen storage among many NPMs, giving rise to the need of predicting at multiple temperatures and pressures. Simulations of many materials...
each at multiple state points result in a big “meta-dataset” containing many small adsorption datasets. By fitting an adsorption isotherm function (AIF) with temperature dependence individually to each dataset, one can predict the adsorption at new state points for materials in the meta-dataset (16). However, an AIF may be only valid for a subset of adsorption systems (17) and is poorly generalizable to predicting adsorption for a different material where only a few adsorption data points are available. Moreover, similarities or other relationships among materials in the meta-dataset are not used.

With regard to these challenges, it is necessary to develop a “meta-isotherm function” for prediction of adsorption on a large number of materials. Such a predictive model can serve as an end-to-end solution to adsorption modeling because human specification of the isotherm functional form is not anymore necessary (18). Anderson et al. (19) trained a neural network that operated on both thermodynamic states and features of the material structure. Nevertheless, well-defined material structure representations are needed, and thus, the approach is unsuitable in cases such as generalizing from simulation data to experimental data. In essence, the prediction of adsorption in a single NPM is a machine learning task (20). The training set includes the state points where adsorption measurements are available, and the isotherm function serves as the machine learning model. Thus, our approach aims to solve the higher-level problem of performing multiple base machine learning tasks simultaneously, which is commonly referred to as meta-learning in the machine learning field (21). Meta-learning studies how to dynamically apply the most appropriate hypothesis (functional form) on each base task (the prediction of adsorption in one NPM) and gain from experiences learned on multiple tasks (22). For a meta-learning model, each base learning task and the base dataset constitute one of its “meta-data” samples. By virtue of being trained on the meta-dataset of many tasks, the meta-learning model is capable of efficiently adapting to new tasks especially when only limited data is available for the new task. Various types of meta-learning algorithms and models have been successfully applied to problems including image recognition and reinforcement learning (23, 24).

Here, we describe a meta-learning method to predict gas adsorption as a function of temperature and pressure for diverse NPMs in a single machine learning model (Fig. 1). We develop the model using an encoder-decoder architecture (25) to extract fingerprints of hydrogen adsorption for NPMs as the dimensionality-reduced representation. In contrast to material representations using structural features (19) or guest-host interaction energy distributions obtained from force fields (26) for machine learning of gas adsorption, our meta-learning method gives the fingerprint for an NPM completely based on its adsorption data. The model directly encodes the hydrogen loading surface of an NPM into a fingerprint representation, and thus, sparse experimental data could also be used as input. Being free of structural or energetic properties in the input also allows us to apply the same or fine-tuned models to NPMs with diverse structures and chemistries. The meta-learning model trained on data for already synthesized zeolites is able to generalize to other types of NPMs and allows for few-shot learning where AIFs fail because of lack of enough data. We apply the meta-learning method to predict the optimal hydrogen storage temperatures for synthesized and hypothetical all-silica zeolites, hyper-cross-linked polymers (HCPs), and MOFs. The adsorption fingerprints generated by the meta-learning model also show distinct features associated with NPMs at the Pareto front of high optimal temperatures.

![Fig. 1. Meta-learning for predicting gas adsorption loading, q, in NPMs. (A) A meta-learning problem setup. To solve the problem of jointly predicting the materials space and the state (temperature and pressure) space, meta-learning consolidates the prediction of all materials into a single model and can be generalized to new materials. (B) The meta-learning architecture for gas adsorption prediction. The encoder network produces the fingerprint vector z from the 3D (q, p, T) loading surface. The decoder network reconstructs the adsorption loading as a continuous function of (p, T) given the adsorption fingerprint z.](http://advances.sciencemag.org/)
and high working capacities. Last, we demonstrate that the predicted optimal temperature and hydrogen storage capacity given by simulation and meta-learning for a cation-exchanged zeolite are in good agreement with experimental results.

RESULTS
Meta-learning model and dataset
In conventional adsorption modeling, a set of coefficients can be obtained by fitting an AIF to adsorption data for a given sorbate-sorbent system; here, we use the Langmuir, Sips, quadratic, and dual-site Langmuir (DSL) functions (see Materials and Methods). The encoder-decoder architecture is used for the meta-learning model (23) because it is able to generate a latent representation of the meta-data sample for each sorbate-sorbent system (Fig. 1B). Moreover, only the fingerprint with five coefficients instead of the whole dataset of all measurements needs to be stored for each material to predict the adsorption surface (loading, \( q \), as function of pressure, \( p \), and temperature, \( T \)).

To obtain the hydrogen loading data as the meta-dataset, Monte Carlo simulations are performed for 216 all-silica zeolites in the International Zeolite Association (IZA) database (27), 70 hypothetical all-silica zeolites with high predicted synthesizability (Predicted Crystallography Open Database, PCOD-syn) (28), 200 hypothetical MOFs in the ToBaCCo database (29), and 45 configurations drawn from nine different HCPs (30). The NPMs considered for the high-throughput simulations are idealized structures that allow for a thorough investigation of the influence of pore topology and channel system on unary hydrogen adsorption and for identification of the promising structures at given hydrogen storage conditions. Subsequent work on specific promising structures should consider the effects of framework defects (e.g., silanol nests due to missing tetrahedral Si atoms or missing ligands in MOFs), of extra framework species (e.g., cations in alumino silicates or residual solvent in MOFs), and of gas impurities on hydrogen storage performance.

For each NPM, the adsorption loading is measured for 64 state points consisting of eight temperature values ranging from 77.0 to 275.9 K and eight pressure values ranging from 1.0 to 403.4 bar (distributed linearly in \( T^{-1} \) and \( \ln p \)). For computational efficiency, all sorbent materials are treated as rigid frameworks in the simulations, which is generally considered to be a good approximation for adsorption of small molecules in zeolites, but may hold less well for HCPs at the higher temperatures and pressures. The meta-learning model is trained on 160 (74%) IZA zeolites and can be directly applied to PCOD-syn zeolites and HCPs, but it requires fine-tuning on 40 (20%) ToBaCCo MOFs because of their larger pore sizes than the IZA zeolites.

Prediction of hydrogen adsorption
A comparison of the predictive accuracy obtained with our meta-learning model and with four commonly used AIFs for all sorbent materials is shown in Fig. 2. Two types of prediction settings are used to evaluate the performance. In the fitting setting, the meta-learning model and the AIFs are fitted to all 64 state points. This “curve fitting” is essentially a reconstruction task because the objective is to obtain a closed-form representation of the loading surface. In our meta-learning model, this means that the input \( D_{\text{train}} = D \) and the network works as an autoencoder. In the few-shot setting, only eight state points are used to fit the AIFs, which resembles a situation where very limited adsorption data are available. For the meta-learning model, this means that the size of \( D_{\text{train}} \) is decreased to 8 while the mean squared error (MSE) is still evaluated on the entire dataset \( D \). Three types of eight-point subsets are investigated: random sampling (where the same eight points are used for all materials) and eight diagonal or eight edge points on the \( p, T \) grid (see Fig. 2). The predictive accuracy separated by NPM type is shown in fig. S1, and numerical data are provided in table S1.

When the sole objective is to reproduce the adsorption loading measured at all 64 state points, meta-learning already yields significantly more accurate predictions than any of the four AIFs. Here, the meta-learning model gives a geometric mean of the MSE (GM-MSE) for all sorbent materials that is lower by a factor of 1.8 than the DSL AIF. As should be expected for this data-rich case, the GM-MSE for the AIFs decreases with the number of fitting parameters from Langmuir with three parameters to DSL with six parameters. To make the comparison more challenging, we also show data
obtained using the best AIF selected for each specific sorbent material and set of $p, T$ points (see Fig. 2 and table S1). Even in this case, the meta-learning model yields a GM-MSE that is about a factor of 2 smaller for the IZA, PCOD-syn, and HCP sorbents but only a factor of 1.2 smaller for the MOF materials.

To help understand the more accurate reconstruction of the continuous loading surface achieved by the meta-learning model, we focus on the IZA zeolite structures where meta-learning outperforms any of the AIFs to a large extent. The hydrogen adsorption isotherms of two of these zeolites, AFG and AWO [zeolite structures are denoted here by their three-letter framework type code assigned to them by the IZA (27)], are shown in fig. S2. The zeolite AFG is an example for a group of zeolites that exhibit a substantial increase in the loading at high pressure even as the temperature approaches 77 K, i.e., the saturation loading is far from being reached at $T = 77 \text{ K}$ and $p = 403.4 \text{ bar}$. While the meta-learning model can accurately capture this adsorption behavior, the AIFs yield a much flatter isotherm in the high pressure region (marked by the three simulation data for $p \geq 54.6 \text{ bar}$). We surmise that applying a temperature dependence for only the equilibrium constants used in the AIFs is not sufficient and that the loading prefactor in the AIFs (representing the number of adsorption sites) would also need to be made temperature dependent.

The zeolite AWO exhibits a unique adsorption behavior (see fig. S2) where only a minimal change in loading occurs for a pressure increase from 150 to 400 bar at $T = 111 \text{ K}$ or a temperature decrease from 160 to 111 K at $p = 403.4 \text{ bar}$, but a 30% increase in loading is observed for the same pressure increase at $77 \text{ K}$ or a temperature decrease from 111 to $77 \text{ K}$ at $p = 403.4 \text{ bar}$. Longer simulation trajectories yield the same loading behavior, and analysis of configurations taken from these trajectories points to a structural change in the adsorbate packing: while almost all hydrogen molecules reside near the center of the channel along the $a$ axis at $p = 148.4 \text{ bar}$, the narrower channel along the $b$ axis sees a substantial increase in occupancy at $T = 77 \text{ K}$ and $p = 403.4 \text{ bar}$ (see fig. S3). Neither the meta-learning model nor any of the AIFs is able to correctly predict the inflection point near 150 bar for the isotherm at 77 K. However, the meta-learning model partially captures the increased loading at $p = 403.4 \text{ bar}$ as $T = 77 \text{ K}$ is approached. The meta-learning model also signals the unique adsorption behavior for AWO through a much larger uncertainty for the low-temperature isotherm. These observations demonstrate that our meta-learning model is able to capture more complex adsorption phenomena compared with AIFs commonly used in adsorption modeling.

The prediction of the Henry’s coefficient (estimated from the data at $p = 1.0 \text{ bar}$) and of the maximum adsorption loading (i.e., at $p = 403.4 \text{ bar}$) for a specific NPM at a given temperature can also be used to evaluate the performance of the meta-learning model. As can be seen from fig. S4, the meta-learning model predicts these properties very well. The meta-learning model can also predict the Henry’s coefficient as the limiting slope of the loading surface, and we observe a slight shift compared to the data at $p = 1.0 \text{ bar}$ (see fig. S4).

The advantage of the meta-learning model trained on a large number of sorbent materials is even more pronounced in the few-shot setting where the entire loading surface has to be predicted from the information at merely eight state points. In most cases, the GM-MSE for the meta-learning model increases only by a factor of 3 compared to using all 64 data points. The exception is random sample (a), where the GM-MSEs are found to increase by factors of 5 and 19 for the PCOD-syn and MOF sorbents, respectively. This specific sample of eight state points lacks data at the three lowest pressures, and, for the MOF subset, the quadratic AIF slightly outperforms the meta-learning model. Clearly, training on IZA zeolites and fine-tuning on a few MOFs do not lead to satisfactory prediction accuracy when the distribution of state points for the few-shot learning lacks coverage, and differences in pore size distributions lead to different loading trends. With diagonal and edge sampling of the $p, T$ grid, the errors for meta-learning and AIFs decrease compared to the random sampling of points. For the AIFs, choosing the eight state points on the edges of the $p, T$ grid is strongly favored, particularly for the quadratic and DSL functions with the larger number of parameters. In contrast, meta-learning on diagonal and edge points yield about the same accuracy. Given that the extremum temperatures and pressures in a state space region are often more difficult to access by experiments or require longer simulations, meta-learning over both the material space and the state space is more advantageous in providing guidance on selecting a minimal set of conditions for studying a new material compared with fitting each material independently. This also implies that meta-learning is potentially capable of (limited) extrapolation over the state space not covered in training data.

### Optimal hydrogen storage temperature for NPMs

After validating its predictive performance, the meta-learning model trained on all 64 state points is then applied for the prediction of the loading surface on a finer grid to find the optimal hydrogen storage temperature, $T_{\text{max}}$, for each NPM placed in a thermodynamic fuel tank model (31). Figure 3 shows the distribution of the optimal

![Fig. 3. Distribution of predicted optimal temperatures and maximum working capacities for $p_{\text{max}} = 2.71 \text{ bar}$ and $p_{\text{max}} = 30.0 \text{ bar}$. Materials with predicted $T_{\text{max}} < 45 \text{ K}$ are not shown. The gray dashed line denotes the working capacity of a compressed hydrogen tank without any sorbent material, and the green line gives twice this value. Highlighted and labeled points are zeolites along the Pareto front for which extra validation simulations are reported. Error bars denote one SD in an ensemble of eight independently trained meta-learning models. The top part shows the fraction of sorbent materials that exceed a given $T_{\text{max}}$.](http://advances.sciencemag.org/)
temperatures and the corresponding maximum working capacities for \( p_{\text{low}} = 2.71 \text{ bar} \) and \( p_{\text{high}} = 30.0 \text{ bar} \). Unexpectedly, the optimal temperature for most IZA and PCOD-syn zeolites and HCPs exceeds 77 K, indicating that the common practice of evaluating hydrogen storage at 77 K does not reflect the best performance of all NPMs. The working temperature of 77 K is a more reasonable choice for MOFs (32), but our study indicates that \( \approx 20\% \) of the ToBaCCo MOFs yield \( T_{\text{max}} > 77 \) K. IZA zeolites along the Pareto front are highlighted in Fig. 3 (numerical data are given in table S2). The highest \( T_{\text{max}} \) among all zeolites predicted by the meta-learning model fall between 130 and 140 K, but zeolites with the highest working capacity yield \( T_{\text{max}} < 77 \) K.

For the eight zeolites on the Pareto front, longer Monte Carlo simulations in the vicinity of \( T_{\text{max}} \) are carried out to validate the meta-learning predictions (which used data from shorter high-throughput simulations for the training). The adsorption isotherms and working capacities as functions of temperature for four of these validated zeolites are shown in Fig. 4 (see fig. S5 for the other four zeolites). Comparing zeolites at the high-temperature (e.g., MTN and DOH) and high-capacity ends (e.g., RWY and IRY), their hydrogen adsorption isotherms exhibit different behaviors in the Henry’s law region and at saturation. At low pressure (<5 bar), the loading in MTN increases sharply with decreasing temperature, whereas that in RWY increases rather gradually. Although the saturation capacity in RWY is three times higher than for MTN, the loading in MTN at 77 K and 1 bar is four times higher than for RWY. As a result, the working capacity of MTN is found to increase when the temperature is raised above 77 K. In contrast to MTN, the loading in RWY remains relatively low at 2.71 bar even at 77 K and rapidly increases at 30 bar when the temperature is decreased to 77 K. Thus, the working capacity for RWY monotonically decreases for \( T > 77 \) K. For zeolites MON and SBN with intermediate \( T_{\text{max}} \) values, the adsorption loading at 77 K and \( p > 50 \) bar is very flat, whereas those for RWY and MTN continue to rise with increasing pressure. This indicates that the latter zeolites have multiple types of adsorption sites that reach saturation at different points on the \( p, T \) surface. The validation simulations for zeolites on the Pareto front confirm the accuracy of the meta-learning model for predicting both the optimal temperature and the associated working capacity (see Fig. 4 and fig. S5).

Given that our meta-learning model predicts \( T_{\text{max}} < 77 \) K for many sorbent materials (see Fig. 3), i.e., outside of the temperature range used for the high-throughput simulations, an important question is whether the meta-learning model is capable of limited extrapolation to points outside of the \( p, T \) grid used for training. Thus, validation simulations for zeolites RWY and IRY are performed for \( 45 \leq T \leq 75 \) K. Figure 5 shows a comparison of the meta-learning predictions to those of a two-layer neural network and the Sips AIF trained on only the 64 \( p, T \) data for zeolite RWY. The meta-learning model successfully predicts the temperature dependence not only of the two isobars but also of the working capacity. Although the Sips AIF is the best performer among the four AIFs for reproducing the 64 \( p, T \) data for this zeolite, it slightly underpredicts the low-pressure loading at lower temperatures and overpredicts the high-pressure loading at higher temperatures, and this results in a slight under-prediction of \( T_{\text{max}} \) and an overprediction of the working capacity for \( T < 80 \) K and \( T > 170 \) K. The two-layer neural network provides more flexibility than the Sips AIF, and the former reproduces the

Fig. 4. Hydrogen adsorption isotherms and working capacities (\( p_{\text{low}} = 2.71 \) bar and \( p_{\text{high}} = 30.0 \) bar) for four representative zeolite structures on the Pareto front: RWY, MON, SBN, and MTN. \( T_{\text{val, max}} \) denotes the temperature with the highest capacity among temperatures used for the validation simulations (see table S2). Circles, triangles, and solid lines denote high-throughput simulations, validation simulations, and meta-learning predictions, respectively. The dashed lines in the bottom row denote the working capacity for a compressed hydrogen tank without sorbent material, respectively.
isobars with satisfactory accuracy over the training range. However, the neural network violates adsorption thermodynamics by predicting a very shallow minimum at $T \approx 170$ K for the low-pressure isobar. Furthermore, the neural network is not able to capture the steep rise in the low-pressure loading for $T < 77$ K. Thus, it fails to predict a maximum in the working capacity. Although the ensemble of eight meta-learning models yields larger uncertainties at temperatures below the training range, its average prediction is in excellent agreement with the validation simulations. This result highlights the benefits of the meta-learning approach that yields vastly more accurate predictions by incorporating the information from a large set of NPMs; that is, the knowledge of the working capacity usually exhibiting a maximum gleaned from all materials helps to extrapolate to the lower temperatures relevant for materials with weaker adsorption sites.

For the eight zeolites on the Pareto front, we find that the mean relative deviation between the optimal temperature predicted by the meta-learning model and the temperature of the validation simulation with the highest working capacity is 4%, whereas those for the Langmuir, Sips, quadratic, and DSL AIFs are 7, 9, 13, and 10%, respectively. Figure S6 illustrates the differences for all materials with $T_{\text{max}} > 45$ K as predicted by the meta-learning model. Overall, the Langmuir AIF with the fewest parameters performs best with particularly small deviations for all MOFs with $T_{\text{max}} < 100$ K (i.e., indicating that adsorption in the ToBaCCo MOFs can be well described, assuming only one type of adsorption site) and larger deviations for IZA zeolites and HCPs with $T_{\text{max}} < 90$ K. The Sips AIF performs slightly better than the Langmuir isotherm for sorbent materials with $T_{\text{max}} < 90$ K but yields much larger deviations for those at the high-temperature end. Unexpectedly, the quadratic and DSL AIFs with their larger number of fitting parameters yield large numbers of outliers also for sorbent materials with $80 < T_{\text{max}} < 140$ K, i.e., within the temperature range of the training data. Thus, it appears that more than 64 $p$, $T$ data points or a higher density of data points near $T_{\text{max}}$ would be required to allow for a reliable prediction of $T_{\text{max}}$ using common AIFs.

The optimal temperature and (maximum) working capacity are dependent on the operating conditions of the fuel tank. To increase the capacity of the tank, one may increase the pressure at which the tank is filled, decrease the pressure at which the tank is depleted, or allow for an increase in the temperature while the tank is being depleted. In this regard, the optimal temperatures and working capacities for IZA zeolites are also calculated, adjusting one of the three conditions to $p_{\text{high}} = 54.6$ bar, $p_{\text{low}} = 1.0$ bar, or a moderate temperature swing of 20 K while keeping the other two at the original values ($p_{\text{low}} = 2.71$ bar, $p_{\text{high}} = 30.0$ bar, and $\Delta T = 0$ K). Here, we define another performance metric as the ratio between working capacities of a tank containing 70% NPM by volume fraction and a tank containing no NPM (see Materials and Methods)

$$r = \frac{\Delta n(p_{\text{low}}, p_{\text{high}}, T, \Delta T, \varepsilon = 0.3)}{\Delta n(p_{\text{low}}, p_{\text{high}}, T, \Delta T, \varepsilon = 1.0)}$$

where $\varepsilon$ is the void fraction. A larger $r$ represents a higher efficiency in hydrogen storage when using the NPM, and $r < 1$ indicates that the NPM is unsuitable for $H_2$ storage due to negative excess adsorption. Comparing different changes in operating conditions (see fig. S7) indicates that the temperature swing yields the largest increase in the working capacity, followed the 1.8-fold increase in $p_{\text{high}}$, and the 2.7-fold decrease in $p_{\text{low}}$ leads to the smallest increase. For zeolites with positive excess adsorption, the increase in $p_{\text{high}}$ (decrease in $p_{\text{low}}$) leads to a slight increase (decrease) in $T_{\text{max}}$ with highly correlated values. Although the temperature swing also lowers $T_{\text{max}}$, there is more scatter in the data with some zeolites showing downward shifts in $T_{\text{max}}$ by more than a factor of 2. Only the temperature swing yields a substantial increase in the performance (i.e., increase in $r$ value), whereas increasing $p_{\text{high}}$ leads to a decrease in $r$ because the density of compressed hydrogen increases more rapidly than the loading in the zeolites. Clearly, exploring temperature swings is the most promising adjustment to the operating conditions. Overall, the high correlation of the data indicates that the ranking of the hydrogen storage performance depends more on the inherent properties of the NPM than on the specific choice of operating conditions. It should be also noted that our assumption of 70% NPM volume fraction in the tank is a conservative estimate and could be improved through improved manufacturing. Assuming an ideal scenario of in situ growth of NPMs in the fuel tank enclosure achieving 100% NPM volume, the working capacity for hydrogen storage can be increased to well above $r = 2$ for some of the NPMs (see fig. S8).

**Hydrogen adsorption fingerprints and heat of adsorption** To gain insights into the adsorption patterns and underlying reasons for high $T_{\text{max}}$ values, we investigate the fingerprint representations from the meta-learning model and the isosteric heat of adsorption for IZA and PCOD-syn zeolites and HCPs. The meta-learning model reconstructs the entire loading surface from the fingerprint vector containing only five coefficients where one positive value represents the estimated saturation capacity of the material and four values in $(-1,1)$ represent the shape of the loading surface. Thus, the fingerprint space contains full information about adsorption patterns and the temperature dependence of the working capacity. Figure 6A depicts the two-dimensional (2D) principal components analysis (PCA) of hydrogen adsorption fingerprints for NPMs in the IZA, PCOD-syn, and HCP databases. The ToBaCCo MOFs are not included in the PCA because, after retraining the meta-learning model for these MOFs, the same fingerprint vector describes a different loading surface. The position of an NPM in the principal component space shows a clear correlation with its $T_{\text{max}}$ and the zeolites on the Pareto front of the temperature-capacity distribution are located at the boundaries of the fingerprint distribution. This demonstrates that the meta-learning model is able to generate coherent fingerprint representations reflecting each NPM’s hydrogen adsorption behavior. It is also noteworthy that a consistent representation for all NPMs is not available using the best isotherm fitting method because different AIFs with different numbers of parameters are needed for different NPMs.

Although the fingerprint representation is entirely obtained from the hydrogen loading surface of an NPM without any structural information, strong correlations are still found between some of the fingerprint values and structural properties of all-silica zeolites in the IZA database (see table S3). The first fingerprint value, $z_0$, gives the ratio between the predicted saturation capacity, $q_{\text{sat}}$, and maximum observed loading, $q_{\text{max}}$ at $T = 77$ K and $p = 403.4$ bar. Its positive correlation with the accessible volume fraction, maximum diffusion and inclusion diameters, and negative correlation with framework density indicates that an NPM is less likely to be saturated at $T = 77$ K and $p = 403.4$ bar when it has a more open pore environment. The magnitudes of the Pearson’s correlation...
coefficient, \( r \), for \( z_0 \) range from 0.36 to 0.47. The weak correlation is likely due to the small SD of 0.05 compared to the mean of 1.05 observed for \( z_0 \) among the IZA zeolites. Fingerprint value \( z_1 \) shows a stronger correlation to the structural properties (0.68 < \( |r| \) < 0.79), and \( z_4 \) also shows a weak correlation (0.35 < \( |r| \) < 0.42). Motivated by the observed correlation of the fingerprint values to the structural parameters, a separate neural network is trained to decode the adsorption fingerprint into structural properties of the IZA zeolites (see fig. S9). As expected for a fingerprint representing the (volumetric) loading surface, this neural network is fairly accurate for the prediction of the accessible volume fraction, but the predictability for the other three structural parameters is only fair.

Measuring the heat of adsorption is an important way to characterize the adsorption behavior (14, 33, 34). For example, Bae and Snurr (33) demonstrated that an optimal heat of adsorption exists that yields the maximum working capacity for a given framework topology with varying sorbate-sorbent interaction strengths. As discussed earlier, high \( T_{\text{max}} \) values result from high loading at low pressures and low temperatures, meaning that strong adsorption sites exist in the NPM, and can be readily occupied by guest molecules at a low temperature due to the decreased entropic penalty. The isosteric heat of adsorption (in atomic units) can be obtained from a set of adsorption isotherms as follows, assuming ideal-gas behavior in the reservoir phase:

\[
Q_{st} = k_B T^2 \left( \frac{\partial \ln \rho}{\partial T} \right)_q = -\left( \frac{\partial \ln \rho}{\partial \beta} \right)_q
\]

where \( q \) is the loading, \( k_B \) is the Boltzmann constant, and \( \beta = 1/k_B T \). The isosteric heat of adsorption is calculated here using automatic differentiation from the meta-learning model, which symbolically transforms the continuous loading surface to find the gradients (35), whereas the energy/particle fluctuations (36) with the assumption of ideal-gas behavior in the reservoir phase are used to obtain \( Q_{st} \) from the molecular simulations. Given that the average and median \( T_{\text{max}} \) values for IZA and PCOD-syn zeolites are close to 92.4 K, the \( Q_{st} \) values are reported at this temperature. Although \( Q_{st} \) values at very low loading are commonly used in the analysis of adsorption behavior, following the lead of Fang et al. (34), we test here also \( Q_{st} \) values obtained at \( p_{\text{low}} = 2.71 \) bar and at the midpoint of \( p_{\text{low}} \) and \( p_{\text{high}} \) in logarithmic pressure space (\( p_{\text{mid}} = 7.39 \) bar). Our data show that \( p_{\text{low}} \) yields \( Q_{st} \) values that are best correlated with \( T_{\text{max}} \) (see Fig. 6B and fig. S10).

Encouragingly, the meta-learning model can also predict \( Q_{st} \), a derivative property, with remarkable accuracy as shown by comparison with the “true” heat of adsorption obtained from the molecular simulation trajectories (see Fig. 6B and fig. S10). The strong positive correlation of \( T_{\text{max}} \) and \( Q_{st} \) indicates that the existence of a sufficient number of strong-binding sites occupied at \( p_{\text{low}} \) is mostly responsible for high \( T_{\text{max}} \) values. Application of the meta-learning model also allows one to obtain a heatmap of \( Q_{st} \) as a function of pressure and temperature (see fig. S11). Here, it is observed that \( Q_{st} \) increases with temperature at constant pressure and decreases with pressure at constant temperature, i.e., increasing the loading decreases \( Q_{st} \) as should be expected for systems with stronger sorbate-sorbent than sorbate-sorbate interactions.

**Experimental validation**

Our predictions for \( T_{\text{max}} \) and the temperature-dependent working capacity obtained from the meta-learning model and the molecular simulations are supported through comparison with experimental results for a specific zeolite. Figure 7 shows the predicted and experimental hydrogen adsorption isotherms for a calcium-exchanged LTA zeolite with 1:1 Si-to-Al ratio (LTA-Ca), and its working capacity as a function of temperature. At higher temperatures (\( T > 130 \) K), the predicted adsorption isotherms agree well with the experimental data. At lower temperatures, however, the predictions are not as satisfactory. Because the meta-learning isotherms remain fairly close to the molecular simulation data (see Fig. 7), the disagreement is mostly due to shortcomings of the force fields with a minor contribution from the simulation approach. The simulation data show a stronger temperature effect for the low-pressure hydrogen adsorption, leading to a substantial overprediction of the loading at \( T = 77 \) K and \( p < 10 \) bar. In addition, the loading at high pressure and low temperature is somewhat overpredicted. This indicates that (i) the hydrogen-Ca binding energy is too favorable and (ii) the accessible pore volume is too large. Here, we explore the effect of increasing the Lennard-Jones diameter of the Ca cation because increasing it will both reduce the hydrogen-Ca binding energy and the accessible pore volume. The data shown in fig. S12 indicate that setting \( \sigma_{\text{Ca}} = 3.6 \) Å yields adsorption loadings that are in better agreement with the experimental data at 77 K and \( p = 2.71 \) and 30.0 bar, but the working capacity and \( T_{\text{max}} \) are not much affected. Further tuning of...

---

**Fig. 5. Hydrogen storage predictions for zeolite RWY.** Working capacities (top) and adsorption isobars (bottom) obtained from short high-throughput and long validation simulations and predicted using ensembles of eight meta-learning models and of multilayer perceptrons (MLP) with two hidden layers of 16 and 8 units and using the Sips isotherm with Arrhenius temperature dependence.
other force field parameters would be needed to improve the match to the experimental data (e.g., reducing the partial charge of the Ca cation would reduce the hydrogen-Ca binding energy but has less impact on the accessible pore volume). Explicitly accounting for nuclear quantum effects may also be beneficial because it would lead to stronger hydrogen-Ca binding at higher temperature and, hence, reduce the temperature dependence of the low-pressure loading.

Although the meta-learning model trained on all-silica zeolites (i.e., not encountering any zeolite with mobile cations) overestimates the maximum working capacity for LTA-Ca compared with the simulation results, it remains very accurate for predicting $T_{\text{max}}$. The already impressive performance of the meta-learning model trained only on all-silica zeolites could be further improved through inclusion of data for a large number of cation-exchanged zeolites in the training set or by fine-tuning the current model using a smaller number of cation-exchanged zeolites as done here for the MOF training set or by fine-tuning the current model using a smaller number of cation-exchanged zeolites.

Compared with the all-silica form, the strong hydrogen–cation interactions lead to a larger $Q_{\text{st}}$ for LTA-Ca (see fig. S11). Thus, LTA-Ca exhibits a much higher $T_{\text{max}}$ (by about 50 K) but also retains about 70% of the working capacity, which, in turn, moves LTA-Ca to the Pareto front (see Fig. 3). The molecular simulation data at 133.1 K ($\approx T_{\text{max}}$) indicate loadings of 0.82 and 1.63 hydrogen molecules per Ca cation at $p_{\text{low}} = 2.71$ bar and $p_{\text{high}} = 30.0$ bar, respectively. Thus, at $T_{\text{max}}$, each of the strong adsorption sites is approximately singly occupied. Prompted by the shift of LTA-Ca to the Pareto front, we also carry out simulations and use the meta-learning model for other cations in the LTA zeolite with 1:1 Si-to-Al ratio. As illustrated in fig. S13, replacing Ca cations with smaller Mg or Li cations leads to an increase of the working capacity, whereas the larger K cation yields a reduction in working capacity. The $T_{\text{max}}$ predictions indicate no significant difference between these four cation-exchanged LTA zeolites.

**DISCUSSION**

This work introduces a meta-learning model that can accurately predict the hydrogen loading surface of NPMs over a wide range of temperatures and pressures. The meta-learning model achieves a much better few-shot prediction performance than fitting the best adsorption isotherm out of a few candidates for each sorbent material and is able to extrapolate to a limited extent beyond the temperature range of the training data. Despite the model being trained on only all-silica zeolites in the IZA database, it accurately predicts the hydrogen adsorption in hypothetical zeolites and HCPs and can readily be tuned for MOFs with larger pores. Besides the meta-learning predictions, the model also yields a fingerprint representation for each NPM that encodes its adsorption behavior. These fingerprints also allow for application of other dimensionality reduction methods, such as PCA, to find regions for promising materials in the PCA space. While the meta-learning model can be used for few-shot prediction, other dimensionality reduction approaches require the number of samples for each material to be fixed.

Exemplified by vehicular hydrogen storage, many real-world applications of NPMs involve adsorption at a range of conditions, particularly in process optimization (37). While it would be prohibitively expensive to perform high-throughput computation in the joint space of materials, chemical structures, and process conditions, meta-learning provides a route toward overcoming this difficulty by making the maximum use of available data. That is, the meta-learning model can be trained on a subset of representative materials with data covering the complete condition space and then be applied via few-shot predictions to the complete set of materials.

A few limitations of the present work should be noted. The task distribution in our meta-learning setting covers multiple NPMs but only for the same adsorbate molecule, namely, hydrogen. Thus, it remains an open question whether the same meta-learning model can also be generalized to the adsorption for a different adsorbate molecule, say, methane. Although meta-learning allows for few-shot prediction and extrapolation in the $p$, $T$ state space, applying a pre-trained model to materials with largely different structures from the training set still requires fine-tuning the model. For example, the model trained on all-silica zeolites has to be fine-tuned for larger-pore MOFs. Furthermore, extending the meta-learning model to multicomponent adsorption data important for chemical separations (38, 39) may markedly increase the problem complexity and require judicious selection of representative training data (e.g., placing data points along edges and/or diagonals in the higher-dimensional space).
The experimental, molecular simulation, and machine learning results presented here demonstrate that some sorbent materials attain their maximum hydrogen storage capacity (using filling and depletion pressures of 30.0 and 2.71 bar, respectively, in a tank containing 30% void volume) at temperature as high as 140 K and retain 90% of their maximum capacity up to 165 K. The capability to use a higher working temperature (compared to 77 K often used for adsorption measurements) may make application of sorbent-based storage system in vehicles more feasible. The meta-learning predictions also indicate that incorporating a modest temperature swing is beneficial for increasing the working capacity than increasing the filling pressure or decreasing the depletion pressure. We also found a Pareto distribution between the maximum working capacity and the optimal temperature for a diverse set of NPMs. Thus, it is of great interest to find and design NPMs that extend the Pareto front of hydrogen storage. On the basis of the computational and experimental results for the all-silica and cation-exchanged forms of the LTA zeolite framework, selecting already high-performing framework topologies and pore sizes from the all-silica zeolite dataset and investigating their cation-exchanged forms may be a promising way to achieve this goal. However, it should be noted that gas adsorption in cation-exchanged zeolites and other porous adsorbents that more strongly adsorb hydrogen through favorable first-order electrostatic interactions will be much more sensitive to impurities, especially water. Therefore, using these materials in vehicular hydrogen storage systems will also impose a higher standard on powertrain and fuel system engineering, for example, in water management (40).

**MATERIALS AND METHODS**

**Meta-learning model architecture**

In a meta-learning problem, the error of the meta-learning model is minimized over a distribution of tasks (24) to obtain the model’s parameters $\hat{\theta}$ as

$$\hat{\theta} = \arg\max_{\theta} \mathbb{E}_{D \sim p(D)}[L_{q}(D)]$$

(3)

where $D$ is the dataset of each learning task that follows the distribution $p(D)$, and $L_{q}(D)$ is the error on the dataset using parameters $\theta$. Specifically for unary adsorption in NPMs, each learning task is to predict the adsorption loading, $q$, in one particular material from the set of state variables. Therefore, a dataset $D = \{x, q\}$ consists of a collection of thermodynamic states $x = (p, T)$ and the corresponding adsorption loadings $q$, and $p(D)$ indicates the distribution of adsorption behavior in the entire materials space of interest. In contrast to base learning models operating on individual entries in a dataset $q = f(x)$, a meta-learning model also uses a training set comprising multiple variable-output pairs as its input (Fig. 1A). Using this approach, the adsorption in any material at any state point can be predicted using a set of example adsorption data and the unknown state to predict

$$\hat{q} = f(D_{\text{train}}, x)$$

(4)

The meta-learning formulation is different from common multi-task learning models that use multiple output units to predict different tasks $\{\hat{q}_1, \cdots, \hat{q}_n\} = f(x)$. For those models, more output units need to be added to predict a new material. In contrast, predicting a new material in meta-learning only requires a different $D_{\text{train}}$ sample.

The encoder network $f_{e}(\cdot)$ produces a latent vector $z$ from the task-specific training set, $D_{\text{train}}$ (adsorption data in one NPM)

$$z = f_{e}(D_{\text{train}})$$

(5)

Similar to other encoder-decoder architectures for meta-learning (41), the encoder network used a pointwise feed-forward architecture followed by a global pooling operation so that the output $z$ is invariant to the permutation of input points, and $f_{e}(\cdot)$ is able to operate on point sets with variable sizes. Apart from meta-learning, this architecture is also present in the PointNet model for 3D object detection (42).

The decoder network $f_{d}(\cdot)$ is a feed-forward network predicting the adsorption loading whose input includes the latent variables encoded in $z$ for a specific task/adsorption system and state variables as the input vector $x = (p, T)$

$$\hat{q} = f_{d}(z, x)$$

(6)

Regularization of the neural network is implemented by penalizing the correlation among dimensions in the latent space. The task-specific loss is formulated in a semisupervised manner as to both reconstruct the training examples and predict test examples

$$L_{q}(D) = \sum_{(x, q) \in D} [q_i - f_{d}(f_{e}(D_{\text{train}}), x_i)]^2 + \lambda \sum_{(i, j) \in E} \text{cov} [f_{e}(D_{\text{train}})]_{ij}$$

(7)
where the training set $D_{train}$ is a subset of the full task-specific dataset $D$, and $\text{cov} [ \cdot ]$ denotes the covariance matrix. $\lambda$ is the hyperparameter controlling the strength of regularization. In practice, $q$ and $\hat{q}$ are normalized to the maximum loading observed in each NPM for all state points; therefore, the model does not have information about the saturation loading, and a single simulation or experiment for a new material is necessary to determine the normalization factor. Implementation details for the meta-learning model are provided in section S2.

### Molecular simulations

Monte Carlo simulations in two types of thermodynamic ensembles were used to calculate the hydrogen adsorption in NPMs. Isothermal-isobaric Gibbs ensemble Monte Carlo (NpT-GEMC) simulations (43) were carried out to calculate the hydrogen adsorption for zeolites in the IZA and PCOD-syn databases, and grand canonical Monte Carlo (GCMC) simulations (44) were carried out for MOFs and HCPs. While the reservoir pressure can be explicitly specified in GEMC simulations, GCMC simulations are performed at equivalent pressures as the GEMC simulations by matching the chemical potential in the hydrogen vapor phase. For each NPM, simulations are carried out for 64 combinations of eight temperatures and eight pressures: $T = 77.00, 92.40, 110.88, 133.06, 159.67, 191.60, 229.92,$ and 275.90 K and $p = 1.00, 2.71, 7.39, 20.09, 30.00, 54.60, 148.4,$ and 403.4 bar. The hydrogen molecules are represented by a model using two Lennard-Jones sites at the positions of the hydrogen nuclei and three partial charges to account for hydrogen's quadrupole moment, two Lennard-Jones sites at the positions of the hydrogen nuclei and three partial charges to account for nuclear quantum effects in the near-critical region. Additional force field and simulation details are provided in section S3.

### Measurement of hydrogen isotherms

A sample of zeolite LTA-Ca was provided by Chemiewerk Bad Köstritz (Köstrolith 5AK) and was used as received. Hydrogen gas with a purity of 5.0 (99.999%) was purchased from Linde AG, business domain (Köstrolith 5AK) and was used as received. Hydrogen gas with a purity of 5.0 (99.999%) was purchased from Linde AG, business domain Linde Gas, and was used without further purification. The hydrogen sorption isotherms were recorded with a home-built closed cycle sorption unit equipped with an Advanced Research Systems closed cycle cryostat and a Lakeshore high-precision temperature controller. Using this setting, isotherms between 77 and 195 K and up to a pressure of 30 bar were recorded. Before the adsorption measurement, 1.04 g of the zeolite sample was activated in vacuum at a temperature of 523 K for 24 hours. After transferring to the measurement cell, the sample was evacuated for another 12 hours at room temperature. Helium expansion for determination of the sample volume was conducted up to a pressure of 300 mbar. Each point of the isotherm was calculated with the ideal gas law using the compression factors of hydrogen for the respective temperature and pressure given by the National Institute of Standards and Technology Chemistry WebBook, (45).

### Temperature-dependent AIFs

As a comparison to the meta-learning model, temperature-dependent AIFs are also used to predict hydrogen loading for all materials. AIFs studied here include the widely used Langmuir, Sips, quadratic, and DSL isotherms (46–48). Their functional forms, temperature-dependent parameters, and the total number of isotherm coefficients are summarized in section S4. The temperature dependence for the equilibrium constant is described using the Arrhenius equation (49), giving two isotherm coefficients per temperature-dependent parameter. To account for the variety in adsorption behavior, the AIF yielding the smallest MSE for a given material is selected for the most stringent test of the performance of the meta-learning model.

### Fuel tank model

For a realistic description of hydrogen storage in a car, the fuel tank is modeled as a mixture of packed sorbent particles and void spaces filled with hydrogen gas (31). The total volumetric loading in the fuel tank at pressure $p$ and temperature $T$ is

$$n_{tot}(p, T) = (1 - \epsilon) q(p, T) + \epsilon p(p, T)$$

where $q(p, T)$ is the volumetric loading for the NPM, $p(p, T)$ is the density of hydrogen gas, and $\epsilon$ is the void fraction. At the adsorption (filling) and desorption (depletion) conditions, it is assumed that the fuel tank is in thermodynamic equilibrium. Therefore, the working capacity is obtained from a difference of the loadings at two state points involving a decrease in pressure and, in some cases, an increase in the temperature by $\Delta T$

$$\Delta n(p_{high}, p_{low}, T, \Delta T) = n_{tot}(p_{high}, T) - n_{tot}(p_{low}, T + \Delta T)$$

On the basis of reducing the infrastructure requirements for fuel stations and the safety concerns for high-pressure tanks, the working pressures are chosen as $p_{high} = 30.0$ or 54.6 bar and $p_{low} = 2.71$ or 1.00 bar. $\epsilon = 0.3$ is used as a modest estimate of the packing efficiency. The working capacities for all NPMs investigated here are predicted with the meta-learning model at temperatures between 45 and 276 K using a step size of 0.1 K. Then, the material-specific optimal temperature for hydrogen storage can be found as

$$T_{max} = T + \frac{\Delta T}{2} = \text{argmax}_{T} [n_{tot}(p_{high}, T) - n_{tot}(p_{low}, T + \Delta T)]$$

A linear search is used to implement the optimization to 0.1-K precision.

### Supplementary Materials

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/30/eabb3983/DC1

### References and Notes

1. M. A. Hannan, F. A. Azidian, A. Mohamed, Hybrid electric vehicles and their challenges: A review. Renew. Sustain. Energy Rev. 29, 135–150 (2014).
2. L. Schlapbach, Hydrogen-fuelled vehicles. Nature 460, 809–811 (2009).
3. A. Yamashita, M. Kondo, S. Goto, N. Ogami, SAE 2015 World Congress and Exhibition (2015).
4. M. Felderhoff, C. Weidenthaler, R. von Helmolt, U. Eberle, Hydrogen storage: The remaining scientific and technological challenges. Phys. Chem. Chem. Phys. 9, 2643–2653 (2007).
5. J. Weitkamp, M. Fritz, S. Ernst, Zeolites as media for hydrogen storage. Int. J. Hydrogen Energy 20, 967–970 (1995).
6. Y. Ye, C. C. Ahn, C. Witham, B. Fultz, J. Liu, A. G. Rinzler, D. Colbert, K. A. Smith, R. E. Smalley, Hydrogen adsorption and cohesive energy of single-walledcarbon nanotubes, Appl. Phys. Lett. 74, 2307–2309 (1999).
7. L. J. Murray, M. Dincă, J. R. Long, Hydrogen storage in metal–organic frameworks. Chem. Soc. Rev. 38, 1294–1314 (2009).
8. Z. Chen, P. Li, R. Anderson, X. Wang, X. Zhang, L. Robison, L. R. Redfern, S. Moribe, T. Islamoglu, D. A. Gómez-Gualdrón, T. Yidirim, J. F. Stoddart, D. K. Farha, Balancing volumetric and gravimetric uptake in highly porous materials for clean energy. Science 368, 297–303 (2020).
9. C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, Large-scale screening of hypothetical metal–organic frameworks. Nat. Chem. 4, 83–89 (2012).
Exceptional hydrogen storage achieved by screening nearly half a million metal–organic frameworks. J. Mater. Chem. A 1688–1700 (2006).

A. Rajendran, R. T. Maruyama, H. O. R. Landa, A. Seidel-Morgenstern, Modelling binary non-linear chromatography using discrete equilibrium data. Adorption 26, 973–987 (2020).

G. Anderson, B. Schweitzer, R. Anderson, D. A. Gómez-Gualdrón, Attainable volumetric targets for adsorption-based hydrogen storage in porous crystals: Molecular simulation and machine learning. J. Phys. Chem. C 123, 120–130 (2019).

C. M. Bishop, Pattern Recognition and Machine Learning (Springer, 2006).

P. Buzádi, C. Giraud-Carrier, C. Soares, R. Vilalta, Metalearning: Applications to Data Mining (Springer, 2008).

R. Vilalta, Y. Drissi, A perspective view and survey of meta-learning. Artif. Intel. Rev. 18, 77–95 (2002).

A. A. Rusu, D. Rao, J. Sygnowski, O. Vinyals, R. Pascanu, S. Osindero, R. Hadsell, Meta-learning with latent embedding optimization. arXiv:1807.05960 (2018).

C. Finn, P. Abbeel, S. Levine, Proceedings of the 34th International Conference on Machine Learning (2017), vol. 70, pp. 1126–1135.

I. Goodfellow, Y. Bengio, A. Courville, Deep Learning (MIT Press, 2016).

B. J. Bucior, N. S. Bobbitt, T. Islamoglu, S. Goswami, A. Gopalan, T. Yildirim, O. K. Farha, N. Bagheri, R. Q. Snurr, Energy-based descriptors to rapidly predict hydrogen storage in metal–organic frameworks. Mol. Syst. Des. Eng. 4, 162–174 (2019).

C. Baerlocher, L. McCusker, Database of zeolite structures, www.iza-structure.org/ databases/ [accessed 25 April 2019].

J. L. S. Perez, M. Haranczyk, N. E. R. Zimmermann, High-throughput assessment of hypothetical zeolite materials for their synthesizability and industrial deployability. Z. Kristallogr. 234, 437–450 (2019).

Y. J. Colón, D. A. Gómez-Gualdrón, R. Q. Snurr, Topologically guided, automated construction of metal–organic frameworks and their evaluation for energy-related applications. Cryst. Growth Des. 17, 5801–5810 (2017).

G. Kugan, T. P. Liyana-Arachchi, C. M. Colina, Pore size tuning of poly(styrene-co-vinylbenzyl chloride-co-divinylbenzene) hypercrosslinked polymers: Insights from molecular simulations. Polymer 99, 173–184 (2016).

H. Zhang, P. Deria, D. K. Farha, J. T. Hupp, R. Q. Snurr, A thermodynamic tank model for studying the effect of higher hydrocarbons on natural gas storage in metal–organic frameworks. Energ. Environ. Sci. 8, 1501–1510 (2015).

N. S. Bobbitt, J. Chen, R. Q. Snurr, High-throughput screening of metal–organic frameworks for hydrogen storage at cryogenic temperature. J. Phys. Chem. C 120, 27328–27341 (2016).

Y. S. Bae, R. Q. Snurr, Optimal isoster heat of adsorption for hydrogen storage and delivery using metal–organic frameworks. Microporous Mesoporous Mater. 132, 300–303 (2010).

H. Fang, A. Kulkarni, P. Kamatoki, R. Awati, P. I. Ravikovitch, D. S. Sholl, Identification of high-CO2-capacity cationic zeolites by accurate computational screening. Chem. Mater. 28, 3837–3896 (2016).

B. Steiner, Z. DeVito, S. Chintala, S. Gross, A. Paszke, F. Massa, A. Lerer, G. Chanan, Z. Lin, E. Yang, A. Desmaison, A. Tejani, A. Kopf, J. Bradbury, L. Antiga, M. Raison, N. Gilmer, S. Chilamkurthy, T. Killeen, L. Fang, J. Bai, Thirty-third Conference on Neural Information Processing Systems (NeurIPS 2019) (2019).

T. J. H. Vlaeg, E. García-Pérez, D. Dubbeldam, S. Ban, S. Calero, Computing the heat of adsorption using molecular simulations: The effect of strong Coulombic interactions. J. Chem. Theory Comput. 4, 1107–1118 (2008).
68. C. D. Wick, M. G. Martin, J. I. Siepmann, Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and alkylbenzenes. J. Phys. Chem. B 104, 8008–8016 (2000).

Acknowledgment

Funding: This research was primarily supported by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under award DE-FG02-17ER16362. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under contract DE-AC02-06CH11357. Additional computer resources were provided by the Minnesota Supercomputing Institute at the University of Minnesota, by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology, and by the Quest high-performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology. The experimental work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project ID 416229255, SFB 1411. We thank C. Bunner for the development of the three-site model for hydrogen and T. Yang for help with the equation of state calculations. Support through a Robert and Jill DeMaster Excellence Fellowship (Y.S.) is also gratefully acknowledged. Author contributions: J.I.S. and M.H. conceived and directed the project. Y.S. developed the meta-learning model. Y.S., R.F.D., and J.I.S. (all-silica and cation-exchanged zeolites), Z.L. and R.Q.S. (MOFs), and D.T., D.S.S., and C.M.C. (HCPs) performed and analyzed the molecular simulations for hydrogen adsorption. S.G., M.T., and M.H. performed and analyzed the hydrogen adsorption measurements. Y.S. and J.I.S. performed the analysis of optimal hydrogen storage conditions and adsorption fingerprints, prepared figures and tables, and wrote the manuscript with contributions from other authors. All authors discussed the results and commented on the manuscript. Competing interests: R.Q.S. has a financial interest in the startup company NuMat Technologies, which is seeking to commercialize MOFs. The authors declare that they have no other competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The Python program implementing the meta-learning model, molecular simulation data, and meta-learning predictions for the hydrogen adsorption loading in all NPMs and for all state points, and scripts to generate figures and tables reported are available from the Data Repository for U of M (DRUM) at the persistent link: https://hdl.handle.net/11299/220168.

Submitted 10 January 2021
Accepted 4 June 2021
Published 21 July 2021
10.1126/sciadv.abg3983

Citation: Y. Sun, R. F. DeJaco, Z. Li, D. Tang, S. Glante, D. S. Sholl, C. M. Colina, R. Q. Snurr, M. Thommes, M. Hartmann, J. I. Siepmann, Fingerprinting diverse nanoporous materials for optimal hydrogen storage conditions using meta-learning. Sci. Adv. 7, eabg3983 (2021).
Fingerprinting diverse nanoporous materials for optimal hydrogen storage conditions using meta-learning
Yangzesheng Sun, Robert F. DeJaco, Zhao Li, Dai Tang, Stephan Glante, David S. Sholl, Coray M. Colina, Randall Q. Snurr, Matthias Thommes, Martin Hartmann and J. Ilja Siepmann

Sci Adv 7 (30), eabg3983.
DOI: 10.1126/sciadv.abg3983