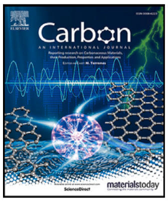
Carbon 257 (2026) 121713

Contents lists available at ScienceDirect

Carbon

journal homepage: www.elsevier.com/locate/carbon



Research paper

Porosity Local Analysis (PoLA) for nanoporous carbons. Porous volume characterization and prediction of gas adsorption isotherms

Maurizio Cossi [a,b],*, Alberto Zoccante [a,b], Valeria Palumbo [a,b], Maddalena D'Amore [a,b], Fernando Vallejos-Burgos [c], Radovan Kukobat [d], Alejandro Moreo [e], Fabrizio Sebastiani [e], Stefano Brandani [f], Enzo Mangano [f], Federico Begni [a,b], Giorgio Celoria [a,b], Leonardo Marchese [a,b],*

[a] Dipartimento di Scienze e Innovazione Tecnologica, Universita del Piemonte Orientale, Viale T. Michel, 11, Alessandria, I-15121, Italy
[b] Centro di Ricerca per il Risanamento e la Protezione Ambientale (RiSPA), Joint-Lab DISIT/Syensqo, Viale T. Michel, 11, Alessandria, I-15121, Italy
[c] Morgan Advanced Materials, Carbon Science Centre of Excellence, 310 Innovation Blvd. Ste 250, State College, PA 16803, USA
[d] Department of Chemical Engineering and Technology, University of Banja Luka, B.V. Stepe Stepanovica 73, Banja Luka, 78000, Bosnia and Herzegovina
[e] Istituto di Scienza e Tecnologie dell'Informazione, CNR, Via G. Moruzzi, 1, Pisa, I-56124, Italy
[f] School of Engineering, University of Edinburgh, Edinburgh, UK

ARTICLE INFO

Keywords:
Characterization of nanoporous carbons
Porous volume distribution
Prediction of gas adsorption isotherms
Machine learning regression

ABSTRACT

We introduce a new computational procedure called PoLA (Porosity Local Analysis), a point-by-point description of the void space in nanoporous materials that surpasses the conventional representation of pores as homogeneous regions of regular geometry (spheres, cylinders, slits). Each volume element is assigned its own porous character through the minimum distance to opposite walls (*MinD*), a quantity directly linked to the host–guest interaction potential and therefore to physisorption behaviour. We apply PoLA to a dataset of 109 atomistic carbon models and correlate the resulting V(*MinD*) distributions with N_2 and H_2 adsorption isotherms at 77 K, simulated by Grand Canonical Monte Carlo. A purpose-built machine learning procedure, based on optimized neural networks, infers V(*MinD*) from a nitrogen isotherm and predicts the corresponding hydrogen uptake. Validation against four commercial activated carbons (Norit Row, Maxsorb, BAX1700, CGF4) shows excellent agreement between predicted and measured H_2 isotherms up to 60 bar, demonstrating that PoLA provides both a transferable porosity descriptor and a predictive tool for adsorbent design.

1. Introduction

Nanoporous solids are studied with great interest for many applications in gas storage and purification [1–11], catalysis [12–18], extraction of trace pollutants from air, water or soil [19–22]. In particular, porous activated carbons are excellent candidates for these applications due to their high affinity for most molecules of interest, ensuring optimal physisorption efficiencies. In addition, carbons can be produced in large quantities at relatively low cost, also from agricultural or industrial wastes, coupling environmental and economical advantages [23–26]. Typically, porous carbons display a large morphological variety, with a strong non-crystalline (amorphous) character.

The efficiency of the physisorption processes relies on the strong interactions between the molecules and the solid walls, enhanced by the high specific surface and the capillary effects in the network of nanosized pores and channels. Clearly, the distribution of the porous volume is of paramount importance to describe such interactions, and possibly to predict the behaviour of a porous solid towards a specific adsorbate.

According to IUPAC [27], "pores" can be identified by their "width": a common classification defines the pores as ultramicro below 7 Å, micro between 7 and 20 Å, meso between 20 and 50 Å and macro above 50 Å. But "pore width" is not further specified, and in many cases its meaning is not evident, except for crystalline solids with regular and geometrical void spaces. For amorphous carbons, the very concept of a pore can be ill-defined, and one should rather refer to a "porous volume" with no clear geometrical shape, which is sometimes described better in terms of a fractal development [28].

Nonetheless, the most used porosity descriptor even for amorphous materials is the so-called Pore Size Distribution (PSD) [29–33], which nowadays is determined experimentally by adsorbing a probe gas (typically, N_2 at 77 K, Ar at 87 K or CO_2 at 273 K) and then fitting

---

* Corresponding authors.
E-mail addresses: maurizio.cossi@uniupo.it (M. Cossi), leonardo.marchese@uniupo.it (L. Marchese).

https://doi.org/10.1016/j.carbon.2026.121713
Received 5 March 2026; Received in revised form 10 May 2026; Accepted 27 May 2026

the adsorption isotherms with DFT (density functional theory) models [34–37]. Though the DFT itself is not constrained to specific space shapes, even in its application to porous solids, all the available models applied to fit the experimental isotherms (non local, NL-DFT, quenched solid, QS-DFT, or 2D-DFT) [38–44] are based on a database of ideal, regular solids (spheres, cylinders or parallel slits): the pore sizes exactly defined for these models (the diameter for spheres and cylinders, the distance between opposite planes for slits) are combined according to the fit to provide the PSD for the sample.

This largely adopted approach became a standard for the characterization of porous materials, for it is simple and fast to use, and provides a recognizable metric for very different systems. However, for the reasons exposed above, the so-obtained PSD is sometimes poorly related to the actual character of the empty volume inside the material, especially in case of strongly amorphous systems as, for instance, activated carbons.

A similar approach is often applied to atomistic models too: for instance, the popular PoreBlazer [45–47] code provides a PSD for porous solid models, based on a collection of overlapping spheres filling the inner voids, and also the Zeo++ model [48,49] describes the porosity of crystalline and amorphous solids starting from spheres inscribed within the empty volumes. Gelb and Gubbins defined the PSD of mesoporous models of silica glasses as the collection of spherical pores, comparing the results to the Barrett–Joyner–Halenda (BJH) analysis of adsorption isotherms simulated in the same models by Monte Carlo [50]. In another work, dealing with the definition of specific porous surfaces [51], the same authors noted that the surface monolayer density is strongly affected by the curvature of the surface, a kind of local effect similar to those described in the following.

The criticism about the use of PSD based on geometric solids motivated to develop 3D-Vis [52,53], a procedure using a large collection of atomistic carbon structures, where nitrogen adsorption isotherms were simulated theoretically, both to fit experimental isotherms and to provide a plausible 3D structure for the experimental samples. Though very useful to correlate the isotherms to possible carbon structures, this approach does not provide any metric to compare different systems.

We have recently presented an alternative description of the porous volume [54], which we believe is more useful to characterize porous carbons as well as other systems like amorphous silica or defective zeolites, where the geometric definition of pores might be less appropriate. Our approach, called PoLA (Porosity Local Analysis) is not based on pores with regular shape, but rather on the point-by-point analysis of the inner void volume, leading to the characterization of the porous volume (V) as a function of the “minimum distance from opposite walls” (*MinD*): note that in Ref. [54] the output of PoLA was called “porous volume distribution (PVD)”, corresponding to the V(*MinD*) used in the present paper. The details of the method are briefly summarized in Section 2, where we show how to obtain the function V(*MinD*).

PoLA is a fast and efficient tool to characterize atomistic models of porous materials and to compare the porosity of very different solids. Even more importantly, V(*MinD*) can be correlated to the adsorption isotherms of various gases, leading to efficient predictions of the experimental physisorption, as we show in the present work.

In this sense, PoLA has a two-fold range of applications: it can be used to describe the porous volume inside atomistic models with greater accuracy than in previous methods, and it can also be applied to real porous materials. In these latter applications, PoLA relies on experimental gas adsorption, like the above-mentioned approaches based on DFT: unlike these methods, however, PoLA does not fit the isotherms on a set of geometrical pores, but performs a machine learning (ML) multiple regression on a dataset of adsorbent models.

We have already shown [54] the correlation between PoLA volume characterization and N_2 adsorption isotherms simulated by Grand Canonical Monte Carlo (GCMC) for a large set of carbon models: a Random Forest ML algorithm, trained with PoLA and GCMC data, predicted the nitrogen isotherms starting from PoLA volumes with a very good precision.

In the present paper, we introduce the reverse procedure: starting from a N_2 adsorption isotherm, a ML procedure provides the PoLA-like V(*MinD*), which can be used to characterize experimental samples with the same metric used for atomistic models. To reinforce the use of PoLA description, instead or besides other porosity definitions like DFT-based PSD, we will show that the same ML algorithm is able to predict the experimental adsorption isotherms of H_2 at 77 K for a number of porous carbons. To do that, a large data set of carbon models has been used, combining the sets already presented in Refs. [52] and [54], and a new neural network ML algorithm has been defined, as described in the following.

The idea of using ML to correlate the structure of porous adsorbents and the adsorption of different gases has been effectively exploited in recent years: [55] in a popular work, Snurr et al. compared different algorithms to predict the adsorption of gas mixtures in a large number of metal organic frameworks (MOF) [56]. Similarly, the adsorption of organic molecules in a huge number of MOF models was computed by GCMC and predicted by ML [57], and recently Wu and Yu have presented a procedure to correlate the adsorption of N_2 and Ar in several MOF on the basis of classical DFT simulations [58]. Most of these applications deal with MOF, which are highly ordered microporous materials, unlike amorphous carbons, and they concern the prediction of isotherms more than the description of the porosity: nonetheless, these examples illustrate the great potential of the ML analysis in this field.

2. PoLA: Porosity Local Analysis

For physisorption applications, the most important characteristic of a porous adsorbent is the shape of the interaction potential felt by the adsorbate molecules in the various points of the inner space. As shown in Fig. 1, if the adsorbate can interact strongly with more than one wall at the same time, the interaction potential appears, at least in one direction, as a single well with a deep minimum (Fig. 1-A); if the walls get further, the potential becomes a double well (Fig. 1-B), and if the adsorbate interacts with just one wall, the potential looks like a typical Morse function (Fig. 1-C) [59,60].

The PoLA algorithm evaluates for each volume element the minimum distance from opposite walls (*MinD*) and uses this number to characterize the porous nature of that element. The V(*MinD*) is hence linked to the shape of the adsorbate-solid interaction potential, and it can therefore be correlated to the gas adsorption isotherms, as mentioned in the Introduction.

To be applied to an atomistic model, PoLA needs the Cartesian coordinates of the atoms along with the three periodic lengths (an orthogonal cell is assumed: other periodic cells can be analysed, provided the atoms are wrapped inside an orthogonal supercell). As illustrated in Fig. 2, the cell is divided in N small cubic elements (blocks) with edge l, so that Nl^3 equals the total cell volume; every block containing an atom is marked as “filled”, as well as all the blocks whose centre falls inside the van der Waals sphere of any atom. To avoid the appearance of spurious voids, the blocks belonging to small void regions trapped within a filled area, without connection with the rest of the empty volume, are also considered filled. The remaining empty blocks constitute the porous volume to be analysed further. Reducing the block edge l leads to a more detailed analysis: we have verified that $l = 1$ Å is a good compromise between accuracy and efficiency, as all the relevant quantities remain stable even with smaller blocks [54].

As shown in Fig. 3, for each void block the code determines the minimum distance from opposite walls: 120 different directions are searched, scanning evenly the spherical angle of an hemisphere centred on the block, and along each direction the blocks are checked until a filled one is found (that is, a wall has been touched). Then the opposite direction is immediately searched (flipping the solid angle)

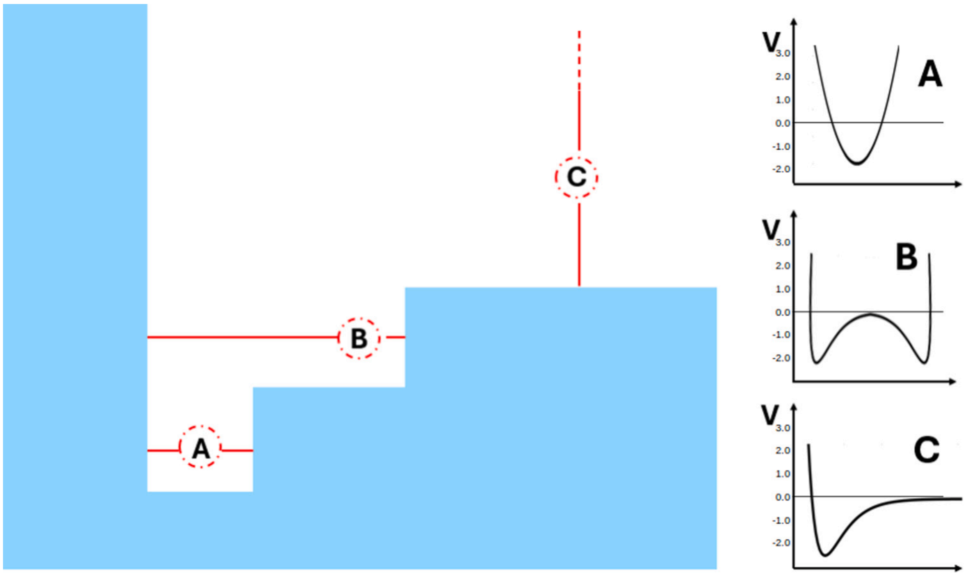


Fig. 1. Classification of volume elements and interaction potentials: (A) ultramicroporous or microporous, (B) mesoporous, (C) macroporous.

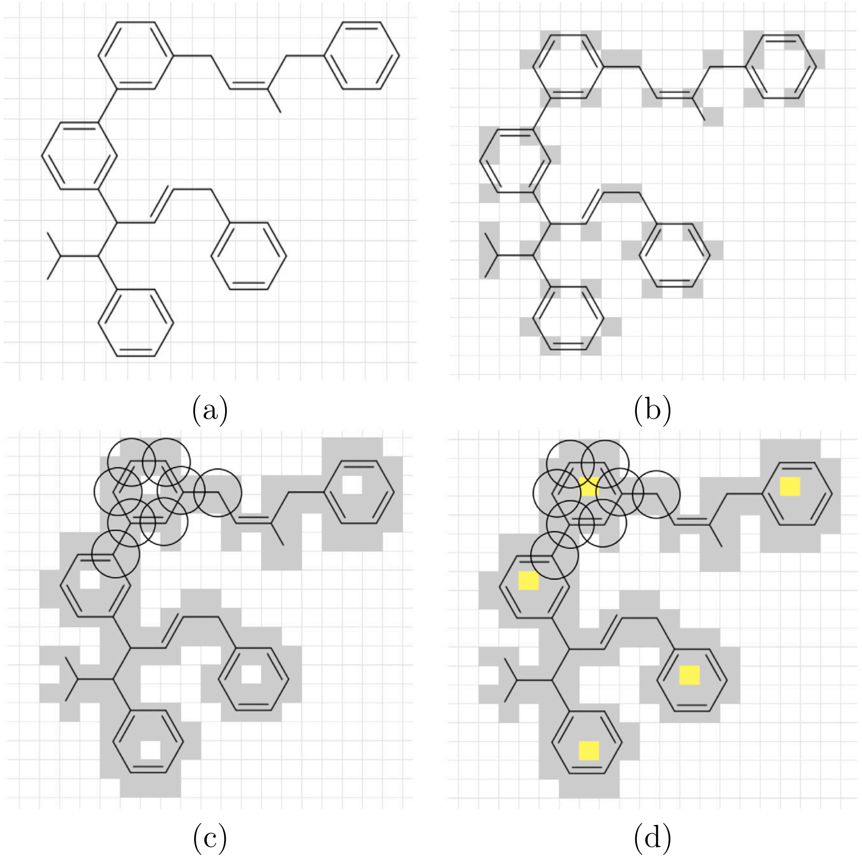

(a) (b) (c) (d)

Fig. 2. Determination of porous volume elements (empty blocks): (a) the atomistic model is embedded in a 3D grid of small volume elements (blocks), here in a 2D representation; (b) the blocks containing an atom are marked as filled (grey); (c) the blocks falling inside the van der Waals sphere of atoms are filled also; (d) the small void elements remaining "trapped" inside a filled region are marked as filled too (yellow). All the remaining blocks (white) are empty and they constitute the porous volume to be analysed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

until another wall is touched, and the distance is stored: when all the directions have been scanned, the minimum distance between opposite walls (*MinD*) is found and assigned to this block.

Eventually, every void block has a *MinD* associated, which signals how a molecule residing in that point would interact with the material walls. The V(*MinD*) function is obtained by summing the volume of

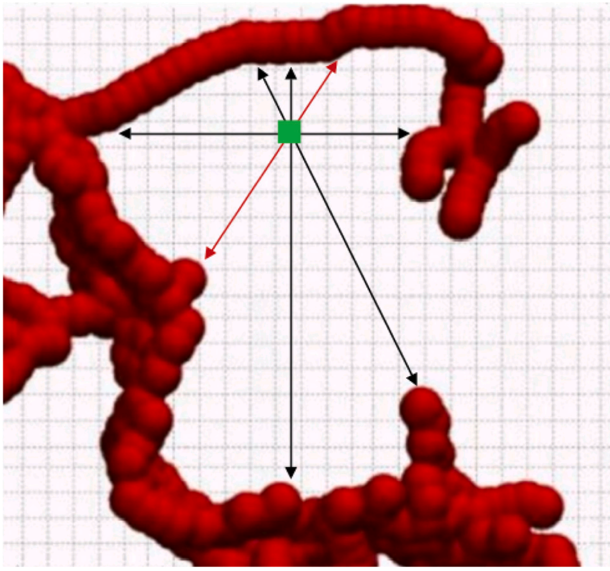


Fig. 3. Starting from a void block (green) many directions are followed until opposite walls are touched. The *MinD* for this block is measured along the red line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

all the blocks with *MinD*: in the practice this function is discretized (as in the histograms shown below) by merging all the distance values in a finite number of intervals. The finest detail is obtained when all the intervals are wide l (i.e., the block size) and this is the V(*MinD*) provided by default by PoLA; other, more concise descriptions of the porous volume can be defined by grouping the block volumes in a small number of broad intervals. A useful classification is obtained, for instance, with four intervals: $MinD \leq 7$ Å, $7 < MinD \leq 20$ Å, $20 < MinD \leq 50$ Å, $MinD \geq 50$ Å, which mimic the IUPAC definition of ultramicro, micro, meso and macropores. We underline, however, that V(*MinD*) does not indicate the amount of micropores, mesopores and so on, as discussed below and detailed in the Supporting Information (SI).

At difference with the PoLA version described in Ref. [54], the present code provides a refined definition of the porous volume too, excluding the fraction of void blocks which could not host a spherical probe (whose radius can be set in the input, with default value of 2 Å). Such blocks are typically placed in narrow inlets or bottlenecks, as illustrated in Figure S1 in the Supporting Information (SI): excluding them produces the so-called accessible porous volume, whose distribution is better related to the physisorption of gas molecules. Though PoLA output provides both total and accessible porous volume functions, the latter only will be used in the following discussions.

3. Models and methods

3.1. Carbon model data set

A large set of 109 carbon models, taken from Refs. [52] and [54], were used to correlate PoLA textural data with N_2 and H_2 adsorption isotherms (obtained for all the models with GCMC as detailed below). All the models are pure hydrocarbons: some of them (models 1 to 33) are made of corannulene molecules randomly assembled [54], others (38 to 109) are graphenic networks with different degrees of structural defects [52], and some (34 to 37) have been prepared with a reactive force field following the DynReaxMas procedure [61,62].

The carbon models used in this work cover a wide range of cell dimensions (cubic edges from 50 to 85.5 Å), densities (0.24 to 1.32 g/cm^3) and porosities: the textural properties of all the models have been analysed with PoLA, with the results thoroughly discussed in Section 4.1.

The Cartesian coordinates of all the carbon models, along with their periodic lengths, are provided in the SI.

3.2. Grand Canonical Monte Carlo simulations

The adsorption isotherms of N_2 and H_2, both at 77 K, were simulated by the GCMC method, in order to set the correlation with the distribution of the porous volume defined by PoLA.

The adsorbent geometries consisted of carbon and hydrogen atoms in some models, carbon only in others with graphitic structure, with fixed positions throughout the simulations; both the adsorbates were described with a single sphere model. Intermolecular interactions were modelled using the Lennard-Jones (LJ) 12-6 potential with parameters: $\epsilon = 28$ K and $\sigma = 0.34$ nm for carbon atoms [63], $\epsilon = 101.5$ K and $\sigma = 0.36154$ nm for N_2 [63] and $\epsilon = 36.7$ K and $\sigma = 0.2958$ nm for H_2 [64]. Solid-fluid interactions were determined using Lorentz–Berthelot combining rules, and all LJ potentials were truncated at 1.4 nm.

Each pressure point involved 3×10^7 Monte Carlo moves with equal probabilities assigned to insertion, deletion, and exchange moves, where the initial 2×10^7 moves were discarded to ensure equilibration. For H_2 adsorption, quantum effects were incorporated through the Feynman–Hibbs correction to the LJ potential, and the excess adsorption was calculated using a quantum-corrected equation of state for H_2 to account for non-ideal gas behaviour [65].

3.3. Neural network regression

A central element of the present work is the machine learning process used to correlate PoLA V(*MinD*) and gas adsorption in the porous materials. In Ref. [54] we have described a Random Forest algorithm able to predict the nitrogen uptake in a number of porous carbon models starting from the V(*MinD*): here, we aim at the reverse process, deducing a plausible V(*MinD*) from the N_2 adsorption isotherm through a neural network (NN) regression procedure. Moreover, to demonstrate the reliability of the deduced V(*MinD*), the same NN will correlate N_2 and H_2 uptakes, both in atomistic models and in some experimental samples, as described below. The NN code is described in detail below, and can be downloaded from the repository cited in [66].

3.3.1. Neural network architecture

Let (X, Y, Z) denote the training data set, where $X = \{\mathbf{x}_i \in \mathcal{X}\}_{i=1}^N$ represents the set of N_2 adsorption isotherms in the N carbon models described above, $Y = \{\mathbf{y}_i \in \mathcal{Y}\}_{i=1}^N$ represents the set of the cumulative porous volumes computed by PoLA in the same carbon models, and $Z = \{\mathbf{z}_i \in \mathcal{Z}\}_{i=1}^N$ represents the set of H_2 adsorption isotherms in the carbon models. The spaces $\mathcal{X} \subset \mathbb{R}^{d_X}$, $\mathcal{Y} \subset \mathbb{R}^{d_Y}$, and $\mathcal{Z} \subset \mathbb{R}^{d_Z}$ denote the corresponding vector spaces, with dimensionalities $d_X, d_Y, d_Z \in \mathbb{N}$.

The learning task consists of regressing Y and Z given X. To this end, we consider two neural regression architectures that encode different dependency structures between the volume-based representation $\mathbf{y}$ and the predicted isotherm $\mathbf{z}$.

The first architecture, denoted as $\mathcal{R}_A$, maps the input isotherm $\mathbf{x}$ onto a latent space composed of the volume-based representation $\mathbf{y}$ concatenated to an additional free latent component $\boldsymbol{l}$. As a result, the regression of $\mathbf{z}$ is conditioned on auxiliary latent variables which are partially conditioned on the regression of physically motivated volume descriptors $\mathbf{y}$ (see Fig. 4, left). In symbols, given two encoding functions E_A^1, E_A^2, $\mathcal{R}_A$is defined as

$$(\hat{\mathbf{y}}, \boldsymbol{l}) = E_A^1(\mathbf{x})$$
$$\hat{\mathbf{z}} = E_A^2(\boldsymbol{l})$$

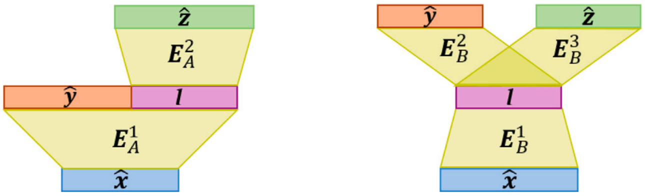


Fig. 4. Two regression models encoding different levels of dependency between $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$.

In contrast, the second architecture, denoted as $\mathcal{R}_B$, maps the input $\mathbf{x}$ into a latent representation that is simultaneously informative for reconstructing both $\mathbf{y}$ and $\mathbf{z}$. In this case, the dependency between volume descriptors and the target isotherm is mediated exclusively through this common latent variable, enforcing a stronger coupling between the two outputs (see Fig. 4, right). In symbols, given three encoding functions E_B^1, E_B^2, E_B^3, $\mathcal{R}_B$is defined as

$$\boldsymbol{l} = E_B^1(\mathbf{x})$$
$$\hat{\mathbf{y}} = E_B^2(\boldsymbol{l})$$
$$\hat{\mathbf{z}} = E_B^3(\boldsymbol{l})$$

The architectural differences between $\mathcal{R}_A$ and $\mathcal{R}_B$ reflect alternative modelling assumptions about how information is shared between representations, rather than claims about the underlying physical or causal mechanisms governing gas adsorption. As such, the models are best interpreted as different modelling choices made for predictive purposes.

Due to the limited amount of labelled data that our data set consists of (and that is likely to be available in practical applications of this method), model capacity plays a critical role in controlling overfitting. In preliminary experiments, we have found that constraining the expressive power of the networks is more effective than applying standard regularization techniques such as dropout or validation-based early stopping. Accordingly, we train all models using the full training data set and limit their capacity by adopting relatively shallow networks for the encoders $E_A^{(\cdot)}$ and $E_B^{(\cdot)}$. In particular, we adopt single-hidden-layer perceptrons with 1024 hidden units for $E_A^1, E_A^2, E_B^1, E_B^3$, while we implement E_B^2 as a simple linear projection; we do not use dropout in any case. The final set of equations that describe our regression models is

$\mathcal{R}_A$:

$$\begin{aligned} \mathbf{h} &= \mathrm{ReLU}(\mathbf{x} \cdot \mathbf{W}^{d_X \times h_1}) \\ (\hat{\mathbf{y}}, \boldsymbol{l}) &= \mathbf{h} \cdot \mathbf{W}^{h_1 \times (d_Y + d_L)} \\ \mathbf{h}' &= \mathrm{ReLU}(\boldsymbol{l} \cdot \mathbf{W}^{d_L \times h_2}) \\ \hat{\mathbf{z}} &= \mathbf{h}' \cdot \mathbf{W}^{h_2 \times d_Z} \end{aligned}$$

$\mathcal{R}_B$:

$$\begin{aligned} \mathbf{h} &= \mathrm{ReLU}(\mathbf{x} \cdot \mathbf{W}^{d_X \times h_1}) \\ \boldsymbol{l} &= \mathrm{ReLU}(\mathbf{h} \cdot \mathbf{W}^{h_1 \times d_L}) \\ \hat{\mathbf{y}} &= \boldsymbol{l} \cdot \mathbf{W}^{d_L \times d_Y} \\ \mathbf{h}' &= \mathrm{ReLU}(\boldsymbol{l} \cdot \mathbf{W}^{d_L \times h_2}) \\ \hat{\mathbf{z}} &= \mathbf{h}' \cdot \mathbf{W}^{h_2 \times d_Z} \end{aligned}$$

where $h_1 = h_2 = d_L = 1024$, and where $\mathrm{ReLU}(x) = \max\{0, x\}$ denotes the rectified linear unit used as the activation function.

3.3.2. *Dimensionality reduction*

While the original dimensionalities d_X, d_Y, d_Z are relatively large, the data exhibit a strong underlying structure. For example, the adsorption isotherms are smooth functions of the indexed variable, i.e. the pressure (with neighbouring values showing limited variation), and the cumulative porous volumes are monotonic by construction. These properties suggest that the effective dimensionality of the data is substantially lower than the original one.

Motivated by this observation, we explore a representation of the signals $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ as linear combinations of a reduced set of basis functions obtained via Principal Component Analysis (PCA). In this context, the principal components can be interpreted as “eigencurves” capturing the dominant modes of variation across the dataset. In all experiments, we retain the first 10 principal components. Figure S2 in the SI illustrates the resulting PCA decompositions for the different signals, together with the cumulative fraction of explained variance associated with each component.

3.3.3. *Training*

We train all neural regression models using stochastic gradient descent with the Adam optimizer [67], as implemented in `PyTorch` [68]. We use a learning rate of 10^{-3} and a weight decay of 10^{-8}. In the training phase we minimize a weighted mean-squared-error (MSE) loss of the form

$$\ell(\mathbf{y}, \mathbf{z}, \hat{\mathbf{y}}, \hat{\mathbf{z}}; \theta) = \lambda_y \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \lambda_z \|\mathbf{z} - \hat{\mathbf{z}}\|^2$$

where θ denotes the set of learnable model parameters $\mathbf{W}^{(\cdot)\times(\cdot)}$ for $\mathcal{R}_A$ or $\mathcal{R}_B$.

We choose the loss weights λ_y and λ_z so as to prioritize accurate prediction of the target gas isotherm, with $\lambda_z = 1$ in all experiments, and $\lambda_y = 10^{-4}$ for $\mathcal{R}_A$ and $\lambda_y = 10^{-3}$ for $\mathcal{R}_B$. This choice is justified by the results shown in the following sections, where the relative errors in the prediction of V(*MinD*) and H_2 isotherms in our carbon data set are discussed.

The training was performed for a maximum of 25 K epochs when using PCA-reduced representations, and 50 K epochs when operating on the raw signals. In both cases, we apply early stopping based on the training loss, terminating optimization if we observe no improvement for 5 K consecutive epochs.

3.3.4. *Ensembling*

In practice, we observe that individual regression models may occasionally exhibit unstable behaviour, with localized noise affecting specific regions of the predicted isotherms. Importantly, such instabilities are not consistent across architectures or input representations, and different model configurations tend to perform well on complementary subsets of the data.

To improve the robustness of the predictions, we thus adopt an ensemble strategy that combines four regression networks obtained from all combinations of our two neural architectures ($\mathcal{R}_A$ and $\mathcal{R}_B$) and two input representations (raw signals and PCA-reduced representations). Given a test input, each model in the ensemble produces an independent prediction of the target V(*MinD*) and isotherm.

From the four resulting predictions, we choose the one which is closest, in terms of mean squared error, to their average. In other words, if

$$\hat{\mathbf{y}}_1 = \mathcal{R}_A^{\mathrm{pca}}(\mathbf{x}), \quad \hat{\mathbf{y}}_2 = \mathcal{R}_B^{\mathrm{pca}}(\mathbf{x}), \quad \hat{\mathbf{y}}_3 = \mathcal{R}_A^{\mathrm{raw}}(\mathbf{x}), \quad \hat{\mathbf{y}}_4 = \mathcal{R}_B^{\mathrm{raw}}(\mathbf{x})$$

are the predictions returned by our four regression networks, we take our final prediction to be

$$\hat{\mathbf{y}}^{\mathrm{ens}} = \underset{\hat{\mathbf{y}} \in \{\hat{\mathbf{y}}_1, \ldots, \hat{\mathbf{y}}_4\}}{\arg\min} \ \mathrm{MSE}\left(\hat{\mathbf{y}}, \ \frac{1}{4}\sum_{i=1}^{4} \hat{\mathbf{y}}_i\right)$$

where MSE stands for mean square error.

This consensus-based selection strategy favours predictions that are most representative of the ensemble as a whole, effectively mitigating model-specific instabilities. Despite its simplicity, in our experiments this ensembling approach leads to a substantial improvement in prediction stability and overall predictive performance.

3.3.5. *Out-of-distribution detection*

Additionally, we equip the model with an out-of-distribution detection mechanism whose goal is to identify atypical input curves and flag them as such. To this end, we model the distribution of the training data in a reduced latent space and assess the plausibility of new samples (i.e., their being in-distribution) via their log-likelihood.

Specifically, we first project the input curves onto a low-dimensional space using Principal Component Analysis (PCA) and retain the first four principal components (eigencurves), which capture the dominant modes of variation while mitigating the curse of dimensionality. We then fit a Kernel Density Estimation (KDE) model with a Gaussian kernel to the resulting representations. We select the kernel bandwidth via grid search over 100 equispaced candidates in the interval $[10^{-4}, 1]$, using 10-fold cross-validation. To calibrate the detection threshold, we estimate the distribution of log-likelihoods on the training set using a leave-one-out procedure, and compute their empirical mean and standard deviation. At inference time, we standardize the log-likelihood of a new sample into a z-score and flag the sample as out-of-distribution if its z-score exceeds the $z_{0.95}$ confidence threshold.

3.4. *Experimental* N_2 *and* H_2 *adsorption isotherms*

To assess the reliability of the machine learning procedure for real materials, N_2 and H_2 adsorption was measured for the following pelletized activated carbons: Maxsorb MSC, Norit Row 0.8 Supra (hereafter Norit Row), BAX1700 and CGF4, purchased from Kansai Coke and Chemical Co., Honeywell Fluka, Cabot Co. and Carbotech, respectively.

Nitrogen adsorption measurements were conducted at 77 K in the relative pressure range of $10^{-7}-1$ P/P_0 using a Quantachrome Autosorb iQ2. Prior to adsorption, the samples were degassed and thermally treated at 423 K for 6 h under vacuum. The pore size distributions were calculated by applying the QSDFT model for N_2 at 77 K on carbon (slit/cylinder/sphere pores, QSDFT adsorption branch).

High pressure H_2 equilibrium isotherms were measured using a recently developed Adsorption Differential Volumetric Apparatus (ADVA-60). The system is a symmetric double branch volumetric apparatus, where the two symmetrical branches (namely, sample and reference) are connected via a differential pressure transducer. Sample thermal activation is carried out in-situ under vacuum (down to 10^{-8} kPa) using an external furnace. Prior to each experiment samples were regenerated at 150 °C overnight using a slow temperature ramp of 1 °C/min to 150 °C with an intermediate 1-hour hold at 110 °C. H_2 (99.9995%) and He (99.999%) were used for the equilibrium isotherm and volume measurements, respectively. Experiments were carried out using about 200 mg for each adsorbent. More details on the use of ADVA and its validation for high pressure equilibrium measurements are reported in previous publications [69,70] and in the SI

4. Results and discussion

4.1. *V(MinD) and simulated adsorption isotherms for the model data set*

The PoLA procedure was applied to all the carbon models, obtaining their (accessible) porous volume distributions: as an example, in Figs. 5 and 6 we show the analysis of the porous volume of two models; the results for the whole set are available as Supporting Information.

PoLA describes the porous volume at fine resolution: Fig. 5 clearly shows that PoLA is not representing the volume with simple geometrical shapes, as the blocks follow the actual distribution of the space, discriminating between points with different *MinD* (and with different host/guest interaction potentials, as discussed above).

PoLA analysis is quantitatively accurate: as shown in Fig. 6, the method computes how much of the volume resides in blocks with any given distance between opposite walls. So, for instance, we know that in model #14 the total volume of blocks with $MinD = 5 \pm 0.5$ Å is 0.062 cm^3/g, and that of blocks with $MinD = 30 \pm 0.5$ Å is 0.038 cm^3/g. In the former case, the very low value of *MinD* informs us that any adsorbate residing in that volume elements would feel a very attractive interaction potential, shaped as a narrow single well, while in the latter we expect that adsorbates would feel a looser double well potential. Clearly, knowing precisely how much of the porous volume is associated to each interaction potential gives important information about the adsorption performance of the material.

As mentioned above, PoLA can represent V(*MinD*) with different functions, reported in Fig. 7 for three other members of our model data set: detailed histograms as in Fig. 6, cumulative volume as a function of *MinD*, simplified histogram with broad intervals (which allow to evaluate at a glance the porosity of the material). So, we see that in model #30 almost all the porous volume is formed by blocks with small *MinD*, lower or equal to 12 Å, while in model #95 the largest contribution comes from volume elements with *MinD* between 7 and 20 Å, with another family around 28–30 Å, and model #6 presents a bimodal distribution of the porous volume.

It is worth noting that the volume distribution illustrated in Fig. 7 is quite different from that provided for instance by PoreBlazer, as examined in detail in Ref. [54]; analogously, the V(*MinD*) graphs obtained for real samples in Section 4.4 will be dissimilar to the DFT-based PSD. Indeed, these differences are expected, since PoLA is based on a distinct characterization of the porous volume inside the materials: this important point is further illustrated in Fig. 8, which depicts a cross-section of model #6.

Fig. 8 depicts a cross-section of model #6: this bi-dimensional representation contains some filled blocks, corresponding to points inside the van der Waals spheres of the atoms, inaccessible to adsorbates; the remaining empty blocks are characterized by a spherical pore model (first row) or by PoLA (second row). In the former model, the empty space is filled by spheres (circles, in this example) touching the material walls and the points in the void are assigned to the larger sphere that contains them; on the contrary, in PoLA every point is characterized by its own *MinD*. One can see that PoLA finds more volume elements with a small or very small *MinD* than those expected in the pore model. Indeed, some points inside large spherical pores have anyway a short distance from opposite walls, receiving a proper classification in the local analysis performed by PoLA. More discussions on the difference between pore models and PoLA can be found in the SI.

As already noted [54], using volume block edge $l = 1$ Å, the procedure is very fast: every model in our data set was analysed by PoLA with a one-processor laptop with CPU times lower than 1 min. Reducing the size of the block edge leads to a greater accuracy, at the cost of longer cpu times: however, we have shown that the V(*MinD*) does not change significantly for $l < 1$ Å, so this value is a very good compromise between speed and accuracy and it has been used for all the calculations in this paper.

Gas adsorption isotherms at 77 K were simulated by the GCMC procedure described above in all the models, for pressures ranging from 10^{-6} to 1 bar for N_2, the latter being above the condensation pressure in all the systems, and from 10^{-2} to 120 bar for H_2. Simulations provided the absolute adsorptions (in terms of number of gas molecules included in the empty space inside the material at the given chemical potential) and they were transformed to excess adsorptions by subtracting the quantity $\rho_{gas} \times V_{por}$, where ρ_{gas} is the density of the free gas (either N_2 or H_2) at 77 K and V_{por} the accessible porous volume provided by PoLA. All the excess adsorption isotherms are available in the SI.

4.2. *V(MinD) from nitrogen adsorption isotherms in models*

As stated in the Introduction, one of the main objectives of this work is a procedure to obtain reliable V(*MinD*)'s for experimental samples, comparable to the PoLA definition for atomistic models. To do this, we exploit the relationship between PoLA volume distribution and gas adsorption, motivated in the previous sections: this relationship has already been used to predict N_2 adsorption isotherms starting from V(*MinD*) in a smaller carbon model data set [54] (note that in that work the volume function was called PVD, porous volume distribution). Here we use the NN machine learning algorithm described in Section 3.3 to perform the reverse job and get the V(*MinD*) from N_2 isotherms.

To demonstrate the applicability of this approach, it was first applied to the model data set in a "leave-one-out" cross-validation scheme: for each one of the models, the NN algorithm was trained with

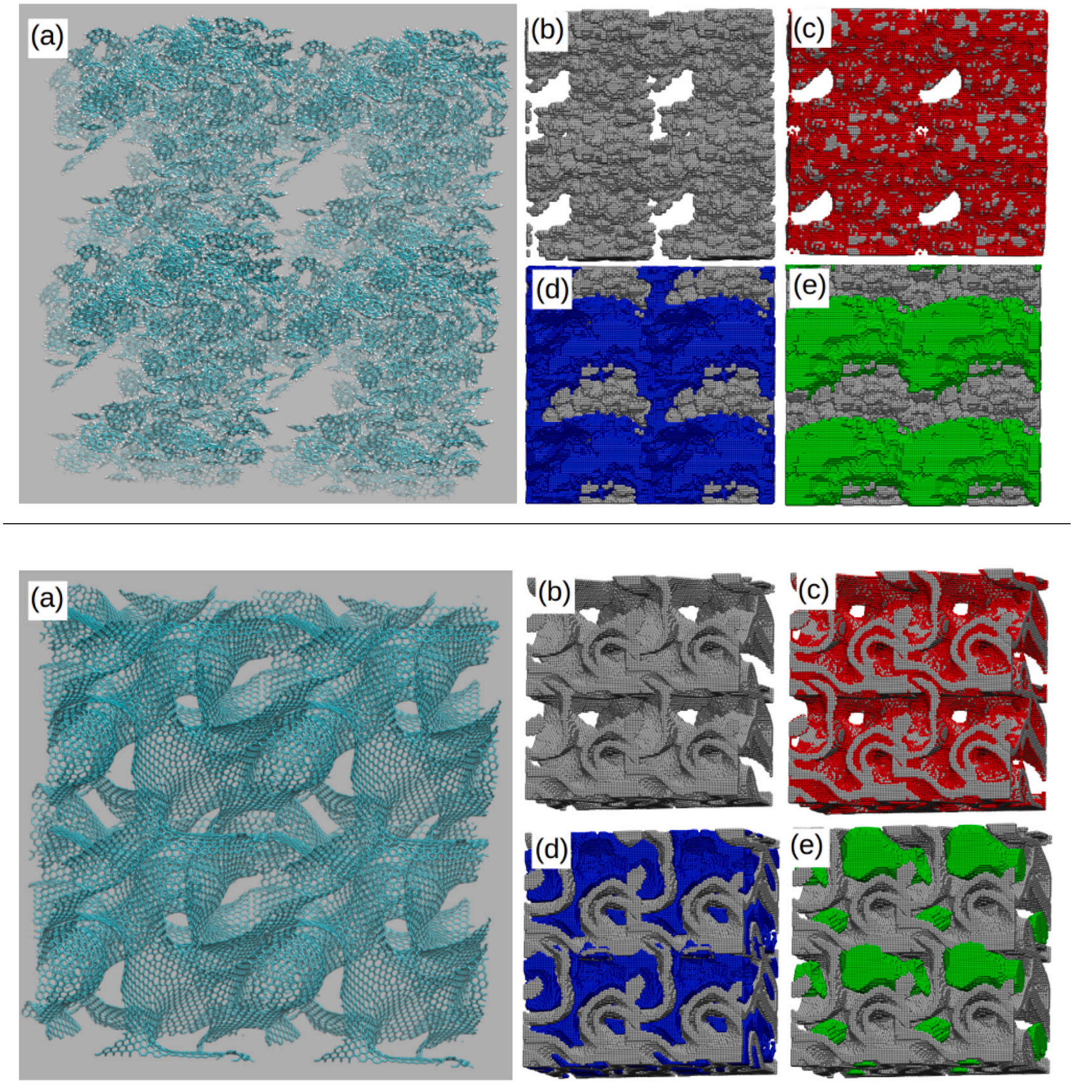


Fig. 5. PoLA analysis of models #14 (top) and #91 (bottom): (a) atomistic representation of models (to improve visibility, a 2 × 2 × 1 supercell is depicted); (b)–(e) filled blocks (grey), representing the volume actually occupied by the material atoms and inaccessible to adsorbates, and blocks with minimum distance from opposite walls: (c) $MinD \leq 7$ Å (red); (d) $7 < MinD \leq 20$ Å (blue); (e) $MinD > 20$ Å (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

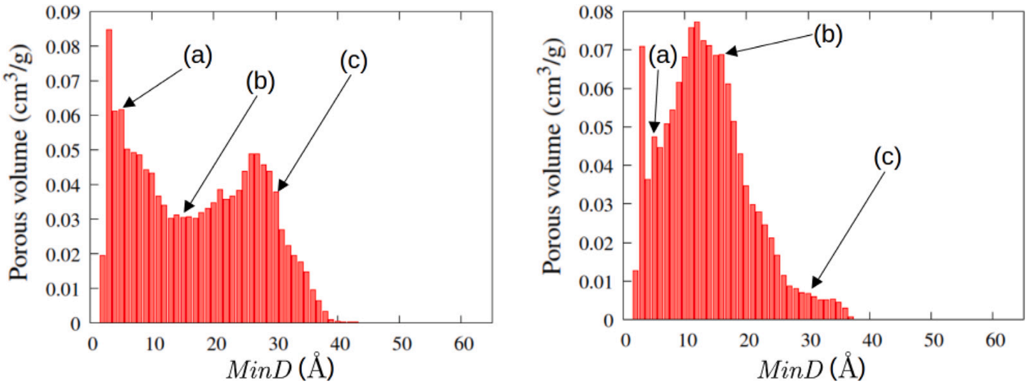
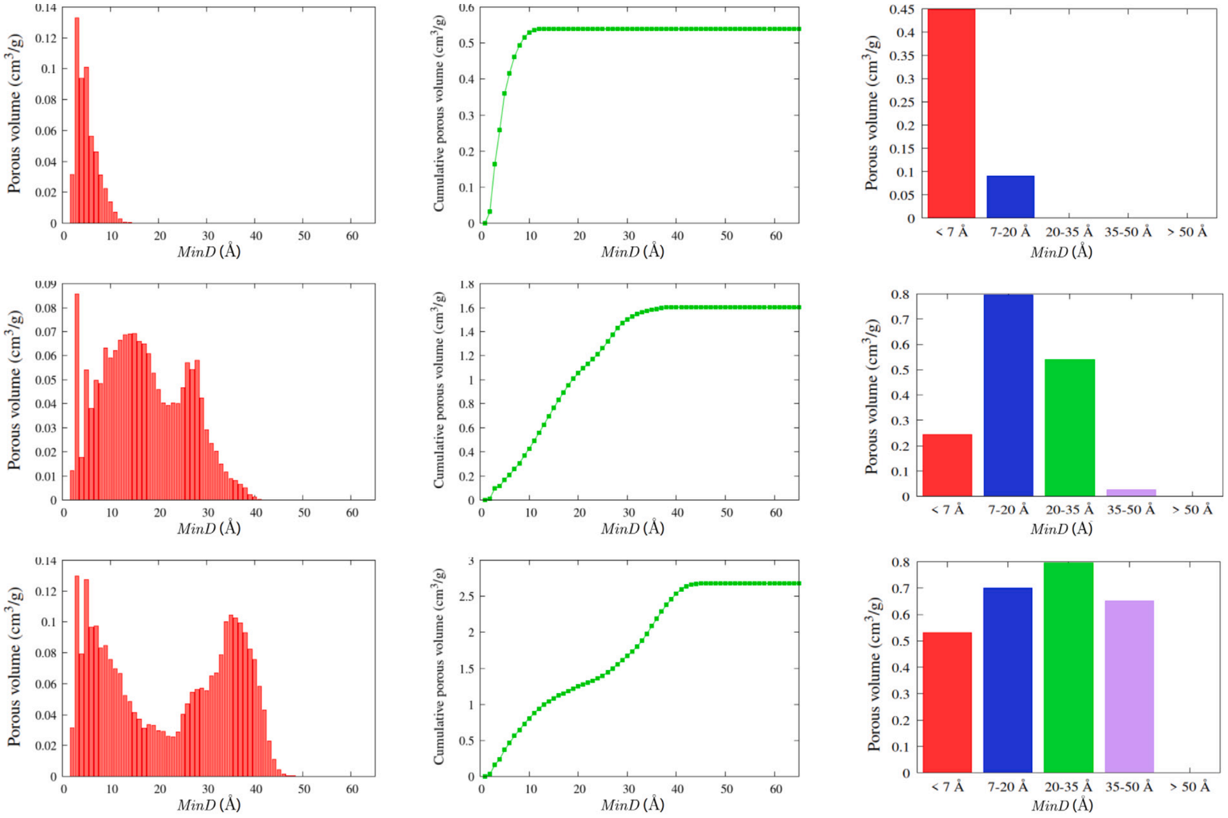


Fig. 6. Numerical analysis of the porous volume functions graphically depicted in Fig. 5 for models #14 (left) and #91 (right): the histograms report the total volume attributed to blocks with different *MinD*. For instance: (a) in models #14 and #91, 0.062 and 0.047 cm^3/g of the porous volume, respectively, belong to blocks with $MinD = 5 \pm 0.5$ Å; (b) 0.031 and 0.069 cm^3/g, respectively, belong to blocks with $MinD = 15 \pm 0.5$ Å; (c) 0.038 and 0.007 cm^3/g, respectively, belong to blocks with $MinD = 30 \pm 0.5$ Å.



V($MinD$) Cumulative V($MinD$) Simplified distribution

Fig. 7. Porous volume as a function of *MinD* for three carbon models. First to third row: model #30, #95, #6, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

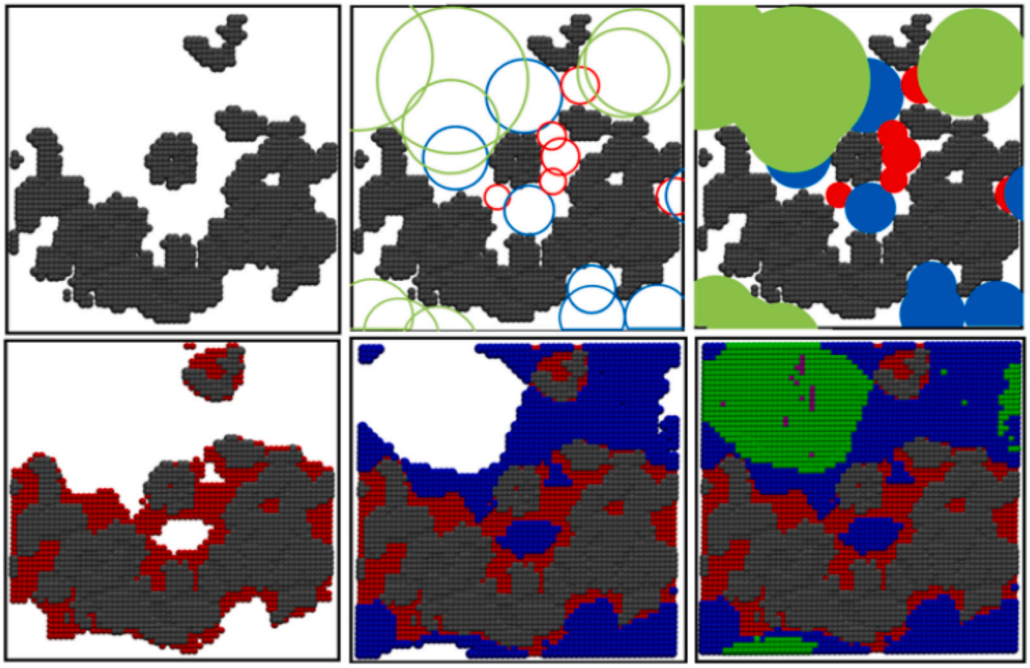


Fig. 8. A cross-section of model #6. Filled blocks (inaccessible to adsorbates) in grey. First row, the empty region (white) is comprised in spherical pores (circles, in this bi-dimensional representation; not all the circles are shown, for clarity): ultramicropores in red, micropores in blue, mesopores in green. Second row, PoLA representation: blocks with $MinD \leq 7$ Å in red, $7 < MinD \leq 20$ Å in blue, $20 < MinD \leq 35$ Å in green, and $MinD > 35$ Å in purple. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

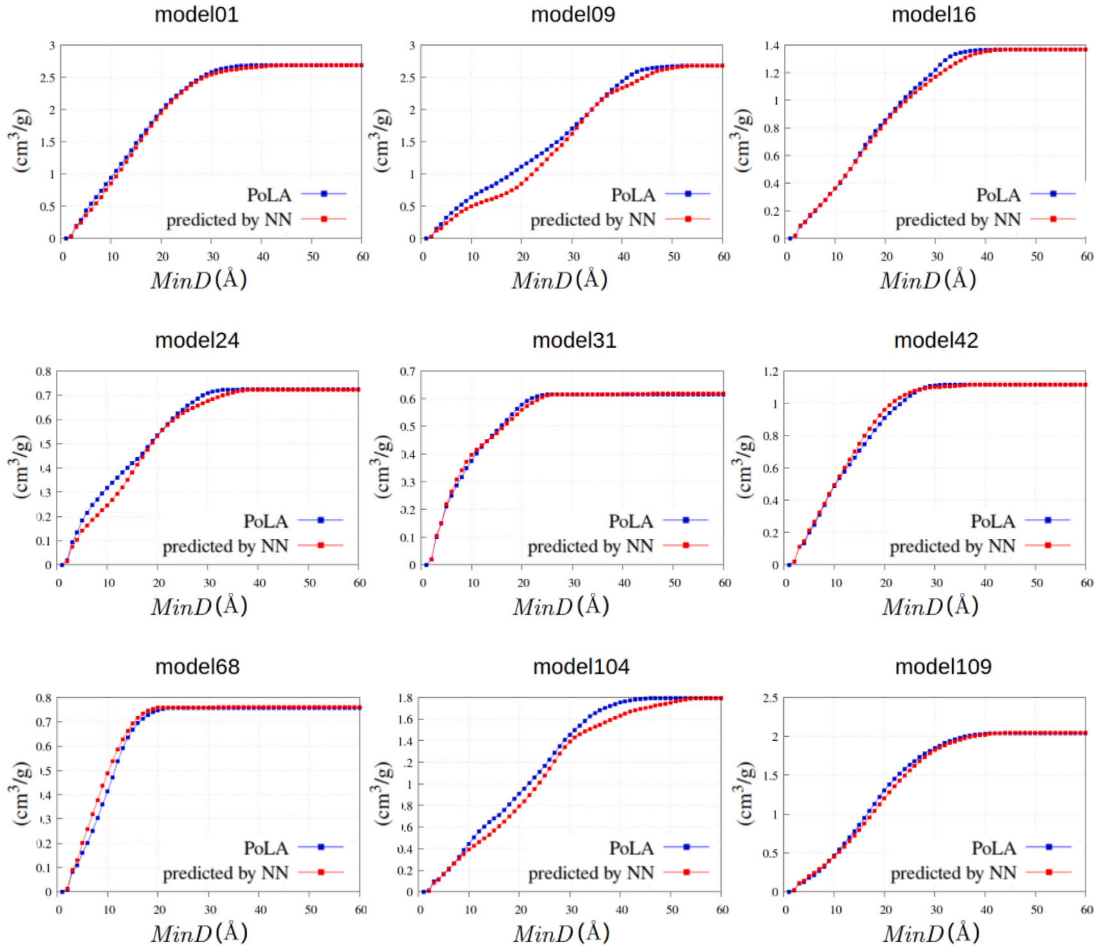


Fig. 9. Cumulative porous volume predicted by the NN machine learning algorithm and computed by PoLA for a selection of carbon models.

the V(*MinD*) and N_2 isotherms of the other 108 models and then used to predict the V(*MinD*) of the left-out from its adsorption isotherm.

To illustrate the results, the predicted cumulative volume for nine models are presented along with the PoLA actual volumes in Fig. 9: the results for the whole data set are provided in the SI. A more informative review of the NN performance is provided by the mean absolute errors (MAE): Fig. 10 shows the MAE divided by the total porous volume computed by PoLA for all the models (the numerical values are reported in Table S1, in the SI). One can see that the agreement is in general very satisfactory: out of 109 models, 46 have a relative MAE lower than 1%, and 40 between 1% and 2%; only two models present a relative MAE larger than 5%.

These results confirm that the machine learning algorithm based on PoLA volume distributions and simulated N_2 isotherms provides a fair description of the porous structure of “unknown” carbons starting from their nitrogen uptake.

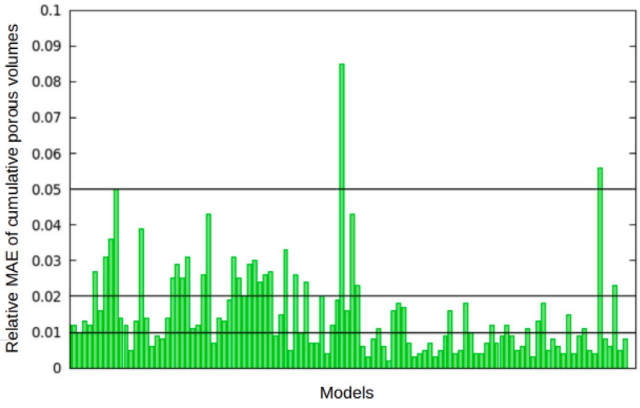


Fig. 10. Mean absolute errors (MAE) of the cumulative porous volumes predicted by NN algorithm, divided by the total PoLA volume, for the model data set. The black lines show the limit of 1%, 2% and 5% relative MAE.

4.3. Prediction of hydrogen adsorption isotherms in models

When the procedure described in the previous Section is applied to real samples, deriving the V(*MinD*) from the measured N_2 isotherm, we cannot compare the result to an independent description of the porous volume. Then, as anticipated above, we use the same NN algorithm to predict also the H_2 isotherms at 77 K, which can be compared to their experimental counterparts.

As a first step, we tested this approach on our carbon model data set. The leave-one-out cross-validation scheme was applied as above, training the NN algorithm with the simulated hydrogen adsorption isotherms, along with the nitrogen uptakes and the PoLA V(*MinD*), for 108 carbon models, and then predicting the H_2 isotherm for the left-out from its N_2 isotherm.

A partial illustration of the results is reported in Fig. 11 for the same models shown in Fig. 9: as above, the full set of results is provided in the SI. Also in this case the NN prediction is satisfactory: for most models, the curves simulated by GCMC and predicted by NN are in good agreement, though some discrepancies remain for some of the carbons as in models #30 or #31 (shown in Fig. 11).

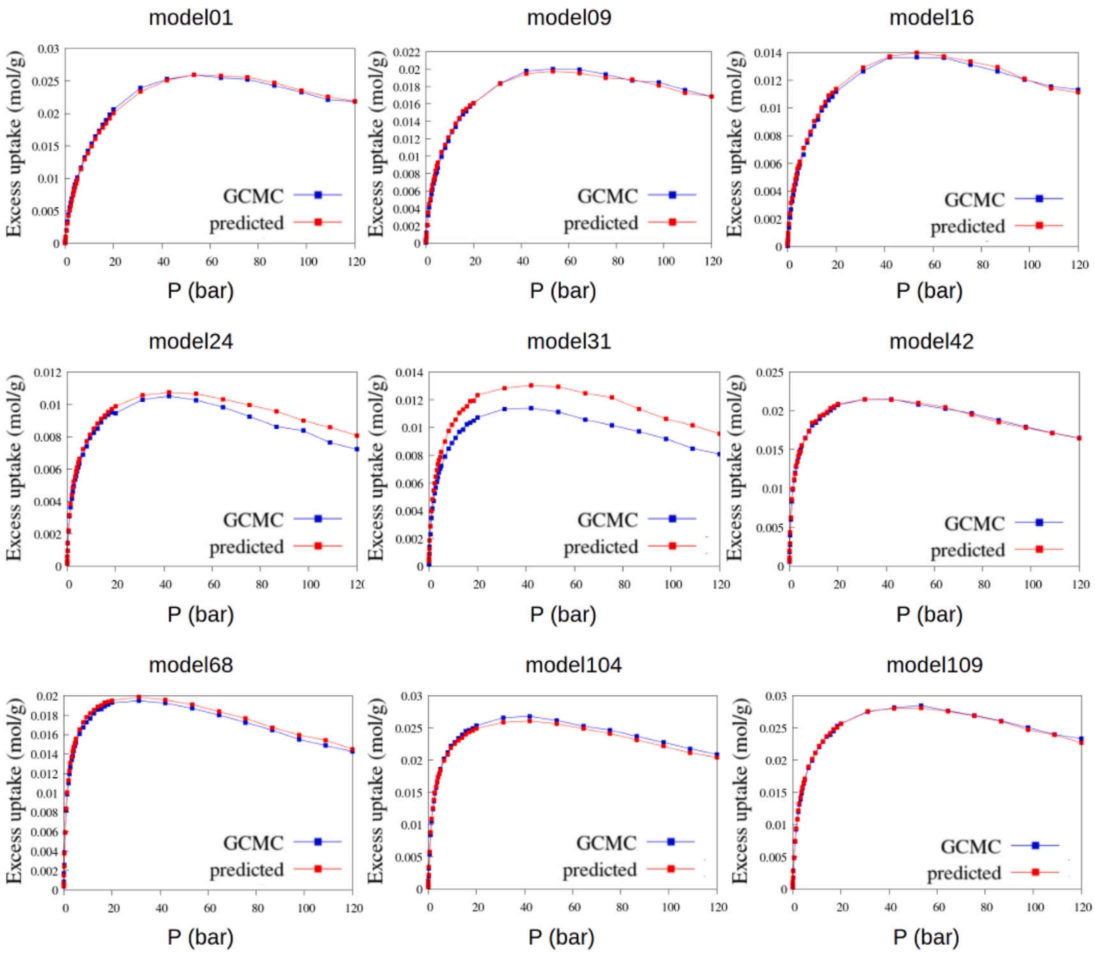


Fig. 11. Excess H_2 adsorption isotherms at 77 K predicted by the NN machine learning algorithm and simulated by GCMC for a selection of carbon models.

The reliability of the NN prediction of H_2 excess uptake can be evaluated using the mean absolute error with respect to the GCMC value. In Fig. 12 we report, as above, the MAE divided by the maximum value of the GCMC curve: the overall performance on the 109 models is again satisfactory, as the relative MAE is below 1% for 49, between 1% and 2% for 26 and between 2% and 5% for other 26 models. Only few of the predicted isotherms deviate significantly from the simulations, with relative MAE around 10% or even above 15% in one case (the above-mentioned model #30): nonetheless, we can conclude that the NN algorithm in general provides reliable predictions of isotherms as well as of porous volume distributions.

4.4. V(MinD) from N_2 *isotherms in real samples*

When the machine learning algorithm is applied to a real porous carbon, starting from the measured N_2 excess isotherm, the output is a detailed description of the PoLA-like porous volume that can be presented in the various forms discussed above.

We have tested this feature on four commercially available porous carbons, namely Norit Row, Maxsorb, BAX1700 and CGF4. The N_2 adsorption isotherms were recorded at 77 K for all the samples, and used as input to the NN which provided the PoLA textural analysis reported in Fig. 13; for comparison, the usual DFT-based analysis in terms of geometrical pores is reported for the same systems in Fig. 14.

PoLA and DFT analyses differ, not surprisingly, as the methods are based on very different models, as diffusely discussed in the previous Sections. As already seen for atomistic models, e.g. in Fig. 8, also for these samples PoLA finds a fraction of the porous volume with small *MinD* (below 7 Å or between 7 and 20 Å) quite larger than the ultramicro and microporous volumes found in the DFT model. Conversely, DFT sees large meso and macroporous volumes much bigger than the PoLA volumes with *MinD* larger that 35 Å: this difference is particularly evident for BAX1700, but clearly visible also for the other samples.

This concept is further illustrated in the SI, where really spherical cavities are considered, showing how PoLA observes some volume elements with very small *MinD* even in large spheres (which are entirely micro- or mesoporous in the pore model, depending on the radius).

For real carbons as well as for models, then, PoLA provides a different characterization of the porous volume than standard pore size distributions: hopefully, this novel description brings some new and useful information, as stated in the previous Sections, related to the behaviour of the materials in physisorption applications.

Clearly, in the case of experimental samples the V(*MinD*) function derives from the ML regression and not from the direct analysis of an atomistic model, but we assume that also in this case PoLA description is related to the interaction potential felt by adsorbates in the various volume elements. Such an assumption is partly justified by the agreement between ML regression and V(*MinD*) actually computed in models, as seen in Section 4.2, but to reinforce it, in the next Section we will compare the hydrogen adsorption isotherms predicted by the same ML algorithm and experimentally measured on the same carbons.

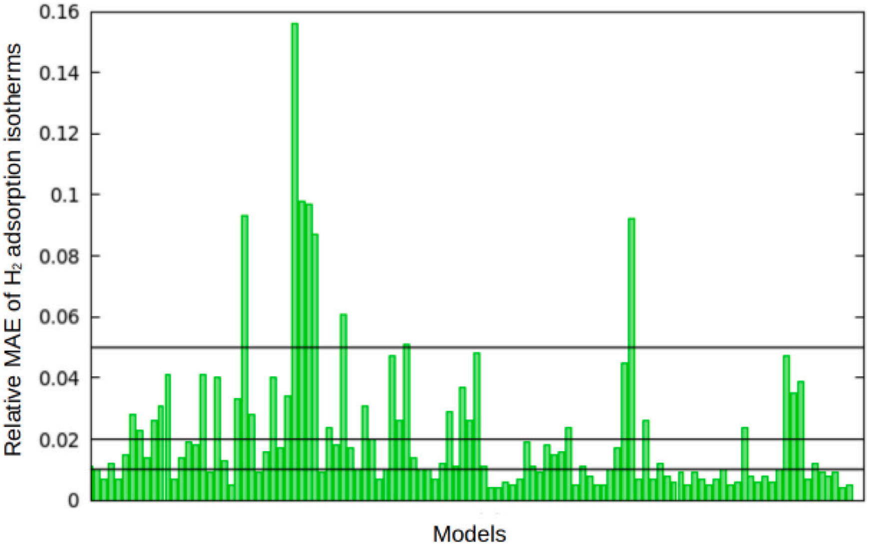


Fig. 12. Mean absolute errors (MAE) of the H_2 excess adsorption isotherms predicted by NN algorithm with respect to GCMC simulations, divided by the maximum GCMC uptake. The black lines show the limit of 1%, 2% and 5% relative MAE.

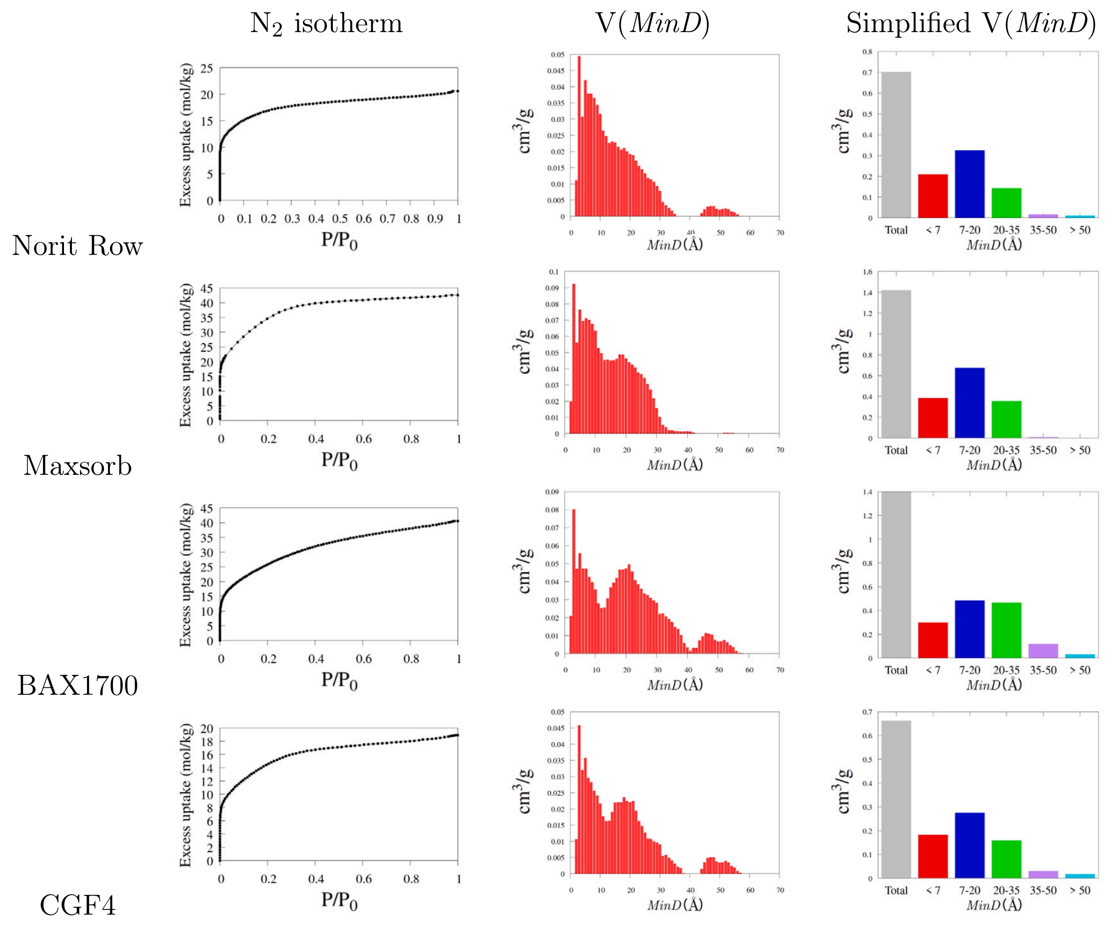


Fig. 13. N_2 adsorption isotherms for commercial porous carbons, along with PoLA textural analyses.

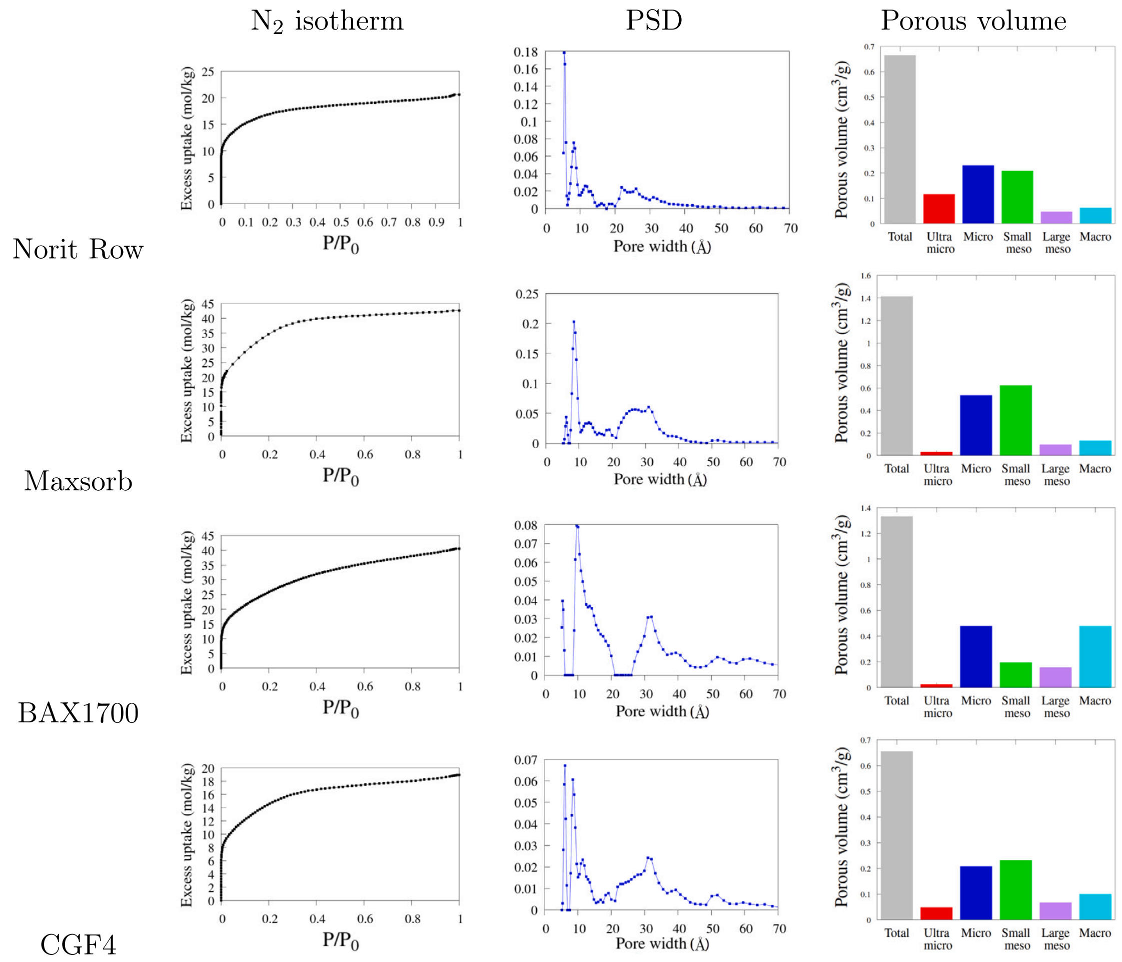


Fig. 14. N_2 adsorption isotherms for commercial porous carbons, along with QSDFT textural analyses. In the last column, the porous volume is reported in analogy with Fig. 13: ultramicro indicates the porous volume with pore size lower than 7 Å, micro between 7 and 20 Å, small and large meso between 20 and 35 Å and between 35 and 50 Å, respectively, and macro with pore size larger than 50 Å.

4.5. Prediction of H_2 *isotherms in real samples*

We have shown in the previous paragraphs that PoLA provides an original description of the porosity: the information carried by V(*MinD*) is indeed quite different from that provided, for instance, by DFT models. As pointed out above, to prove the usefulness of this new model we can show that it is actually related to the physics of adsorption: as done for the carbon models in our data set, we used the NN algorithm to predict H_2 adsorption isotherms along with V(*MinD*) for the commercially available porous carbons mentioned above.

In Fig. 15 the predicted H_2 excess isotherms are compared with the experimental curves, recorded by us on all the samples as explained in Section 3.4.

The agreement between predicted and measured isotherms is very satisfactory, over the entire range of pressures for which the experimental data were available. This result is clearly of great interest per se, since predicting reliable hydrogen uptakes up to high pressures can assist the design of new adsorbents, reducing time and costs strongly. Moreover, this agreement reinforces the plausibility of PoLA V(*MinD*), since this quantity is obtained through the same regression process, starting from our model data set: these results support the physical basis of V(*MinD*) as an effective descriptor of the porous volume.

5. Conclusions

We have presented a new procedure to describe the textural properties of porous materials: PoLA analyses the void space inside the material point by point, assigning a porous nature to each small volume element, based on the distance of the given point from opposite material walls.

We propose to adopt the function V(*MinD*) generated by PoLA as a descriptor of porosity, since it provides a common metric for crystalline and amorphous materials, can be applied to atomistic models as well as to real samples and brings information about physisorption processes.

To support these claims, a large data set of 109 carbon models, with various densities and porosities, have been characterized by PoLA, and N_2 and H_2 adsorption isotherms at 77 K were simulated by GCMC for all the models. The results have been used to train a neural network algorithm, specifically optimized to provide the V(*MinD*) and predict the hydrogen uptake from the nitrogen isotherm. This NN regressor was able to reproduce the data of carbon models, and to provide the V(*MinD*) and the hydrogen isotherms for some real porous carbon samples (for which the predicted H_2 uptakes can been compared to experimental curves).

In all the cases, the predictions of the NN regressor have proved strongly reliable, confirming the relationship between PoLA V(*MinD*) and physisorption properties. Future work will extend the procedure

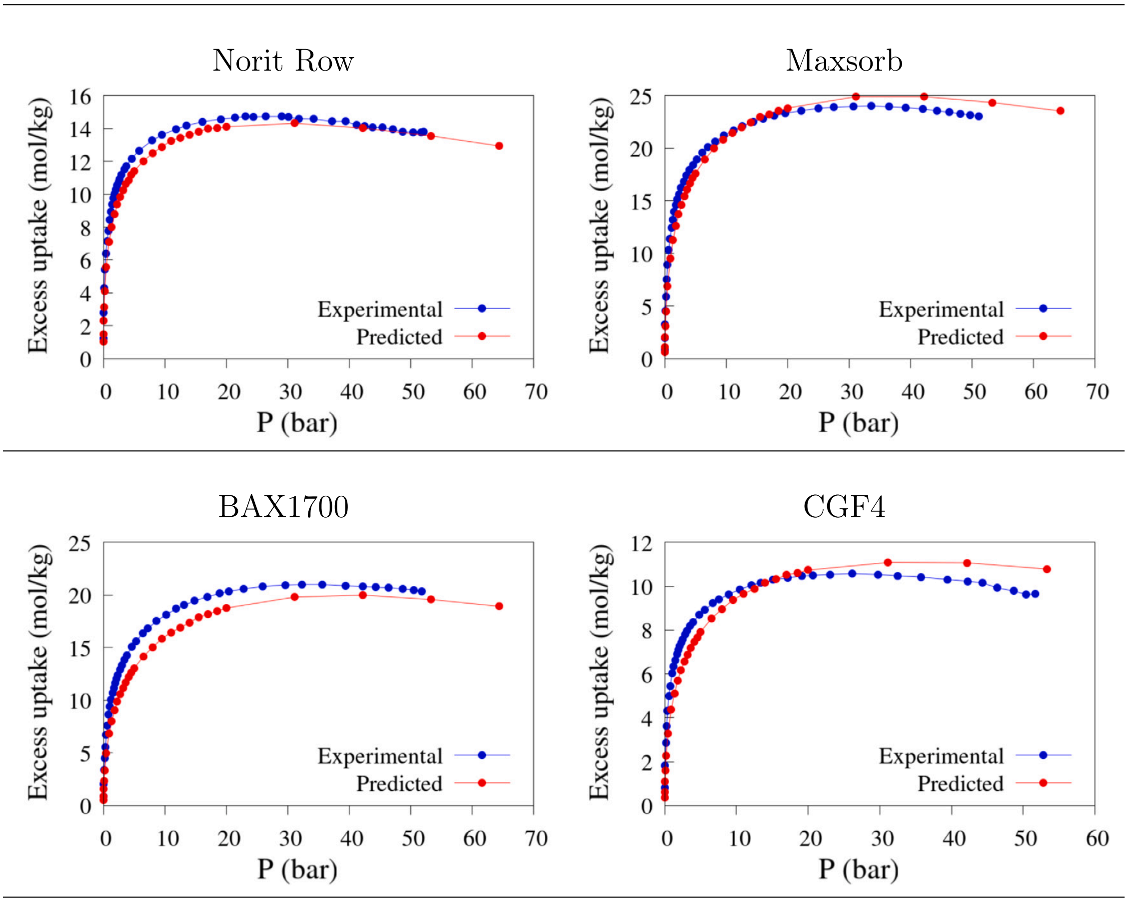


Fig. 15. Excess H_2 adsorption isotherms at 77 K, measured experimentally and predicted by NN regression from the nitrogen isotherms.

to other adsorbates (in particular, CO_2 and methane), as well as to other adsorbent models, including zeolites and silica, and derivatized carbons.

CRediT authorship contribution statement

Maurizio Cossi: Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Conceptualization. **Alberto Zoccante:** Software, Methodology, Investigation, Conceptualization. **Valeria Palumbo:** Software, Investigation. **Maddalena D'Amore:** Investigation. **Fernando Vallejos-Burgos:** Writing – original draft, Methodology, Investigation, Conceptualization. **Radovan Kukobat:** Software. **Alejandro Moreo:** Writing – original draft, Software, Methodology, Conceptualization. **Fabrizio Sebastiani:** Methodology, Conceptualization. **Stefano Brandani:** Methodology, Investigation. **Enzo Mangano:** Methodology, Investigation. **Federico Begni:** Investigation. **Giorgio Celoria:** Investigation. **Leonardo Marchese:** Methodology, Investigation, Funding acquisition, Conceptualization.

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.

Acknowledgements

Prof. Matthias Thommes (Friedrich-Alexander-Universität, Erlangen) is warmly thanked for helpful discussions. This research is part of the ECOSTORE-H2 project funded under the Italian National Recovery and Resilience Plan (NRRP), M2C2-3.5 of Ministero dell'Ambiente e della Sicurezza Energetica funded by the European Union-NextGenerationEU (project code: RSH2 A_000031, CUP: F67G22000140004). The financial support by Syensqo to Centro RiSPA at the Joint-Lab DISIT /Syensqo (Alessandria, Italy) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.carbon.2026.121713.

References

[1] K.M. Thomas, Hydrogen adsorption and storage on porous materials, Catal. Today 120 (2007) 389–398, http://dx.doi.org/10.1016/j.cattod.2006.09.015.

[2] R.E. Morris, P.S. Wheatley, Gas storage in nanoporous materials, Angew. Chem. Int. Ed. 47 (2008) 4966–4981, http://dx.doi.org/10.1002/anie.200703934.

[3] K.L. Lim, H. Kazemian, Z. Yaakob, W.R.W. Daud, Solid-state materials and methods for hydrogen storage: A critical review, Chem. Eng. Technol. 33 (2010) 213–226, http://dx.doi.org/10.1002/ceat.200900376.

[4] M.M. Deegan, M.R. Dworzak, A.J. Gosselin, K.J. Korman, E.D. Bloch, Gas storage in porous molecular materials, Chem. Eur. J. 27 (2021) 4531–4547, http://dx.doi.org/10.1002/chem.202003864.

[5] Y. Wu, B.M. Weckhuysen, Separation and purification of hydrocarbons with porous materials, Angew. Chem. Int. Ed. 60 (2021) 18930–18949, http://dx.doi.org/10.1002/anie.202104318.

[6] J. Wu, X. Zhu, F. Yang, R. Wang, T. Ge, Shaping techniques of adsorbents and their applications in gas separation: a review, J. Mater. Chem. A 10 (2022) 22853–22895, http://dx.doi.org/10.1039/d2ta04352a.

[7] M. Ding, X. Liu, P. Ma, J. Yao, Porous materials for capture and catalytic conversion of CO2 at low concentration, Coord. Chem. Rev. 465 (2022) http://dx.doi.org/10.1016/j.ccr.2022.214576.

[8] V. Sharma, A. Agrawal, O. Singh, R. Goyal, B. Sarkar, N. Gopinathan, S.P. Gumfekar, A comprehensive review on the synthesis techniques of porous materials for gas separation and catalysis, Can. J. Chem. Eng. 100 (2022) 2653–2681, http://dx.doi.org/10.1002/cjce.24507.

[9] Y. Liu, L. Chen, L. Yang, T. Lan, H. Wang, C. Hu, X. Han, Q. Liu, J. Chen, Z. Feng, X. Cui, Q. Fang, H. Wang, L. Li, Y. Li, H. Xing, S. Yang, D. Zhao, J. Li, Porous framework materials for energy & environment relevant applications: A systematic review, Green Energy Environ. 9 (2024) 217–310, http://dx.doi.org/10.1016/j.gee.2022.12.010.

[10] W. Yan, H. Huang, A. Zhang, H. Dong, W. Liao, Z. He, X. Yang, Q. He, Frontiers in applications of porous materials in CO2 gas separation membranes: Mechanisms, membrane properties, and future perspectives of porous aromatic frameworks (PAFs), J. Environ. Chem. Eng. 12 (2024) http://dx.doi.org/10.1016/j.jece.2024.113509.

[11] L.A. Mahmoud, J.L. Rowlandson, D.J. Fermin, V.P. Ting, S. Nayak, Porous carbons: a class of nanomaterials for efficient adsorption-based hydrogen storage, 2024, http://dx.doi.org/10.1039/d4lf00215f.

[12] T. Zhu, Y. Han, S. Liu, B. Yuan, Y. Liu, H. Ma, Porous materials confining single atoms for catalysis, Front. Chem. 9 (2021) http://dx.doi.org/10.3389/fchem.2021.717201.

[13] J. Dai, H. Zhang, Recent advances in catalytic confinement effect within micro/meso-porous crystalline materials, Small 17 (2021) http://dx.doi.org/10.1002/smll.202005334.

[14] W.G. Cui, T.L. Hu, Incorporation of active metal species in crystalline porous materials for highly efficient synergetic catalysis, Small 17 (2021) http://dx.doi.org/10.1002/smll.202003971.

[15] S. Chongdar, S. Bhattacharjee, P. Bhanja, A. Bhaumik, Porous organic-inorganic hybrid materials for catalysis, energy and environmental applications, Chem. Commun. 58 (2022) 3429–3460, http://dx.doi.org/10.1039/d1cc06340e.

[16] T. Wang, B. Tian, B. Han, D. Ma, M. Sun, A. Hanif, D. Xia, J. Shang, Recent advances on porous materials for synergetic adsorption and photocatalysis, Energy Environ. Mater. 5 (2022) 711–730, http://dx.doi.org/10.1002/eem2.12229.

[17] K. Xue, Y. Mo, B. Long, W. Wei, C. Shan, S. Guo, L. Niu, Single-atom catalysts supported on ordered porous materials: Synthetic strategies and applications, InfoMat 4 (2022) http://dx.doi.org/10.1002/inf2.12296.

[18] Y. Tang, G. Qi, S. Wang, X. Meng, F.S. Xiao, Recent development of bio-inspired porous materials for catalytic applications, Chem. Res. Chin. Univ. 39 (2023) 13–18, http://dx.doi.org/10.1007/s40242-022-2164-0.

[19] V.K. Gupta, T.A. Saleh, Sorption of pollutants by porous carbon, carbon nanotubes and fullerene- An overview, Environ. Sci. Pollut. Res. 20 (2013) 2828–2843, http://dx.doi.org/10.1007/s11356-013-1524-1.

[20] M.B. Mosbah, L. Mechi, R. Khiari, Y. Moussaoui, Current state of porous carbon for wastewater treatment, Processes 8 (2020) 1–24, http://dx.doi.org/10.3390/pr8121651.

[21] T. Fu, B. Zhang, X. Gao, S. Cui, C.Y. Guan, Y. Zhang, B. Zhang, Y. Peng, Recent progresses, challenges, and opportunities of carbon-based materials applied in heavy metal polluted soil remediation, Sci. Total Environ. 856 (2023) http://dx.doi.org/10.1016/j.scitotenv.2022.158810.

[22] J. Florek, M. Negoro, Y. Hu, K. Kanamori, K. Nakanishi, F. Kleitz, The role of nanoporous adsorbents in the circular economy—Closing the loop of critical materials recovery, 2025, http://dx.doi.org/10.1002/adfm.202409462.

[23] M. Lewoyehu, Comprehensive review on synthesis and application of activated carbon from agricultural residues for the remediation of venomous pollutants in wastewater, J. Anal. Appl. Pyrolysis 159 (2021) http://dx.doi.org/10.1016/j.jaap.2021.105279.

[24] X.L. Zhou, H. Zhang, L.M. Shao, F. Lü, P.J. He, Preparation and application of hierarchical porous carbon materials from waste and biomass: A review, Waste Biomass Valorization 12 (2021) 1699–1724, http://dx.doi.org/10.1007/s12649-020-01109-y.

[25] Y. Yin, Q. Liu, J. Wang, Y. Zhao, Recent insights in synthesis and energy storage applications of porous carbon derived from biomass waste: A review, Int. J. Hydrog. Energy 47 (2022) 39338–39363, http://dx.doi.org/10.1016/j.ijhydene.2022.09.121.

[26] K. Cao, S. Zhang, Y. Shi, X. Diao, R. Wei, N. Ji, Catalytic upgrading of plastic wastes into high-value carbon nanomaterials: Synthesis and applications, 2025, http://dx.doi.org/10.1021/acsnano.5c03391.

[27] M. Thommes, K. Kaneko, A.V. Neimark, J.P. Olivier, F. Rodriguez-Reinoso, J. Rouquerol, K.S. Sing, Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC technical report), Pure Appl. Chem. 87 (9–10) (2015) 1051–1069.

[28] S.V. Sklepova, N. Ivanichok, P. Kolkovskyi, V. Kotsyubynsky, V. Boychuk, B. Rachiy, A. Uhryński, M. Bembenek, L. Ropyak, Porous structure and fractal dimensions of activated carbon prepared from waste coffee grounds, Materials 16 (18) (2023) http://dx.doi.org/10.3390/ma16186127.

[29] N. Seaton, J. Walton, N. quirke, A new analysis method for the determination of the pore size distribution of porous carbons from nitrogen adsorption measurements, Carbon 27 (6) (1989) 853–861, http://dx.doi.org/10.1016/0008-6223(89)90035-3.

[30] G.M. Davies, N.A. Seaton, V.S. Vassiliadis, Calculation of pore size distributions of activated carbons from adsorption isotherms, Langmuir 15 (1999) 8235–8245, http://dx.doi.org/10.1021/la9902643.

[31] G.M. Davies, N.A. Seaton, Development and validation of pore structure models for adsorption in activated carbons, Langmuir 15 (1999) 6263–6276, http://dx.doi.org/10.1021/la990160s.

[32] K.A. Cychosz, R. Guillet-Nicolas, J. García-Martínez, M. Thommes, Recent advances in the textural characterization of hierarchically structured nanoporous materials, Chem. Soc. Rev. 46 (2017) 389–414, http://dx.doi.org/10.1039/c6cs00391e.

[33] M. Thommes, C. Schlumberger, Characterization of nanoporous materials, Annu. Rev. Chem. Biomol. Eng. 55 (2021) 27, http://dx.doi.org/10.1146/annurev-chembioeng.

[34] C. Ebner, W. Saam, D. Stroud, Densit-functional theory of simple classical fluids. I. Surfaces, Phys. Rev. A 14 (1976) 2264–2273.

[35] W.F. Saam, C. Ebner, Density-functional theory of classical systems, Phys. Rev. A 15 (1977) 2566–2568, http://dx.doi.org/10.1103/PhysRevA.15.2566.

[36] J. Wu, Density functional theory for chemical engineering: From capillarity to soft materials, AIChE J. 52 (2006) 1169–1193, http://dx.doi.org/10.1002/aic.10713.

[37] J. Wu, Z. Li, Density-functional theory for complex fluids, Annu. Rev. Phys. Chem. 58 (2007) 85–112, http://dx.doi.org/10.1146/annurev.physchem.58.032806.104650.

[38] C. Lastoskie, K.E. Gubbins, N. Quirke, Pore size distribution analysis of microporous carbons: a density functional theory approach, J. Phys. Chem. 97 (18) (1993) 4786–4796, http://dx.doi.org/10.1021/j100120a035.

[39] A.V. Neimark, The method of indeterminate Lagrange multipliers in nonlocal density functional theory, Langmuir 11 (10) (1995) 4183–4184, http://dx.doi.org/10.1021/la00010a090.

[40] P.I. Ravikovitch, A.V. Neimark, Density functional theory model of adsorption on amorphous and microporous silica materials, Langmuir 22 (2006) 11171–11179, http://dx.doi.org/10.1021/la0616146.

[41] J.F. Lutsko, Recent Developments in Classical Density Functional Theory, vol. 144, John Wiley & Sons, Inc, 2010.

[42] J. Landers, G.Y. Gor, A.V. Neimark, Density functional theory methods for characterization of porous materials, Colloids Surf. A: Physicochem. Eng. Asp. 437 (2013) 3–32, http://dx.doi.org/10.1016/j.colsurfa.2013.01.007.

[43] J. Jagiello, J.P. Olivier, A simple two-dimensional NLDFT model of gas adsorption in finite carbon pores. application to pore structure analysis, J. Phys. Chem. C 113 (2009) 19382–19385, http://dx.doi.org/10.1021/jp9082147.

[44] J. Jagiello, C. Ania, J.B. Parra, C. Cook, Dual gas analysis of microporous carbons using 2D-NLDFT heterogeneous surface model and combined adsorption data of N2 and CO2, Carbon 91 (2015) 330–337, http://dx.doi.org/10.1016/j.carbon.2015.05.004.

[45] L. Sarkisov, A. Harrison, Computational structure characterisation tools in application to ordered and disordered porous materials, Mol. Simul. 37 (15) (2011) 1248–1257.

[46] L. Sarkisov, J. Kim, Computational structure characterization tools for the era of material informatics, Chem. Eng. Sci. 121 (2015) 322–330.

[47] L. Sarkisov, R. Bueno-Perez, M. Sutharson, D. Fairen-Jimenez, Materials informatics with PoreBlazer v4.0 and the CSD MOF database, Chem. Mater. 32 (2020) 9849–9867, http://dx.doi.org/10.1021/acs.chemmater.0c03575.

[48] T.F. Willems, C.H. Rycroft, M. Kazi, J.C. Meza, M. Haranczyk, Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials, Microporous Mesoporous Mater. 149 (1) (2012) 134–141, http://dx.doi.org/10.1016/j.micromeso.2011.08.020, URL https://www.sciencedirect.com/science/article/pii/S1387181111003738.

[49] R.L. Martin, B. Smit, M. Haranczyk, Addressing challenges of identifying geometrically diverse sets of crystalline porous materials, J. Chem. Inf. Model. 52 (2) (2012) 308–318.

[50] L.D. Gelb, K.E. Gubbins, Pore size distributions in porous glasses: a computer simulation study, Langmuir 15 (2) (1999) 305–308, http://dx.doi.org/10.1021/la9808418.

[51] L.D. Gelb, K.E. Gubbins, Characterization of porous glasses: Simulation models, adsorption isotherms, and the brunauer-emmett-teller analysis method, Langmuir 14 (8) (1998) 2097–2111, http://dx.doi.org/10.1021/la9710379.

[52] F. Vallejos-Burgos, C. de Tomas, N.J. Corrente, K. Urita, S. Wang, C. Urita, I. Moriguchi, I. Suarez-Martinez, N. Marks, M.H. Krohn, R. Kukobat, A.V. Neimark, Y. Gogotsi, K. Kaneko, 3D nanostructure prediction of porous carbons via gas adsorption, Carbon 215 (2023) http://dx.doi.org/10.1016/j.carbon.2023.118431.

[53] https://3d-vis.streamlit.app.

[54] A. Zoccante, M. D'Amore, C.A. Guido, A. Fortunelli, G. Conter, L. Marchese, M. Cossi, Porosity local analysis (PoLA): a new approach to describe the porous volume distribution in amorphous carbons, ACS Omega 10 (29) (2025) 31623–31637, http://dx.doi.org/10.1021/acsomega.5c02479.

[55] D. Morgan, R. Jacobs, Opportunities and challenges for machine learning in materials science, Annu. Rev. Mater. Res. 50 (2020) http://dx.doi.org/10.1146/annurev-matsci-070218-010015.

[56] A. Rosen, S. Iyer, D. Ray, Z. Yao, A. Aspuru-Guzik, L. Gagliardi, J. Notestein, R. Snurr, Machine learning the quantum-chemical properties of metal-organic frameworks for accelerated materials discovery, Matter 4 (2021) http://dx.doi.org/10.1016/j.matt.2021.02.015.

[57] H. Tang, J. Jiang, Active learning boosted computational discovery of covalent–organic frameworks for ultrahigh CH4 storage, AIChE J. 68 (2022) http://dx.doi.org/10.1002/aic.17856.

[58] X. Wu, Y. Liu, Predicting gas adsorption without the knowledge of pore structures: a machine learning method based on classical density functional theory, J. Phys. Chem. Lett. 14 (2023) 10094–10102, http://dx.doi.org/10.1021/acs.jpclett.3c02708.

[59] S.J. Gregg, K.S.W. Sing, Adsorption, Surface Area, and Porosity, second ed., Academic Press, London ; New York, 1982, OCLC: 8908132.

[60] M. Thommes, Physical adsorption characterization of ordered and amorphous mesoporous materials, in: Nanoporous Materials: Science and Engineering, 2004, pp. 317–364, http://dx.doi.org/10.1142/9781860946561_0011.

[61] S. Monti, G. Barcaro, W.A. Goddard, A. Fortunelli, Diverse phases of carbonaceous materials from stochastic simulations, ACS Nano 15 (4) (2021) 6369–6385, http://dx.doi.org/10.1021/acsnano.0c08029.

[62] G. Conter, S. Monti, G. Barcaro, W.A. Goddard, A. Fortunelli, Functionalized amorphous carbon materials via reactive molecular dynamics simulations, ACS Appl. Mater. Interfaces 16 (36) (2024) 48043–48057, http://dx.doi.org/10.1021/acsami.4c06527.

[63] D.D. Do, E.A. Ustinov, H.D. Do, Porous texture characterization from gas-solid adsorption, in: Adsorption By Carbons, Elsevier, 2008-01-01, pp. 239–271, http://dx.doi.org/10.1016/B978-008044464-2.50015-8, URL https://www.sciencedirect.com/science/chapter/edited-volume/abs/pii/B9780080444642500158.

[64] H. Tanaka, H. Kanoh, M. Yudasaka, S. Iijima, K. Kaneko, Quantum effects on hydrogen isotope adsorption on single-wall carbon nanohorns, J. Am. Chem. Soc. 127 (20) (2005-05-25) 7511–7516, http://dx.doi.org/10.1021/ja0502573, Publisher: American Chemical Society.

[65] J.W. Leachman, R.T. Jacobsen, S.G. Penoncello, E.W. Lemmon, Fundamental equations of state for parahydrogen, normal hydrogen, and orthohydrogen, J. Phys. Chem. Ref. Data 38 (3) (2009) 721–748, http://dx.doi.org/10.1063/1.3160306.

[66] https://github.com/AlexMoreo/porous-materials.

[67] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, in: Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, US, 2015.

[68] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E.Z. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, PyTorch: An imperative style, high-performance deep learning library, in: Proceedings of the 32nd Conference on Neural Information Processing Systems, NeurIPS 2019, 2019, pp. 8024–8035.

[69] H.G.T. Nguyen, C.M. Sims, B. Toman, et al., A reference high-pressure CH4 adsorption isotherm for zeolite Y: results of an interlaboratory study, Adsorption 26 (2020) 1253–1266, http://dx.doi.org/10.1007/s10450-020-00253-0.

[70] A. Tavitian, D.P. Broom, M.D. Gee, M.J. Benham, S. Brandani, E. Mangano, Validated methane adsorption isotherms up to 7.5 MPa on a reference Na-Y zeolite at near ambient temperatures, Adsorption 30 (2024) 1529–1538, http://dx.doi.org/10.1007/s10450-024-00518-y.