Electronic Supplementary Material (ESI) for RSC Advances

Supplementary information:

Prediction of thermodynamic properties of adsorbed gases on Zeolitic Imidasolate frameworks using a molecular simulation-QSPR approach

Hedi Amrouche[1,2], Benoit Creton[1], Flor Siperstein[2], Carlos Nieto-Draghi[1,*]

[1]IFP Energies nouvelles, 1 et 4, Avenue de Bois-Préau 92852 Rueil-Malmaison, France,
[2]School of Chemical Engineering and Analytical Science The University of Manchester, PO
*carlos.nieto@ifpen.fr

1 Solid structures

The structures used are distributed around three different topologies : The SOD structures includes ZIF-8 ($C_4H_5N_2$, **mIm**), ZIF-90 ($C_4H_3N_2O$, Ica), ZIF-COOH ($C_4H_3N_2O_2$,**carboIm**), ZIF-NO_2 ($C_3H_2N_4O_2$, **nIm**) and ZIF-Cl ($C_3N_2H_2Cl$, **cIm**). The functional groups are placed on position 2 of the imidazolate linker. In ZIF-8, sodalite cages are connected by 6-rings, i.e. windows with 6 Zn atoms. The free diameter of the sodalite cages is about 1.16 nm. The imidazolate linkers point towards the centre of the 6-ring, which connects two sodalite cages. The free diameter of the opening of the 6-ring is about 0.3-0.34 nm. The RHO topology contains four structures : ZIF-71 ($C_3H_2N_2Cl_2$ **dcIm**), ZIF-93 ($C_5H_6N_2O$ **almeIm**), ZIF-96 ($C_4H_4N_4$ **cyamIm**) and ZIF-97 ($C_5H_8N_2O$ **hymeIm**). The functional groups are placed in the -4 and -5 positions. According the literature[1],[2],[3],[4], these positions are favourable to produce ZIFs with RHO topology. This topology is constructed from a body-centered arrangement in which the largest cage is [$4^{12}.6^8.8^6$] cage. (the symbol [....m^n....] indicates that n faces of the cage are m-membered rings). Their structures and their capacity to capture carbon dioxide have been investigated by W. Morris et al[5], who provides solid structures that were used in this work. The GME series is made up of five solids: ZIF-68 ($C_7H_5N_2$ **bIm**), ZIF-69 ($C_7H_4N_2Cl$ **cbIm**), ZIF-78 ($C_7H_5N_3O_2$ **nbIm**), ZIF-79 ($C_8H_8N_2$ **mIm**), ZIF-81 ($C_7H_5N_2Br$ **bbIm**). GME structures are made up of two different linkers: $C_3H_2N_4O_2$ **nIm**

linkers, which are common to the whole GME family, and another substituted bIm (which are **bIm, cbIm, nbIm, mIm** and **bbIm**). As illustrated in Figure 1, the GME structures can be described as the assembling of *kno* [$4^3.8^3.12^2$] cages forming channels of 12-membered rings, cross-linked *gme* [$4^9.6^2.8^3$] cages and *hpr* [$4^6.6^2$] cages in a 1:1:1 ratio. The **nIm** links occupy the same position in all solids (two edges which are part of the *hpr* cages), the others substituted bIms occupy the remaining edges. The substituted bIm's point into the voids of the *kno* cage. Consequently, depending on the functional groups, the pore diameter of the *kno* cage vary from 7.3 to 9.9 Å. Due to their adjustable pore size and highly electrostatic nature the GME are one of the most studied ZIFs in the literature[6,7,8,9,10]. A detailed information of the volume fraction of different pores an channels of GME structure is provided in Table B.

An interesting point for discussion is the permeability of the selected structures for all gases considered. The concept of pore window is somewhat artificial for judging the permeability of ZIF materials, due to the high flexibility of the organic linkers. In fact, ZIF-7 and SIM-1 are ZIF materials with small window apertures (≤ 3 Å) that allow CO_2 and bulky hydrocarbon molecules to diffuse inside their pores.[11,12,13]

The topological and structural information of the ZIF solids studied in this chapter are listed in Table B. The molecular structure of each material was taken from the CCDB[14] or from the original references. In certain cases, where hypothetical solids are considered, the framework has been optimized through VASP periodic DFT calculations.[15]

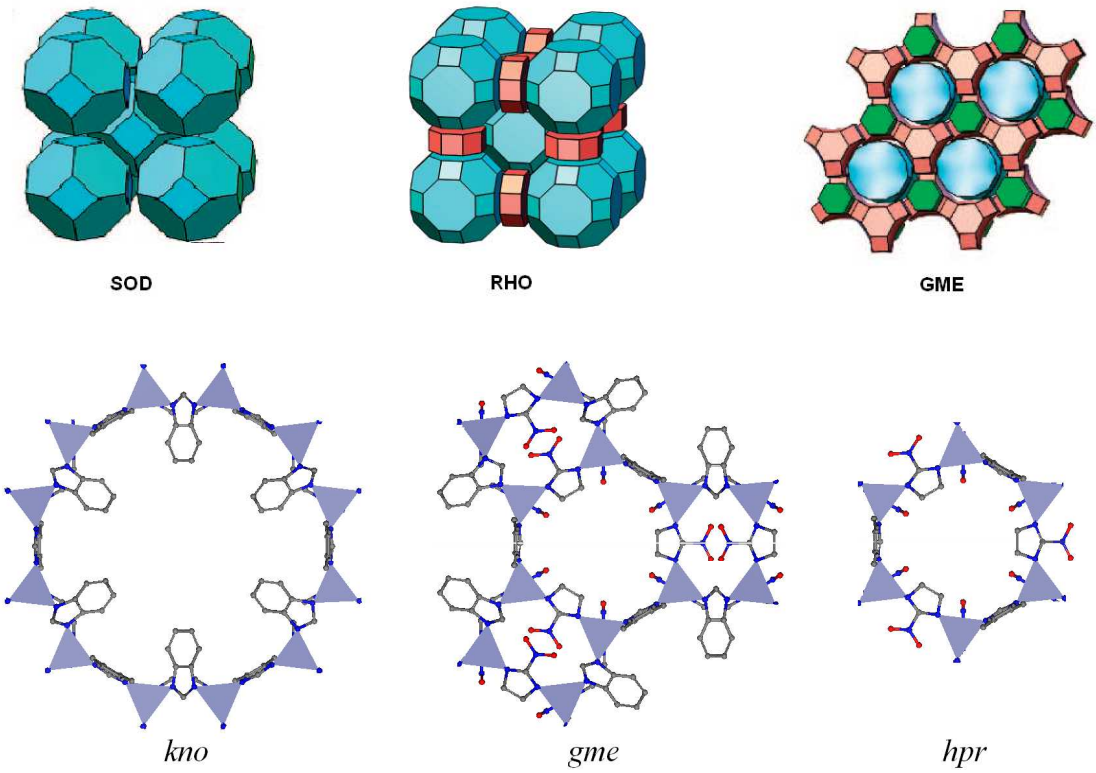


Figure 1 : The tiling of the SOD, RHO and GME structures representing the subdivisions of space (top). Schematic representation of the different types of cages present in GME : *kno*, *gme* and *hpr*. Image of topologies were taken from ref.[3],ref.[5] and ref.[2], respectively.

Tab. A: Details of pore volume fraction (ϑ) for different channels (hpr) and pores (kno) of the GME structures (see Figure 1 for details).

gme	μ RbIm (D)	μ nIm (D)	V_T of *hpr* (Å^3)	V_T of *kno* (Å^3)	V-free (Å^3)	$\vartheta_{channel}$	ϑ_{pore}	Total |μ|
ZIF-68	1.48	7.44	12966.98	2031.16	14998.14	0.86	0.14	6.63
ZIF-79	1.72	8.67	8509.99	829.16	9339.15	0.91	0.09	8.05
ZIF-81	1.81	8.02	11879.74	985.47	12865.21	0.92	0.08	7.55
ZIF-69	1.88	8.02	12320.75	924.74	13245.49	0.93	0.07	7.59
ZIF-78	5.51	9.25	15722.25	924.74	16646.99	0.94	0.06	9.04

μ RbIm and μ nIm are the dipole moment of the substituted (R) bIm and nIm linkers. VT hpr and VT kno corresponding to the total pore volume occupy for the hpr channel and kno pores in GME. V-free corresponds to the total free volume of the GME structure.

Tab. B . List of ZIFs materials used in the QSPR database.

Name	Topology	Organic Linker	Cell volume [Å^3]	d_p [Å]	S_a [Å^2/g]	$\lvert\vec{\mu}_{OL}\rvert$ [D]	$\lvert Q_{OL}\rvert$ [D.m]
ZIF-8	SOD	meIm	4917.50	11.47	1395.14	1.22	8.97
ZIF-90	SOD	IcaIm	5080.34	10.88	1269.42	3.51	13.46
ZIF-COOH*	SOD	carboIm	5131.07	10 .99	1261.1	1.84	13.83
ZIF-NO_2*	SOD	nIm	4980.17	10.61	949.78	3.40	12.65
ZIF-Cl	SOD	cIm	4950.98	11.10	1200.41	0.33	3.56
ZIF-71	RHO	dcIm	23280.71	17.76	1137.2	0.52	10.30
ZIF-93	RHO	almeIm	22801.20	17.04	968.13	5.65	25.54
ZIF-96	RHO	cyamIm	22800.96	16.96	1251.7	5.20	28.02
ZIF-97	RHO	hymeIm	22983.57	16.48	872.67	5.04	12.72
ZIF-68	GME	bIm	22871.06	9.90	1060.83	1.48	16.23
ZIF-69	GME	cbIm	22871.06	7.62	1003.1	1.88	17.05
ZIF-78	GME	nbIm	22871.06	7.62	928.77	5.51	19.60
ZIF-79	GME	mbim	22871.06	7.34	1012.54	1.72	10.49
ZIF-81	GME	bbIm	22871.06	7.78	984.41	1.81	17.26

dp is the pore diameter calculated with the material Studio package. S_a is the accessible surface area calculated using the method developed by Düren *et al*[16] . $\lvert\vec{\mu}_{OL}\rvert$ and $\lvert Q_{OL}\rvert$ are, respectively, the dipole and quadrupole moment of the organic linkers (see next section for details.). *Hypothetical solids obtained by periodic DFT simulations

1.1 Correlative QSPR Method

We can identify two main types of descriptors, some associated with the solid and some associated with the adsorbed gases. With this procedure we obtain about 60 descriptors which are ready to feed into our database. Due to the large amount of data care must be taken to avoid redundant information. The correlation between descriptors has been analyzed as follows.

1.1.1 Correlation Between Descriptors

Given that the total number of molecular descriptors outweighs the number of molecules in the database (concerning solid linkers and gases), only the most relevant descriptors are retained for building the predictive QSPR equations. For example, descriptors which are highly correlated, i.e., hold essentially the same information, with respect to the target thermodynamic property, can be kept. Therefore, a correlation matrix was constructed. This $m \times m$ matrix, with m equal to the number of descriptors plus one for the simulated property, contains the correlation coefficients as defined by Eq. 1 :

Eq. 1

$$r_{(X,Y)} = \frac{C_{(X,Y)}}{\sqrt{V_X V_Y}},$$

where X and Y denote descriptors or simulated values of the target property. V and C are respectively the variance and the covariance. Their expressions are detailed in Eq. 2 and Eq. 3 where n runs over the number of molecules of the data set.

Eq. 2

$$V_X = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \langle X \rangle\right),$$

and

Eq. 3

$$C_{(X,Y)} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \langle X \rangle\right)\left(Y_i - \langle Y \rangle\right),$$

For two descriptors where X and Y are strongly correlated, $r_{(X,Y)}$ is close to 1. Using the correlation matrix, only the descriptors that correlate well with the target property and poorly correlate with another descriptor can be conserved. Materials Studio 5.0 software was used in this work to analyze the correlation between descriptors as well as to perform the QSPR models.[17]

1.1.2 Multivariance Analysis

In most cases, QSPR studies consist of determining an equation which gives a reliable property reproduction. In this work, we have chosen the target predictive model to be linear as described by Eq. 4

Eq. 4

$$P_{calc.} = \lambda_0 + \sum_i \lambda_i X_i \,,$$

where *i* runs over 1 to 5, which corresponds to the maximum number of descriptors admitted in the predictive equation and λ_i are weight factors. The method used to select the representative descriptors and to optimize the associated constants is the Genetic Function Approximation (GFA) available in the Materials Studio software. This approach starts by establishing an initial population of equations randomly chosen. The equation terms are viewed as strings, and the population evolves through iterative operations: selection, crossover, and mutation. During the evolution process, the constructed equations are scored using a slightly modified Friedman's lack-of-fit (LOF) method[17] evaluated by:

Eq. 5

$$LOF = \frac{SSE}{M\left[1 - \lambda\left(\frac{c + dp}{M}\right)\right]^2},$$

where *SSE* is the sum of squares of errors, *c* is the number of terms in the models, *d* a scaled smoothing parameter, *p* the total number of descriptors contained in all model terms, *M* the

number of samples in the training set and λ is a safety factor, it ensure that the denominator of the expression can never be zero. By adding more terms to the regression models, the SSE may reduce, but also the values of c and p increase. Consequently it tends to increase the LOF score. Thus the choice of LOF model favour the equations with a low number of descriptors. The equations evolve until the convergence is observed, i.e., the scores are not further improved. Among all returned equations the best predictive equation is the one having the lowest Mean Absolute Error (MAE), defined as:

Eq. 6

$$\mathbf{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|P_{sim.} - P_{calc.}\right|,$$

where, *i* runs over the n compounds of the data set, $P_{sim.}$ and $P_{calc.}$ are respectively the experimental and the calculated property. Figure 2 represents a schematically the evolution of equations on the GFA approach.

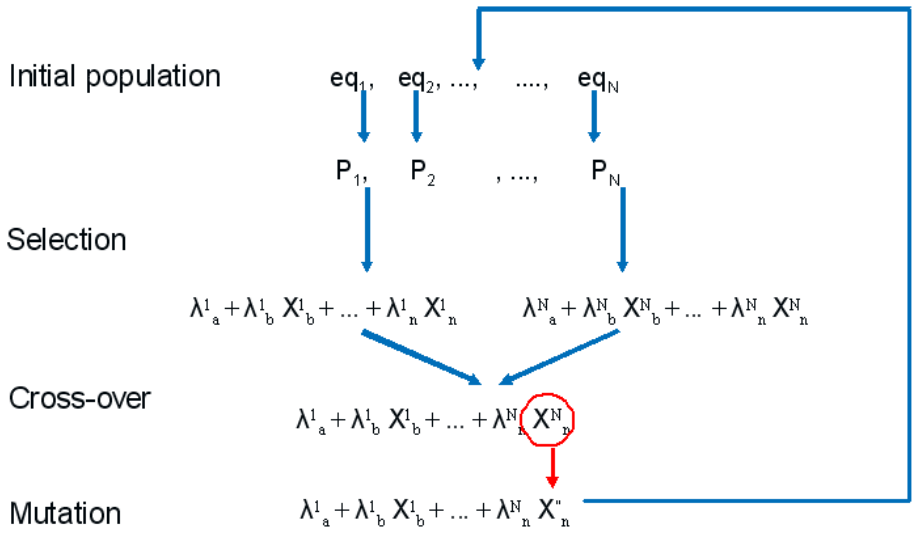


Figure 2: Schematic representation of equation evolution on the GFA approach.

Before carrying out calculations using the GFA approach, the total database was splitted into two data subsets:

The *training set* which represents 90% of the database. This is intended to build the predictive models, and is chosen to be representative of the solids and gases in the database;

The *test set* constituted from the remaining solid/gas couples of the database. There are viewed as external values and used to test the predictive power of the equations and to select the best predictive model. This Training/Test sets ratio used is commonly used in QSPR studies.[18],[19]

2 Simulation details

2.1 Simulation Methods:

The ESP fitting methodology has been applied by using the Jaguar[20] package: the B3LYP functional combined with the pseudo potential LanL2DZ for the transition metal and the double-ζ basis set 6 31G** for the rest of the atoms was applied.

To design the isoreticular series and localise the crystallographic positions of each solids, periodic DFT optimisation were performed using VASP code[21]. Then the atomic positions were frozen, and the host-guest interaction energy grid was constructed before the MC simulation.

To compute fugacity simulations in the isobaric-isothermal ensemble (NPT) combined with the particle insertion method were performed[22] for pressure choose. Probabilities for MC bias move were set to 0.63 for rigid-body translations, 0.35 for rigid-body rotations (except for monoatomic gases such as Ar and CH_4), and 0.02 for volume changes.

The complete list of force field parameters to describe the solid-gas interactions are listed in Tables C to G

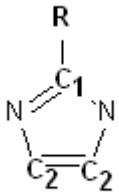
Tab. C: ESP partial charges for GME structures presented in Table B. The charges of the R-bIm, b) charges of the nIm

bIm	ZIF-Cl	ZIF-Br	ZIF-CH3	ZIF-NO2	ZIF-H	Moyenne	Variance
Zn	1.07	1.04	1.08	1.02	1.095	1.0525	0.0008
N	-0.63	-0.63	-0.6	-0.66	-0.67	-0.6300	0.0006
C_1	0.55	0.5	0.5	0.55	0.51	0.5250	0.0008
C_2	0.225	0.1	0.2	0.33	0.25	0.2138	0.0089
C_3	-0.49	-0.245	-0.12	-0.38	-0.19	-0.3088	0.0259
C_4	-0.14	-0.175	-0.35	-0.3	-0.13	-0.2413	0.0100
C_5	0.008	0.01	0.27	0.08	-0.13	0.0920	0.0152
C_6	-0.13	-0.37	-0.4	-0.38	-0.19	-0.3200	0.0162
R	**ZIF-Cl**	**ZIF-Br**	**ZIF-CH3**	**ZIF-NO2**	**ZIF-H**	**Moyenne**	**Variance**
Cl	-0.2	-	-	-	-	-0.2	-
Br	-	-0.2	-	-	-	-0.2	-
N	-	-	-	0.76	-	0.76	-
O	-	-	-	-0.51	-	-0.51	-
C	-	-	-0.15	-	-	-	-
H	-	-	0.02	-	-	-	-

nIm	ZIF-Cl	ZIF-Br	ZIF-CH3	ZIF-NO2	ZIF-H	Moyenne	Variance
Zn	1.07	1.04	1.08	1.02	1.095	1.0610	0.0009
N	-0.5	-0.45	-0.48	-0.45	-0.53	-0.4820	0.0012
C1	0.52	0.5	0.52	0.46	0.54	0.5080	0.0009
C2	-0.12	-0.14	-0.1	-0.085	-0.1	-0.1090	0.0005
C3	-0.12	-0.14	-0.1	-0.085	-0.1	-0.1090	0.0005
N	0.62	0.62	0.6	0.66	0.6	0.6200	0.0006
O	-0.48	-0.48	-0.5	-0.49	-0.46	-0.4820	0.0002

Tab. D: ESP partial charges for SOD structures presented in Table B.

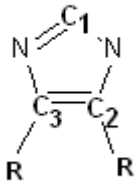
SOD	ZIF-8	ZIF-COOH	ZIF-90	ZIF-NO2	ZIF-Cl	Moyenne	Variance
Zn	1.100	0.900	0.850	0.792	0.790	0.945	0.0117
N	-0.540	-0.290	-0.310	-0.120	-0.170	-0.3875	0.0130
C1	0.640	0.435	0.480	0.450	0.470	0.5212	0.0077
C2	-0.080	-0.150	-0.200	-0.300	-0.320	-0.135	0.0027
H	0.144	0.133	0.173	0.172	0.173	0.1541	0.0003
R	-CH3	-COOH	-HCO	-NO2	-Cl		
Cl	-	-	-	-	-0.230	-	-
C	-0.670	0.570	0.300	-	-	-	-
O	-	-0.620	-0.520	-0.530	-	-	-
N	-	-	-	0.710	-	-	-
H	0.144	0.400	-0.030	-	-	-	-



SOD

Tab. E: ESP partial charges for RHO structures presented in Table B.

RHO	ZIF-71	ZIF-93	ZIF-96	ZIF-97	Moyenne	Variance
Zn	0.94	1.05	1.18	1.108	1.0695	0.0103
N_1	-0.4	-0.5	-0.56	-0.34	-0.4500	0.0097
C_1	0.51	0.32	0.21	0.31	0.3375	0.0157
N_2	-0.4	-0.5	-0.56	-0.67	-0.5325	0.0128
C_2	0.09	0.35	0.305	0.49	0.3088	0.0275
C_3	0.09	-0.1	-0.1	-0.36	-0.1175	0.0342
H_1	-0.06	0.055	0.085	0.046	0.0315	0.0040
R	**$(-Cl)_2$**	**$-HCO/-CH_3$**	**$-CN/-NH_2$**	**$-CH_3/-CH_2OH$**		
Cl	-0.15	-	-0.92	-	-	-
H/C/O	-	-0.02/0.45/-0.56	-	-	-	-
C/H	-	-0.50/0.16	-	-0.76/0.18	-	-
C/N	-	-	0.39/-0.6	-	-	-
N/H	-	-	-0.7/0.35	-	-	-
$C/H_2/O/H$	-	-	-	0.52/0.18/-0.68/0	-	-



RHO

The vdW potential parameters for the force field used are summarized on Tab. F:

Tab. F : vdW potential parameters for the elements used in this work

Non-bonded interactions		
Atom Type	ε(K)	σ(nm)
Zn	43.084	0.2338
N	23.974	0.3997
O	20.847	0.3118
Cl	78.872	0.3516
C	36.483	0.3259
H	15.288	0.2440

The gas models used during this work derive from several sources:

Tab. G : List of force fields (FF) used in the QSPR analysis to represent the adsorbed gases

Molecule name	Formula	Polar nature	T_b[K]	Type of FF/family	Ref.
Argon	Ar	non polar	87.3	All atoms	[23]
Methane	CH_4	non polar	111.6	United atom	[24]
Ethane	C_2H_6	non polar[a]	184.5	Anisotropic United Atom (AUA4)	[25]
Molecular nitrogen	N_2	quadrupolar	77.9	All atoms	[26]
Molecular oxygen	O_2	quadrupolar	90.0	All atoms	[27]
Carbon dioxide	CO_2	quadrupolar	186.5	All atoms (EMP2)	[28]
Carbon monoxide	CO	dipolar	77.9	All atoms	[29]
Sulfur dioxide	SO_2	dipolar/quadrupolar	263.0	All atoms	[30]
Hydrogen sulfide	H_2S	dipolar/quadrupolar	211.0	All atoms	[31]
Aceto-nitrile	CH_3CN	dipolar	383.0	Hybrid (united atom for CH_3)	[32]
Water	H_2O	dipolar/quadrupolar	263.0	Hybrid TIP4P	[33]

2.2 Adsorption simulation details

To model the gas adsorption at low coverage, Grand Canonical Monte Carlo (GCMC) simulations combined with a bias scheme for the insertion of the centre of mass off the guests molecules were performed with the Gibbs Code v.8.3[34]. The probabilities bias moves[35] were used with arbitrary proportion: 0.35 for the rigid-body translation, 0.10 for rigid-body rotations (except for monoatomic gases such as Ar or CH_4), and 0.55 for the insertions or deletions.

All simulations were performed in a simulation box incorporating 2x2x2 unit cells. LJ interactions and real-space electrostatic contributions were calculated by using a cut off radius of ~17Å. No LJ tail corrections were considered[36], but standard long-range electrostatic interactions were calculated by using the Ewald methodology with ten vectors on the reciprocal space and a screening factor nα of 2.5.

2.3 Volumes and surfaces

The free volume was computed by using the volume integral in Eq. 7 in which is the LJ interaction between a single helium atom is (ε_{He}=10.22 K ; σ_{He}=0.258 nm) and the complete structure of the adsorbent:

Eq. 7
$$V_{free} = \int \exp(-U_{ads}^{He}(r)/k_bT)dr$$

The porosity Φ is simply the ratio between the free volume (V_{free}) with respect to the total unit cell volume (V_{uc}) (Eq. 8):

Eq. 8
$$\phi = \frac{V_{free}}{V_{uc}}$$

The accessible surface area was computed from Monte Carlo integration technique using a "probe" molecule which is rolled along the surface of the solid. Then the surface area is obtained from the ratio of the probe molecules that did not overlap with other framework atoms to the total number of attempts.

2.4 Isosteric heat of adsorption computation

Isosteric heats of adsorption q_{st} were calculated from Eq. 9[37] in which H_b is the enthalpy of the bulk phase and U_a is the energy of the adsorbed phase.

Eq. 9

$$q_{st} = H_b - \left[\frac{\delta U_a}{\delta N}\right]_{T,V}$$

By using a fluctuations method with ideal gas assumptions under Henry's regime, the isosteric heats can be readily calculated from GCMC simulation according to Eq. 10, in which U^s_{ext} is the intermolecular energy of the adsorbed phase and N is the number of adsorbed molecules. The angle brackets denote averages in the grand canonical ensemble.

Eq. 10

$$q_{st} = RT - \frac{\left\langle U^s_{ext} N\right\rangle - \left\langle U^s_{ext}\right\rangle\left\langle N\right\rangle}{\left\langle N^2\right\rangle - \left\langle N\right\rangle^2}$$

Experimentally, the isosteric heat of adsorption at low coverage can be estimated by means of the Van't Hoff equation (Eq. 11):

Eq. 11

$$\left[\frac{d\ln P}{dT}\right]_N = \frac{\Delta H_0}{RT^2},$$

A series of adsorptions isotherms at differents temperatures are required. Then at constant coverage, the pressure is derivate as a function of the temperature as presented in Eq. 11.

3 Results:

3.1 Analysis of the q_{st} simulation results in function of the different descriptors

Therefore a series of analysis are proposed, where $q_{st}°$ is studied as a function of variations in the magnitude of a single descriptor. As an example the two descriptors proposed are the organic linker dipolar moment $|\vec{\mu}_{OL}|$ and the gas atmospheric boiling temperature T_b.

3.1.1 Variation of q_{st} with $|\vec{\mu}_{OL}|$

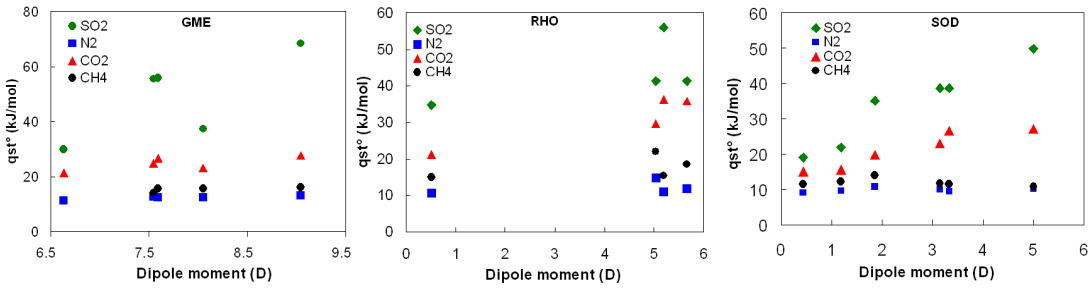


Figure 3 : Variation of the isosteric heat of adsorption for some selected gases as function of $|\vec{\mu}_{OL}|$ for the three topologies studied.

In Figure 3, we studied the variation of the isosteric heat of adsorption as a function of $|\vec{\mu}_{OL}|$. Each topology was treated separately and we choose to present only the main representative gases owing different polarities. The influence of the linkers dipolar moment $|\vec{\mu}_{OL}|$ is not the same for the different gases. As expected for non polar gas, such as CH_4, the dipolar moment does not impact on the $q_{st}°$. For polar gases, the higher the linker $|\vec{\mu}_{OL}|$, the higher will be the $q_{st}°$. Moreover, the analysis reveals that the degree of polarity of the gas is important. For example, gases having weak quadrupolar moment (polarity), such as N_2, are less influenced by solids having highly polar linkers than CO_2 or SO_2 (which possess high

quadrupole and dipole moment respectively). The higher the gas polarity, the higher will be the increase of $q_{st}°$ when we increase $\left|\vec{\mu}_{OL}\right|$. We can note that this tendency is strongest for dipolar gases than for quadrupolar ones.

3.1.2 Variation of q_{st} with T_b

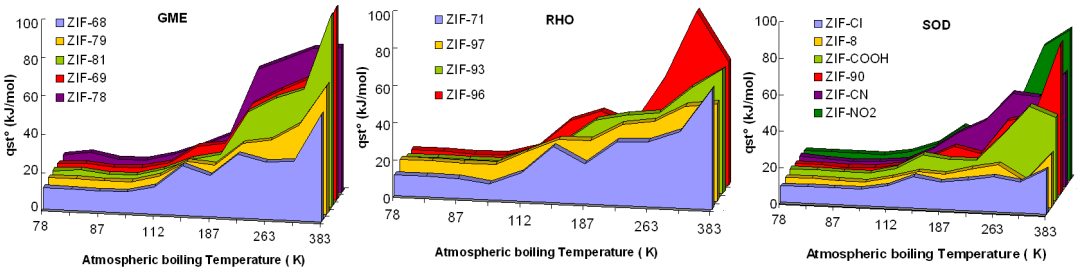


Figure 4: Variation of the isosteric heat of adsorption of all gases studied as function of T_b and the three topologies studied.

The evolution of $q_{st}°$ as function of T_b for all gases was investigated in Figure 4. Each topology was also treated separately for the same reasons explained above. According to the results, the higher the T_b, the higher will be $q_{st}°$. The value of T_b is correlated with the cohesive energy intrinsic to the gas molecules. We can extend this intrinsic cohesive energy to the surface-gas energy. Thus, the higher the intrinsic cohesive energy, the higher the surface-gas and then, the higher $q_{st}°$ will be. As expected the influence of T_b influence seems independent to the solid topology.

3.1.3 Variation of q_{st} with *H*

Similar analyses for other descriptors are complicated due to the difficulty in isolating their effect over $q_{st}°$. We compared different topologies having different pores sizes, but also having different organic linker with different chemical nature (different n_{fg}). It is interesting to notice that with some rare exceptions (RHO ZIF-93[5] and SOD SIM-1 have **almeIm** linker[], it is difficult to obtain two different framework topologies with the same organic linkers. However, one may expect that materials with narrow pores would increase the confinement effect (this is the case for deallumined zeolites, where the surface chemistry is composed exclusively by Si and O).[13]

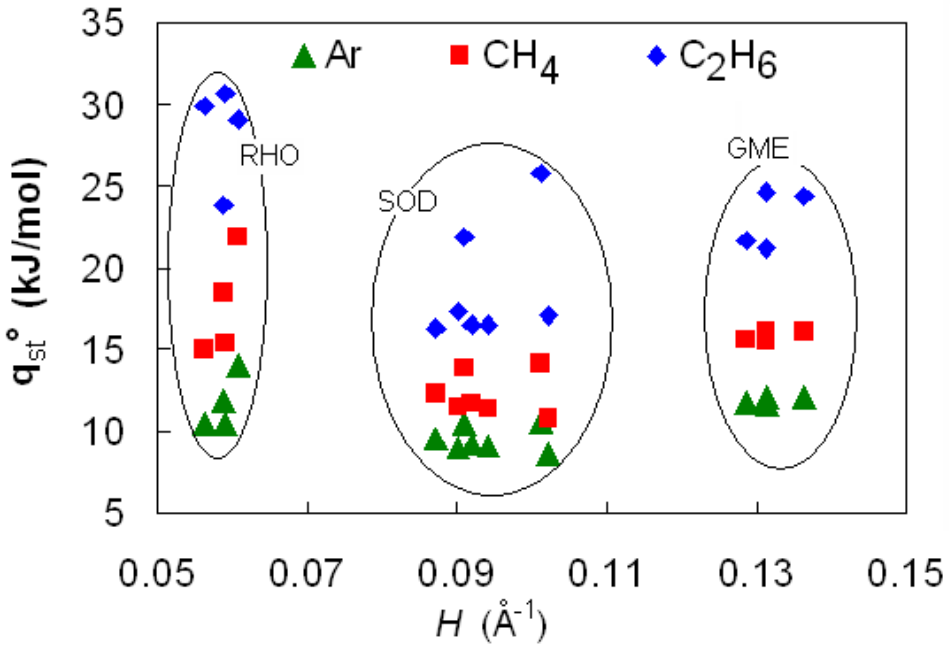


Figure 5: Variation of the isosteric heat of adsorption as function of the mean curvature H for non polar gas. The topology corresponding to the pore curvature points have been indicated.

The variation of $q_{st}°$ with the mean curvature can be observed in Figure 5. We selected non polar gases to avoid the solid-gas electrostatic interactions. The variation of the isosteric

heat is not monotone. The solid topology and the chemical nature of the organic linkers affect the results obtained. Indeed the mean curvature are grouped by topology, the smaller values of the *H* are obtained for the RHO, followed by the SOD and finally by the GME structures. The q_{st}° tends to decrease between *H* values corresponding to RHO and SOD solids, and increase between the SOD-GME *H* values. In addition, it is important to notice that solids with RHO topologies have bi-substituted organic linkers (n_{fg} = 2), while SOD and GME possess linkers with only one functional group (n_{fg} = 1). It is possible that although having the biggest pore diameters (low *H* values), the RHO solids have strongest vdW interactions than the other two solids due to the additional functional group. This example reveals the limit of descriptor *H* to appropriately characterize the vdW interactions of the different solids compared in this study.

3.2 Analysis of the QSPR model for q_{st}

It is important to keep in mind that the QSPR method used in our analysis (GFA) impose a linear relation between descriptors (Eq. 4). However, it is possible to introduce some degree of non linear behavior by considering mathematical transformations of the descriptors used in the database (by taking the inverse, the natural logarithm or the exponential of the magnitudes). In Table H we summarized the response of the target properties (increase or decrease) when a single descriptor is increased or decreased while keeping constant the other ones.

After taking into account the expected trends shown in Eq. 4, the most appropriate equation obtained with the GFA procedure describing q_{st}^{0} is shown below,

Eq. 12 $$q_{st}^{\circ} = \lambda_0 + \left(\lambda_1 \left|Q_{OL}\right| + \lambda_2 \left[\mathrm{n_{fg}} \cdot H\right]\right)_s + \left(\lambda_3 \ln\left(T_b\right)\right)_g + \left(\lambda_4 H \cdot \left[\left|\vec{\mu}_{OL}\right| \cdot \left|\vec{\mu}_g\right|\right]\right)_{s/g},$$

Tab. H . Expected response of the target property (q_{st}^0) to an increase or decrease of individual descriptors. Increments and decrements are represented by (↑) and (↓) arrows, respectively.

Descriptor	Sense of variation	Expected response for q_{st}^0
$\lvert\vec{\mu}_{OL}\rvert$	↑	↑
	↓	↓
$\lvert Q_{OL}\rvert$	↑	↑
	↓	↓
$\lvert\vec{\mu}_{g}\rvert$	↑	↑
	↓	↓
T_b	↑	↑
	↓	↓
n_{fg}	↑	↑
	↓	↓
K or H	↑	↑
	↓	↓

with H the mean curvature, n_{fg} the number of functional groups per organic linker, $\lvert\vec{\mu}_{OL}\rvert$ and $\lvert\vec{\mu}_{g}\rvert$ are the organic linker and gas molecular dipolar moment respectively, $\lvert Q_{OL}\rvert$ is the module of the quadrupolar matrix of the organic linker and $\ln(T_b)$ is the natural logarithm of the atmospheric boiling temperature of the adsorbed gas. We can classify the selected descriptors according to three type of interactions, i.e.: solid, gas and solid/gas. The different coefficient λ_i obtained for the correlation is reported in Table I. Four solids descriptors were taken into account : the polar nature of the solid (represented by Q_{OL} and μ_{OL}), the van der Waals interaction of the solid represented by the degree of confinement (H and n_{fg}).

Tab. I Constants (λ_i) obtained in the QSPR approach for the prediction of the isosteric heat of adsorption (q_{st}^0) as described with Eq. 12. The dipolar moment are expressed in Debye (D).

	Coefficients	Units
λ_0	*- 97.2651*	*kJ mol-1*
λ_1	*0.52011*	*kJ mol-1 D-1 Å-1*
λ_2	*95.8506*	*kJ mol-1 Å*
λ_3	*20.1537*	*kJ mol-1 K-1*
λ_4	*9.6137*	*kJ mol-1 D-2 Å*

The adsorbed gas is represented by μ_g and T_b, which are respectively a measure of its polar nature and its cohesive energy, i.e. the higher of μ_g or T_b , the stronger will be the gas-gas interactions. The Eq. 12 agrees with the expectations formulated in Table H.

The error obtained to reproduce the q_{st}^0 simulated values over the *test* and the *training* sets are presented in Table J.

Tab. J Mean Absolute Error (MAE) as well as the Mean Absolute Percentage Error (MAPE) obtained with Eq. 12 for the training and test sets.

Set	MAE	MAPE(%)
Test	6.7	23.8
Training	5.6	24.6

The results obtained using Eq. 12 are encouraging. The predicted results agree fairly well with simulated data for both the test and training set. The predictive power of this equation is hence interesting.

The data set can also be divided as a function of the gases polarity. The comparison between the simulation results and the QSPR correlation as a function of the polar nature of gas is illustrated in Figure 6. Its predictive accuracy is analyzed in Table K.

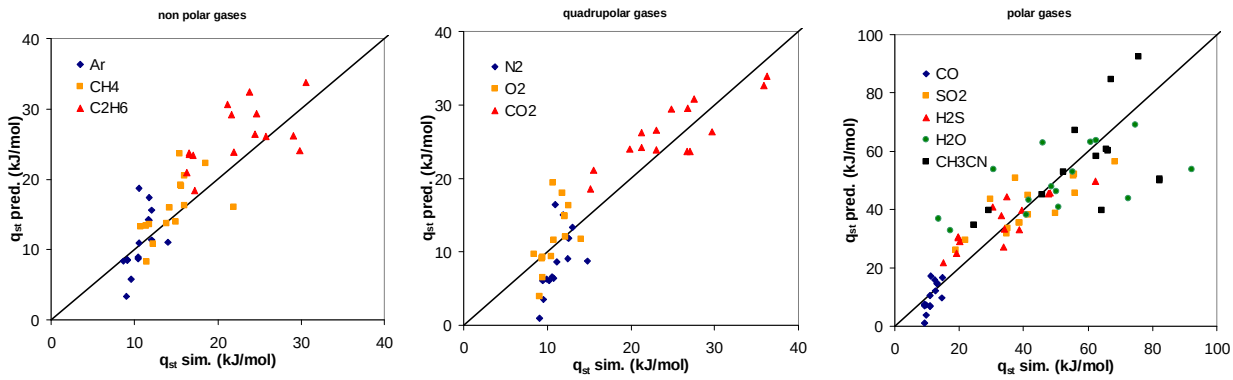


Figure 6 : Comparison of the simulated isosteric heat of adsorption of different gases on different solids with prediction issued by the QSPR model as function of the gas polar nature.

Tab. K . Mean Absolute Error (MAE) as well as the Mean Absolute Percentage Error (MAPE) obtained with Eq. 12 for the different gases studied. The electrostatic nature of each gas is classified according to: non polar (np), quadrupolar (Q) and dipolar (μ_g)

Gas	**(polar nature)**	**MAE**	**MAPE (%)**
Ar	np	2.7	25.2
CH_4	np	2.9	19.5
C_2H_6	np	4.8	23.4
O_2	*Q* (0.99 D.Å)	2.6	23.8
N_2	*Q* (1.80 D.Å)	3.7	34.8
CO_2	*Q* (5.06 D.Å)	3.4	15.1
CO	μ_g (0.03 D)/*Q* (0.06 D.Å)	3.3	30.3
H_2S	μ_g (1.43 D)/*Q* (2.50 D.Å)	6.5	23.8
SO_2	μ_g (1.63 D)/*Q* (7.19 D.Å)	6.7	18.3
H_2O	μ_g (2.18 D)/*Q* (3.04 D.Å)	11.7	34.0
CH_3CN	μ_g (4.12 D)/*Q* (1.19 D.Å)	13.8	21.6

It is not straight forward to discriminate which kind of fluid (non polar, quadrupolar or polar) is better (or worst) described by the model. It seems that there is no particular

correlation between polar nature of the gases and the obtained deviation. We can notice however that the lowest MAPE deviations are for CO_2 and CH_4 and the highest for H_2O and N_2. We can observe the following order from lowest to highest MAPE,

$CO_2 < SO_2 < CH_4 < CH_3CN < C_2H_6 < O_2 < H_2S < Ar < CO < H_2O < N_2$

The main deviations obtained in function of the solid topology is presented in Table L. Our analysis reveals that the QSPR model for $q_{st}°$ is slightly more accurate for the family of solids with GME topology (17.8%), followed by the RHO (23.5%) and SOD (33.1%) topologies respectively. The main difference between families (excepting the evident crystallographic topology) is the degree of confinement and the type of organic linker. GME has two types of linkers, the nIm is common to all solids and it is exposed in the small channel and, substituted bIm which are distributed in the large cavities. The GME have the smallest pores of the three solids studied as illustrated in the Figure 5. The RHO family has the largest pores but every organic linker possess two functional groups in positions -4 and -5 instead of one for the other two families. Finally the SOD family have pores sizes ranging in between of those GME and RHO and have an unique type of functionalized Im linker in position -2. It is extremely difficult to find a physical explanation that relates the observed tendency of the QSPR model with any structural difference between the solids studied here

Tab. L . Mean Absolute Error (MAE) as well as the Mean Absolute Percentage Error (MAPE in %) obtained with Eq. 12 for the different solids studied for all gases studied.

***Solid*[a]**	**MAE**	**MAPE(%)**
ZIF-68	5.8	21.0
ZIF-69 (Cl)	4.0	13.4
ZIF-78 (NO_2)	6.7	20.4
ZIF-79 (CH_3)	6.2	17.2
ZIF-81 (Br)	4.6	16.8
Average GME	**5.4**	**17.8**
ZIF-8 (CH_3)	7.1	49.3
ZIF-90 (OH-CH_3)	5.9	21.8
ZIF-(Cl)	6.6	50.1
ZIF-(COOH)[(b)]	3.5	18.5
ZIF-(NO_2)[(b)]	8.4	25.6
Average SOD	**6.3**	**33.1**
ZIF-71 (Cl2)	5.2	19.8
ZIF-93 (CO-CH_3)	4.2	22.2
ZIF-96 (CN-NH_2)	9.2	34.4
ZIF-97 (CH_3-OH)	3.8	17.4
Average RHO	**5.6**	**23.5**

[(a)]chemical functionalization is included for clarity, [(b)] virtual solids.

It is important to mention that, if the QSPR method is only fed with CO_2 and the SOD family of solids, a simple correlation is obtained between the exponential of dipolar moment of the solid (μ_{OL}) and q_{st}^0, such as the behaviour observed in our previous work.[15] If the analysis is extended to CO_2 and CH_4 adsorbed over the complete list of solids, a more complex equation is obtained,

Eq. 13 $$\ln\left(q_{st}^{\circ}.H\right) = 0.08\left|\mu_{OL}\right| - 0.035H^{-1} + 0.019Mw_g + 0.001,$$

Where Mw_g is the molecular weight of the gas. In this case q_{st}^0 is predicted with an R^2 about 0.9 and MAE of about 10.7%. As usual, any physical interpretation of the obtained equation is complicated. The important point to keep in mind is the large variation in the form of the resulting equation as a function of the size of the database used. The difference between Eq. 12 and Eq. 13 clearly illustrates this idea.

In the last part of this section we analyse the evolution observed between a couple of descriptors, while keeping the rest of descriptors constant. In Figure 7 we observe the variation of the q_{st}^0 as function of the solid and gas dipolar moment. To better illustrate this idea, we have considered a factious solid with pore diameter of 8 Å and with linker quadrupole moment $|Q_{OL}|$ equal to 28 D Å. The gas is assumed to have a high boiling temperature of 383 K. The highest q_{st}° values are obtained for the combination of both highest values of the organic linker dipolar moment (μ_{OL}) and gas dipolar moment (μ_g). The q_{st}° value is correlated to both gas and solids dipolar moment intensity. The q_{st}° value could also be restricted by one of this dipolar moment intensity. Indeed if one of these descriptors is weak, the q_{st}° value increases slowly until a limit value even if the other descriptor reaches high values. This phenomenon is illustrated by the vertical and horizontal asymptote delimiting the border between two ranges of q_{st}° values at the extreme of dipolar value.

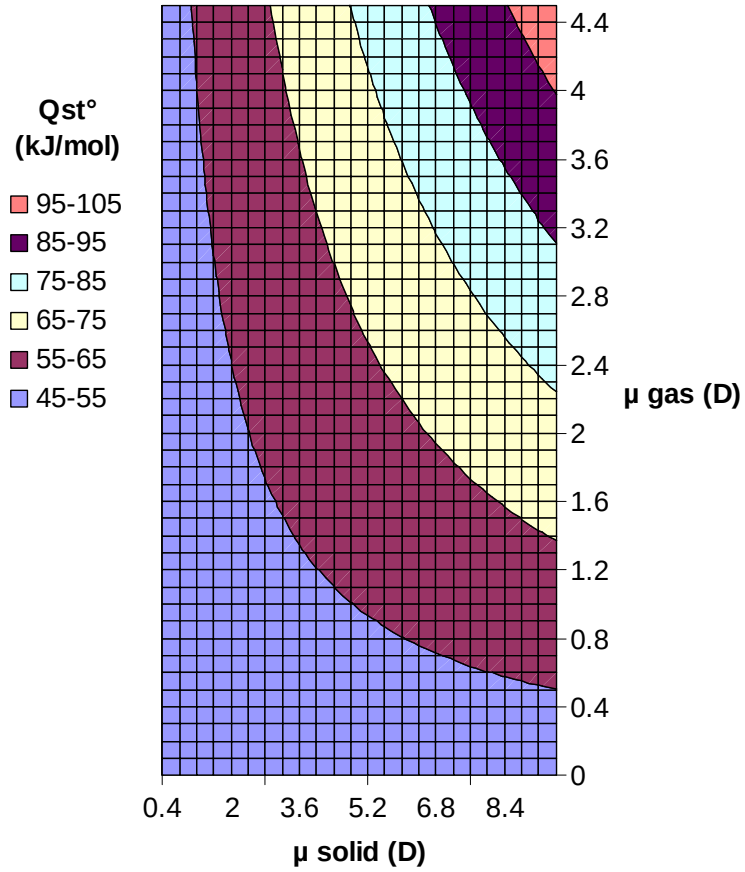


Figure 7. Variation of q_{st}° with the organic linker dipolar moment (μ_{OL}) and gas dipolar moment (μ_g). In this example we have considered a solid with pore diameter of 8 Å and |Qs| of 28 D.Å together with a gas having a high boiling temperature of 383 K.

The isosteric heat of adsorption is directly related to the separation and storage ability of a solid. Thus the selectivity between gases with different dipolar moment could be strongly enhanced using solid with high dipolar moment. Nevertheless, the choice of one material for PSA application should be done carefully, because too high value of q_{st}° will strongly impact on the desorption phase cost.

3.3 Correlative equation using Artificial Neural Network (ANN)

The comparison between the simulation results and the ANN for the isosteric heat of adsorption at zero coverage can be observed in Figure 8. The accuracy of the ANN model can be analyzed from the deviation obtained in reproducing the simulated qst° values. The mean absolute error (MAE) and the MAPE are used to quantify the errors between predicted and simulated values. The lower MAE and MAPE, the better are the predictions. The predicted results agree fairly well with simulated data for both the test and training set (R2 of 0.891, MAE about 3.4 kJ/mol and 5.6 kJ/mol and a MAPE about 13.7 % and 12.3 % for the test and the training set). The ANN model is more accurate than the correlation given in Eq. 12, the problem is that this approach do not provide a useful working equation.

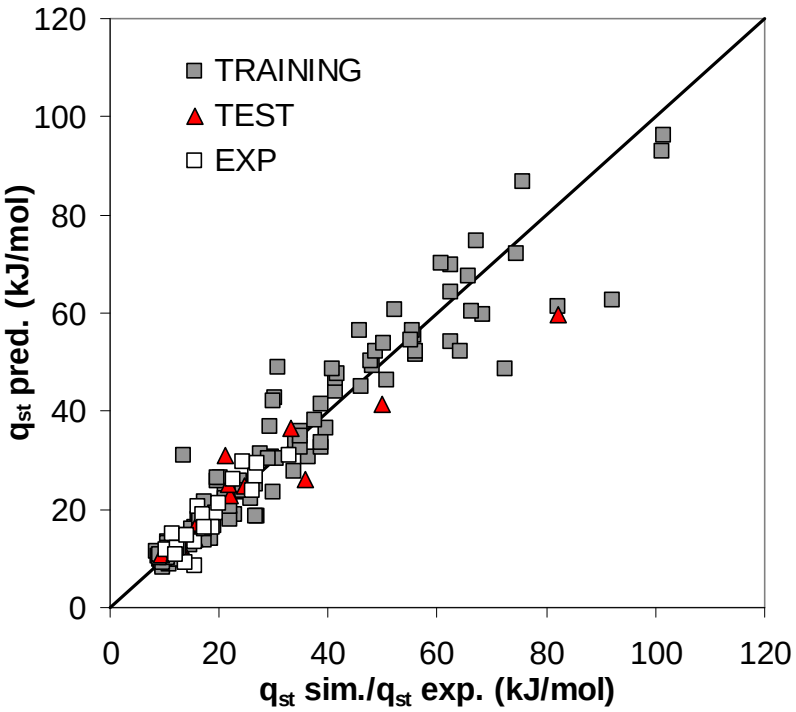


Figure 8. Comparison of the simulated and experimental isosteric heat of adsorption of different gases on different solids with prediction obtained using the ANN model. Both sets of training and test are shown for consistency

Database :

Solid	Topology	Set type	Gas	Qst sim (kJ/mol)	Qst pred (eq 11) (kJ/mol)
ZIF-93 (OH-CH3)	RHO	TEST	CO_2	35.910000	32.6350882
ZIF-71 (Cl2)	RHO	TEST	CH_4	14.990000	13.9096125
ZIF-96 (CN-NH2)	RHO	TEST	CH_4	15.420000	23.6340357
ZIF-69(Cl)	GME	TEST	CH_4	15.530000	19.2127878
ZIF-CN	SOD	TEST	CO	9.271000	7.48338923
ZIF-81(Br)	GME	TEST	CO	13.180000	14.5362226
ZIF-78(NO2)	GME	TEST	N_2	12.990000	13.2837375
ZIF-81(Br)	GME	TEST	Ar	11.720000	14.0958139
ZIF-68	GME	TEST	H_2S	33.100000	38.012356
ZIF-8	SOD	TEST	SO_2	21.880000	29.6858815
ZIF-CN	SOD	TEST	SO_2	49.790000	38.5862076
ZIF-90	SOD	TEST	CH_3CN	82.220000	49.8982612
ZIF-69(Cl)	GME	TEST	C_2H_6	24.630000	29.3501103
ZIF-78(NO2)	GME	TEST	C_2H_6	21.170000	30.6764002
ZIF-81(Br)	GME	TEST	C_2H_6	21.610000	29.1909523
ZIF-8	SOD	TRAINING	CO_2	15.580000	21.1196764
ZIF-90	SOD	TRAINING	CO_2	23.010000	23.9088807
ZIF-CN	SOD	TRAINING	CO_2	27.110000	23.6269241
ZIF-COOH	SOD	TRAINING	CO_2	19.870000	24.0131402
ZIF-NO2	SOD	TRAINING	CO_2	26.670000	23.7118793
ZIF-Cl	SOD	TRAINING	CO_2	15.140000	18.585827
ZIF-71 (Cl2)	RHO	TRAINING	CO_2	21.250000	24.2484722
ZIF-96 (CN-NH2)	RHO	TRAINING	CO_2	36.260000	33.9728955
ZIF-97 (CH2-OH)	RHO	TRAINING	CO_2	29.670000	26.3506334
ZIF-68	GME	TRAINING	CO_2	21.250000	26.2208784
ZIF-69(Cl)	GME	TRAINING	CO_2	26.750000	29.5516476
ZIF-78(NO2)	GME	TRAINING	CO_2	27.570000	30.8779375
ZIF-79(CH3)	GME	TRAINING	CO_2	23.090000	26.6093701
ZIF-81(Br)	GME	TRAINING	CO_2	24.850000	29.3924895
ZIF-8	SOD	TRAINING	CH_4	12.280000	10.7808166
ZIF-90	SOD	TRAINING	CH_4	11.760000	13.5700209
ZIF-CN	SOD	TRAINING	CH_4	10.770000	13.2880643
ZIF-COOH	SOD	TRAINING	CH_4	13.870000	13.6742804
ZIF-NO2	SOD	TRAINING	CH_4	11.430000	13.3730195
ZIF-Cl	SOD	TRAINING	CH_4	11.450000	8.24696725
ZIF-93 (OH-CH3)	RHO	TRAINING	CH_4	18.510000	22.2962284
ZIF-97 (CH2-OH)	RHO	TRAINING	CH_4	21.890000	16.0117737
ZIF-68	GME	TRAINING	CH_4	14.180000	15.8820186
ZIF-78(NO2)	GME	TRAINING	CH_4	16.030000	20.5390777
ZIF-79(CH3)	GME	TRAINING	CH_4	16.030000	16.2705103
ZIF-81(Br)	GME	TRAINING	CH_4	15.630000	19.0536297

ZIF-8	SOD	TRAINING	CO	9.772000	3.83375125
ZIF-90	SOD	TRAINING	CO	9.622000	7.14574598
ZIF-COOH	SOD	TRAINING	CO	10.970000	6.91054738
ZIF-NO2	SOD	TRAINING	CO	9.200000	7.0169053
ZIF-Cl	SOD	TRAINING	CO	9.173000	1.12281845
ZIF-71 (Cl2)	RHO	TRAINING	CO	10.850000	6.75572862
ZIF-93 (OH-CH3)	RHO	TRAINING	CO	12.310000	15.991355
ZIF-96 (CN-NH2)	RHO	TRAINING	CO	11.140000	17.2569638
ZIF-97 (CH2-OH)	RHO	TRAINING	CO	14.570000	9.63249058
ZIF-68	GME	TRAINING	CO	11.030000	10.523417
ZIF-69(Cl)	GME	TRAINING	CO	13.080000	14.7684444
ZIF-78(NO2)	GME	TRAINING	CO	14.910000	16.6302147
ZIF-79(CH3)	GME	TRAINING	CO	12.660000	12.1068856
ZIF-8	SOD	TRAINING	N_2	9.580000	3.52547639
ZIF-90	SOD	TRAINING	N_2	9.945000	6.31468074
ZIF-CN	SOD	TRAINING	N_2	10.240000	6.03272414
ZIF-COOH	SOD	TRAINING	N_2	10.800000	6.41894017
ZIF-NO2	SOD	TRAINING	N_2	9.507000	6.11767927
ZIF-Cl	SOD	TRAINING	N_2	9.043000	0.99162705
ZIF-71 (Cl2)	RHO	TRAINING	N_2	10.610000	6.65427225
ZIF-93 (OH-CH3)	RHO	TRAINING	N_2	11.860000	15.0408882
ZIF-96 (CN-NH2)	RHO	TRAINING	N_2	11.000000	16.3786955
ZIF-97 (CH2-OH)	RHO	TRAINING	N_2	14.800000	8.75643345
ZIF-68	GME	TRAINING	N_2	11.110000	8.62667844
ZIF-69(Cl)	GME	TRAINING	N_2	12.480000	11.9574476
ZIF-79(CH3)	GME	TRAINING	N_2	12.420000	9.01517013
ZIF-81(Br)	GME	TRAINING	N_2	12.510000	11.7982895
ZIF-8	SOD	TRAINING	O_2	9.508000	6.52838108
ZIF-90	SOD	TRAINING	O_2	9.484000	9.31758544
ZIF-CN	SOD	TRAINING	O_2	9.337000	9.03562883
ZIF-COOH	SOD	TRAINING	O_2	10.530000	9.42184487
ZIF-NO2	SOD	TRAINING	O_2	9.326000	9.12058396
ZIF-Cl	SOD	TRAINING	O_2	9.058000	3.99453175
ZIF-71 (Cl2)	RHO	TRAINING	O_2	8.419000	9.65717694
ZIF-93 (OH-CH3)	RHO	TRAINING	O_2	11.790000	18.0437929
ZIF-96 (CN-NH2)	RHO	TRAINING	O_2	10.650000	19.3816002
ZIF-97 (CH2-OH)	RHO	TRAINING	O_2	14.090000	11.7593382
ZIF-68	GME	TRAINING	O_2	10.780000	11.6295831
ZIF-69(Cl)	GME	TRAINING	O_2	12.070000	14.9603523
ZIF-78(NO2)	GME	TRAINING	O_2	12.510000	16.2866422
ZIF-79(CH3)	GME	TRAINING	O_2	12.200000	12.0180748
ZIF-81(Br)	GME	TRAINING	O_2	12.110000	14.8011942
ZIF-8	SOD	TRAINING	Ar	9.569000	5.82300079
ZIF-90	SOD	TRAINING	Ar	9.164000	8.61220514
ZIF-CN	SOD	TRAINING	Ar	8.685000	8.33024853
ZIF-COOH	SOD	TRAINING	Ar	10.450000	8.71646457

ZIF-NO2	SOD	TRAINING	Ar	9.085000	8.41520366
ZIF-Cl	SOD	TRAINING	Ar	9.002000	3.28915145
ZIF-71 (Cl2)	RHO	TRAINING	Ar	10.450000	8.95179664
ZIF-93 (OH-CH3)	RHO	TRAINING	Ar	11.800000	17.3384126
ZIF-96 (CN-NH2)	RHO	TRAINING	Ar	10.520000	18.6762199
ZIF-97 (CH2-OH)	RHO	TRAINING	Ar	13.980000	11.0539579
ZIF-68	GME	TRAINING	Ar	10.560000	10.9242028
ZIF-69(Cl)	GME	TRAINING	Ar	11.650000	14.254972
ZIF-78(NO2)	GME	TRAINING	Ar	12.090000	15.5812619
ZIF-79(CH3)	GME	TRAINING	Ar	12.050000	11.3126945
ZIF-8	SOD	TRAINING	H_2S	19.190000	25.1144824
ZIF-90	SOD	TRAINING	H_2S	19.590000	30.4705318
ZIF-CN	SOD	TRAINING	H_2S	34.160000	33.2303346
ZIF-COOH	SOD	TRAINING	H_2S	20.120000	28.907784
ZIF-NO2	SOD	TRAINING	H_2S	19.590000	30.6080854
ZIF-Cl	SOD	TRAINING	H_2S	16.970000	21.7105976
ZIF-71 (Cl2)	RHO	TRAINING	H_2S	33.860000	27.2273076
ZIF-93 (OH-CH3)	RHO	TRAINING	H_2S	39.500000	39.783172
ZIF-96 (CN-NH2)	RHO	TRAINING	H_2S	30.350000	40.7662356
ZIF-97 (CH2-OH)	RHO	TRAINING	H_2S	38.660000	33.1324371
ZIF-69(Cl)	GME	TRAINING	H_2S	48.210000	45.8307777
ZIF-78(NO2)	GME	TRAINING	H_2S	62.360000	49.7912082
ZIF-79(CH3)	GME	TRAINING	H_2S	34.890000	44.2728661
ZIF-81(Br)	GME	TRAINING	H_2S	47.690000	45.3159146
ZIF-90	SOD	TRAINING	SO_2	38.750000	35.4005201
ZIF-COOH	SOD	TRAINING	SO_2	35.020000	33.6051219
ZIF-NO2	SOD	TRAINING	SO_2	38.750000	35.5851806
ZIF-Cl	SOD	TRAINING	SO_2	19.000000	26.1602879
ZIF-71 (Cl2)	RHO	TRAINING	SO_2	34.850000	31.6565207
ZIF-93 (OH-CH3)	RHO	TRAINING	SO_2	41.380000	44.7948763
ZIF-96 (CN-NH2)	RHO	TRAINING	SO_2	55.970000	45.7289103
ZIF-97 (CH2-OH)	RHO	TRAINING	SO_2	41.350000	38.0941504
ZIF-68	GME	TRAINING	SO_2	29.790000	43.6720204
ZIF-69(Cl)	GME	TRAINING	SO_2	55.830000	52.1268659
ZIF-78(NO2)	GME	TRAINING	SO_2	68.370000	56.4430015
ZIF-79(CH3)	GME	TRAINING	SO_2	37.420000	50.7516137
ZIF-81(Br)	GME	TRAINING	SO_2	55.530000	51.5543208
ZIF-8	SOD	TRAINING	CH_3CN	29.230000	39.7084678
ZIF-CN	SOD	TRAINING	CH_3CN	62.610000	58.383955
ZIF-COOH	SOD	TRAINING	CH_3CN	45.710000	45.1966562
ZIF-NO2	SOD	TRAINING	CH_3CN	82.220000	50.6626249
ZIF-Cl	SOD	TRAINING	CH_3CN	24.490000	34.6673781
ZIF-71 (Cl2)	RHO	TRAINING	CH_3CN	64.200000	39.9099068
ZIF-93 (OH-CH3)	RHO	TRAINING	CH_3CN	66.230000	60.3164714
ZIF-96 (CN-NH2)	RHO	TRAINING	CH_3CN	65.750000	60.625618

ZIF-97 (CH2-OH)	RHO	TRAINING	CH_3CN	52.230000	52.974515
ZIF-68	GME	TRAINING	CH_3CN	55.960000	67.2844648
ZIF-69(Cl)	GME	TRAINING	CH_3CN	101.400000	83.5648227
ZIF-78(NO2)	GME	TRAINING	CH_3CN	75.770000	92.4570565
ZIF-79(CH3)	GME	TRAINING	CH_3CN	67.170000	84.5833698
ZIF-81(Br)	GME	TRAINING	CH_3CN	101.200000	82.3673903
ZIF-8	SOD	TRAINING	C_2H_6	16.310000	20.9181392
ZIF-90	SOD	TRAINING	C_2H_6	16.540000	23.7073435
ZIF-CN	SOD	TRAINING	C_2H_6	17.080000	23.4253869
ZIF-COOH	SOD	TRAINING	C_2H_6	21.930000	23.8116029
ZIF-NO2	SOD	TRAINING	C_2H_6	16.540000	23.510342
ZIF-Cl	SOD	TRAINING	C_2H_6	17.260000	18.3842898
ZIF-71 (Cl2)	RHO	TRAINING	C_2H_6	29.880000	24.046935
ZIF-93 (OH-CH3)	RHO	TRAINING	C_2H_6	23.770000	32.433551
ZIF-96 (CN-NH2)	RHO	TRAINING	C_2H_6	30.630000	33.7713583
ZIF-97 (CH2-OH)	RHO	TRAINING	C_2H_6	29.070000	26.1490962
ZIF-68	GME	TRAINING	C_2H_6	25.770000	26.0193412
ZIF-79(CH3)	GME	TRAINING	C_2H_6	24.440000	26.4078329
ZIF-8	SOD	TRAINING	H_2O	13.600000	36.6955137
ZIF-90	SOD	TRAINING	H_2O	41.700000	43.3917061
ZIF-CN	SOD	TRAINING	H_2O	48.600000	47.7396842
ZIF-COOH	SOD	TRAINING	H_2O	50.850000	40.9589229
ZIF-NO2	SOD	TRAINING	H_2O	72.550000	43.7042282
ZIF-Cl	SOD	TRAINING	H_2O	17.240000	32.8379607
ZIF-71 (Cl2)	RHO	TRAINING	H_2O	40.840000	38.2784344
ZIF-93 (OH-CH3)	RHO	TRAINING	H_2O	55.290000	53.0100605
ZIF-96 (CN-NH2)	RHO	TRAINING	H_2O	92.100000	53.8085419
ZIF-97 (CH2-OH)	RHO	TRAINING	H_2O	50.140000	46.1689753
ZIF-68	GME	TRAINING	H_2O	30.670000	53.6695756
ZIF-69(Cl)	GME	TRAINING	H_2O	62.570000	63.8356511
ZIF-78(NO2)	GME	TRAINING	H_2O	74.590000	69.1516064
ZIF-79(CH3)	GME	TRAINING	H_2O	46.000000	62.9891497
ZIF-81(Br)	GME	TRAINING	H_2O	60.840000	63.1285149

Bibliography

1 K. S. Park.; Z. Ni, A. P. Côté, J. Y. Choi, R. Huang, F.J. Uribe-Romo, H. K. Chae, M. O'Keefe, O. M. Yaghi, Proc. Natl. Acad. Sci. U.S.A. 2006, 103, 10186-10191

2 Banerjee, R.; Phan, A.; Wang, B.; Knobler, C.; Furukawa, H.; O'Keeffe, M.; Yaghi, O. M. Science 2008, 319, 939–943

3 Liu, Y.; Kravtsov, V. Ch.; Larsen, R.; Eddaoudi, M. Chem. Commun. 2006, 1488–1490.

4 Huang, X.-C.; Lin, Y.-Y.; Zhang, J.-P.; Chen, X. M. Angew. Chem., Int. Ed. 2006, 45, 1557–1559.

5 Morris W., Leung B., Furukawa H., Yaghi O. K., He N., Hayashi H., Houndonougbo Y., Asta M., Laird B. B., and Yaghi O. M., J. Am. Chem. Soc., 2010, 132, 11006

6 R.B. Rankin, J. Liu,A.D. Kulkarni, J.K. Johnson, J. Phys. Chem. C, Vol. 113, No. 39, 2009, 16906-16914

7 R. Babarao, S. Dai, D-e. Jiang, J. Phys. Chem. C, 2011, 115, 8126–8135

8 R. Banerjee, H. Furukawa, D. Britt, M. Knobler O'Keeffe, O. Yaghi, J. Am. Chem. Soc. 2009, 131, 3875-3877

9 Liu Y., Liu H. L., Hu Y. and Jiang J. W., J. Phys. Chem. B, 2010, 114, 2820.

10 J. Liu, S. Keskin, David S. Sholl, J. Karl Johnson, J. Phys. Chem. C 2011, 115, 12560–12566

11 Aguado, S.; Bergeret, G.; Titus, M. P.; Moizan, V.; Nieto-Draghi, C.; Bats, N.; Farrusseng, D. New J. of Chem. 2011, 35, 546.

12 Gücüyener, C.; van den Bergh, J.; Gascon, J.; and Kapteijn, F. J. Am. Chem. Soc. 2010, 132, 17704.

13 Aguado, S.; Nicolas, C.H.; Moizan-Baslé, V.; Nieto, C.; Amrouche,H.; Bats, N.; Audebrand, N.; Farrusseng, D. New J. Chem. 2011, 35, 41.

14 Cambridge Crystallographic Data Center, http://www.ccdc.cam.ac.uk/

15 H. Amrouche, S. Aguado, J. Pérez-Pellitero, C. Chizallet, F. Siperstein, D. Farrusseng, N. Bats, Carlos Nieto-Draghi, J. Phys. Chem. C, 2011, 115 (33), 16425–16432

16 T. Düren, F. Millange, G. Férey, K.S. Walton and R.Q. Snurr J. Phys. Chem. C, 2007, Vol. 111, No.42

17 Materials Studio, Release 5.0, Accelrys Software Inc., San Diego, 2009.

18 Gharagheizi, F. Energy and Fuels 2008, 22, 3037.

19 Patel, S. J.; Ng, D.; Mannan, M. S. Ind. Eng. Chem. Res. 2009, 48, 7378.

20 Jaguar version 7.0, Schrödinger, LLC., New York, NY, 2007

21 a) Kresse, G.; Hafner, J. Phys. Rev. B 1994, 49, 14251-14269; b) Kresse, G.; Furthmüller, J. Phys. Rev. B 1996, 54, 11169:11161-11118

22 D. Frenkel, B. Smit in Understanding Molecular Simulation: From Algorithms to Applications, Academic Press, San Francisco, 2001

23 Vrabec, J.; Stoll, J.; Hasse, H. . J. Phys. Chem. B 2001, 105, 12126.

24 D. Moller, J. Oprzynski, A. Muller, J. Fischer, Molecular Physics, 1992, 75, 363

25 Ungerer, P.; Beauvais, C.; Delhommelle, J.; Boutin, A.; Rousseau, B.; Fuchs, A. J. Chem. Phys. 2000, 112, 5499.

26 J. Delhommelle, PhD Thesis, Universite de Paris XI, Orsay, France, 2000

27 Y. Boutard, Ph. Ungerer, J.M. Teuler, M.G. Ahunbay, S.F. Sabater, J. Perez-Pellitero, A.D. Mackie, E. Bourasseau, Fluid Phase Equilib. 236 (2005) 25–41.

28 G. Harris and K. H. Yung, J. Phys. Chem., 1995, 99, 12021.

29 J. Pérez-Pellitero, H. Amrouche, F. Siperstein, G. Pirngruber, C. Nieto-Draghi, G. Chaplais, A. Simon-Masseron, D. Bazer-Bachi, D. Peralta and N. Bats, Chem.–Eur. J., 2010, 16, 1560–1571

30 E. El Ahmar, B. Creton, A. Valtz, C. Coquelet, V. Lachet, D. Richon, P. Ungerer, Fluid Phase Equilibria 304 , 2011, 21–34
31 Kristóf, T.; Liszi, J. J. Phys. Chem. B 1997, 101, 5480.
32 Edwards, D. M. F.; Madden, P. A.; McDonald, I. R. Mol. Phys. 1984, 51, 1141.
33 Jorgensen, W. L.; Madura, J. D.; Impey, R. W.; Klein, M. L. J. Chem. Phys. 1983, 79, 926.
34 Ungerer, P.; Nieto-Draghi, C.; Lachet, V.; Wender, A.; di Lella, A.; Boutin, A.; Rousseau, B and Fuchs, A. Mol. Sim., 2007, 33, 287-304
35 A. D. Mackie, B. Tavitian, A. Boutin, A. H. Fuchs, Mol. Phys., 1992, 75, 363-378
36 M. P. Allen, D. J. Tildesley in Computer Simulation of Liquids, Oxford University Press, New York, 1989
37 T. Vuong, P. A. Monson, Langmuir, 1996, 12, 5425-5432