PCCP

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2013, **15**, 4436

Predicting adsorption on metals: simple yet effective descriptors for surface catalysis†

Erik-Jan Ras,*[ab] Manuel J. Louwerse,[b] Marjo C. Mittelmeijer-Hazeleger[b] and Gadi Rothenberg*[b]

Received 23rd August 2012,
Accepted 18th January 2013

DOI: 10.1039/c3cp42965b

www.rsc.org/pccp

We present a simple and efficient model for predicting the adsorption of molecules on metal surfaces. This heuristic model uses six descriptors for each metal (number of d-electrons, surface energy, first ionization potential and atomic radius, volume and mass) and three for each adsorptive (HOMO–LUMO energy gap, molecular volume and mass). Strikingly, despite its simplicity and low computational cost, this model predicts well the chemisorption of a range of adsorptives (H_2, HO•, N_2, CO, NO, O_2, H_2O, CO_2, NH_3 and CH_4) on a range of metals (Fe, Co, Ni, Cu, Mo, Ru, Rh, Pd, Ag, W, Ir, Pt and Au) as calculated with DFT and taken from the literature. Using only a third of the data for fitting, the rest of the data were predicted with Q^2 = 0.91–0.95 and RMSEP = 0.94–1.16 eV. Furthermore, we measured experimental adsorption data for CO, CO_2, CH_4, H_2, N_2 and O_2 on Ni, Pt and Rh supported on TiO_2. Using the same descriptors, we then constructed a model for this experimental data set. Once again, the model explained the data well, with R^2 = 0.95 and Q^2 = 0.86, respectively.

Introduction

The majority of bulk chemicals and fuels derived from crude oil are produced using heterogeneous catalysis. Moreover, an increasing number of processes for biomass conversion rely on solid catalysts, holding the key to a sustainable energy and chemicals cycle.[1] The market size for catalysts (which are by and large rather inexpensive materials) was €10 Bn in 2009. Thus, heterogeneous catalysis is big business, which also attracts much research attention in both industry and academia.

Nevertheless, despite its economic importance, heterogeneous catalysis remains an empirical science. The complexity of the surface reactions and the large number of independent parameters make the prediction of catalyst performance a formidable task. Numerous man-years of research notwithstanding, we are still closer to Mittasch's "test everything" philosophy[2] than to predicting structure–activity relationships *a priori* for a new target reaction and catalyst. Moreover, unlike homogeneous catalysis where predictive modelling is proven and accepted,[3–5] studies applying descriptors to solid catalysts are scant.[6] This reflects the difficulty of finding a relationship between bulk/surface structure and performance. It can be done by combining detailed mechanistic studies, *in situ* spectroscopy,[7–12] and DFT modelling,[13–18] but unfortunately this is too costly for large sets of catalysts.[13–15,19,20] Therefore, catalyst discovery needs simpler models.[21]

To meet this challenge, we present here a "quick and dirty" approach for describing and modelling adsorption – a fundamental step in surface-catalysed reactions. Adsorption is a crucial step for catalysis, not least because of the relation between the heat of adsorption and the activation energy, following the so-called Brønsted–Evans–Polanyi relationship.[22,23] This is because transition states often have adsorption properties comparable to those of the dissociated[24,25] or non-dissociated[26] reactants. Furthermore, a Sabatier-type trade-off exists between adsorption of reactants and desorption of products, such that only those with optimal adsorption energies will give maximal catalytic activity.[24,27]

The importance of adsorption is easily illustrated using the Fischer–Tropsch (F–T) process as an example: classically, F–T reactions are carried out using supported Fe and Co catalysts. This is because these metals have suitable adsorption capabilities for both CO and H_2. Therefore, one can envisage that metals with similar adsorption capabilities to those of Fe and Co

[a] *Avantium Chemicals B.V., Zekeringstraat 29, 1014BV Amsterdam, The Netherlands. E-mail: erikjan.ras@avantium.com; Web: www.avantium.com; Fax: +31 (0)20 586 8085; Tel: +31 (0)20 586 8080*

[b] *Van't Hoff Institute of Molecular Sciences, University of Amsterdam, Science Park 904, 1098XH Amsterdam, The Netherlands. E-mail: g.rothenberg@uva.nl; Web: www.science.uva.nl/hims/hcsc; Fax: +31 (0)20 525 5604; Tel: +31 (0)20 525 6963*

† Electronic supplementary information (ESI) available: Experimental adsorption isotherms, modelling details including details of the separate models for H_2 and HO• and final model coefficients and equations. See DOI: 10.1039/c3cp42965b

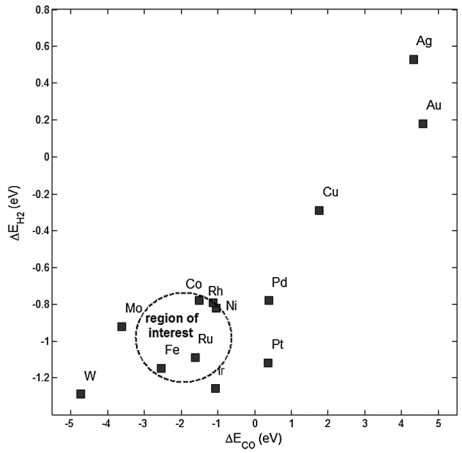
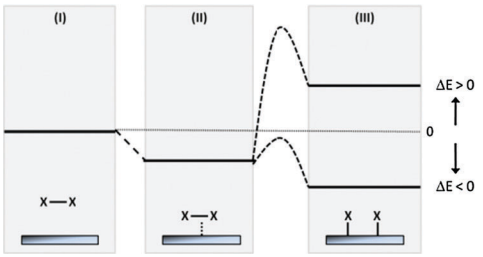


Fig. 1 Heats of adsorption for CO (*x*-axis) and H_2 (*y*-axis) on a number of pure metals. The dotted circle indicates the region of interest for metals with expected Fischer–Tropsch activity.



Scheme 1 Simplified diagram showing the energy changes when a molecule from the gas phase (I) first adsorbs (II) and then dissociatively chemisorbs (III) on a surface. The data discussed here are for dissociative chemisorption, and can thus have either positive or negative values.

should also be good F–T catalysts. Fig. 1 shows the adsorption energies calculated by Bligaard *et al.*,[15] circling a region of heats of adsorption around these two metals. We see in the circle also Rh, Ru and Ni. Indeed, all three are known as F–T catalysts, though they are not used for commercial reasons (Ru in particular is a very effective catalyst[28]). Just outside the circle we find Mo and Ir, which are also capable of catalyzing F–T reactions.[29,30]

In the following sections, we apply a set of simple descriptors that capture the essence of molecule–surface interactions. For simplicity, we use here the term adsorption referring to both chemisorption and physisorption. Using these descriptors, we model a set of 39 metal–adsorptive combinations, and then test the model on 91 new combinations. This approach is both simple and powerful. With only a standard desktop PC, we obtain very good predictions for adsorption energies (Q^2 = 0.91–0.95 and RMSEP = 0.94–1.16 eV for 70–87 observations in the validation set). This heuristic model for adsorption has value as such, but more importantly, it is a step towards predicting catalytic activity. Once new potentially interesting systems are identified using these rough predictions, one can use in-depth research methods (experimental and/or computational) to investigate these candidates further. Exploring further the validity of our method, we use the key descriptors also for explaining the experimental adsorption data for Ni, Pt and Rh supported on titania.

Methods

Data set selection

We used the data set published by Bligaard *et al.*[15] as a base case. This data set describes the dissociative chemisorption (Scheme 1) of ten adsorptives (H_2, HO•, N_2, CO, NO, O_2, H_2O, CO_2, NH_3, CH_4) on 13 metals (Fe, Co, Ni, Cu, Mo, Ru, Rh, Pd, Ag, W, Ir, Pt, Au) at an isosteric coverage of 1/6 monolayer. We then modelled three subsets of these data: the entire set, the data set excluding H_2 and the data set excluding both H_2 and HO•. For model validation purposes each of these subsets was partitioned into a training set and a validation set by excluding 1/3, 1/2 and 2/3 of the data (the excluded part then used as an ''unseen'' validation set). This partitioning was done by combining d-optimal and space filling design methods in Matlab (ESI† contains the partitioning of the data set for all models herein).

As descriptors we use steric and electronic parameters to characterize both the metals and the adsorptives (see Table 1). The metals are described by the number of valence electrons, surface energy, first ionization potential, atomic radius, volume and mass. Similarly, the adsorptives are described by molecular volume, the HOMO–LUMO gap and molecular mass. For the metals in particular, we rely on tabulated values for the descriptors for the sake of accessibility. This makes them easier to use than, for example, the d-band center. Note that the adsorptives are described using simple gas-phase properties, which can be calculated in seconds using standard software. Finally, no metal–adsorptive interaction is assumed in the definition of these parameters. A table with the values of the descriptors for all metals and adsorptives is included in ESI.†

Variable selection and modelling methods

To identify the best model for correlating the descriptors with heats of adsorption, we combined genetic algorithm variable selection and PLS regression. This routine is implemented in Matlab, in combination with the PLS toolbox.[32] We studied three variations in our modelling approach (see Fig. 2): A – the initial data set selection, B – the initial variable space selection

Table 1 Overview of the descriptors used for model construction

Variable	Description	Source
x_1	Number of valence electrons of metal	[a]
x_2	Surface energy of metal (J m^{-2})	[b]
x_3	First ionization potential of metal (V)	[a]
x_4	Atomic radius of metal (Å)	[a]
x_5	Atomic volume of metal (Å^3)	Calculated from x_4
x_6	Atomic mass of metal (g mol^{-1})	[a]
x_7	Molecular volume of adsorptive (Å^3)	[c]
x_8	Adsorptive HOMO–LUMO gap (eV)	[c]
x_9	Molecular mass of adsorptive (g mol^{-1})	[c]

[a] From tabulated data. [b] Experimentally determined surface energies from ref. 31. [c] Calculated using HyperChem version 7.5 using an STO-3G basis set.

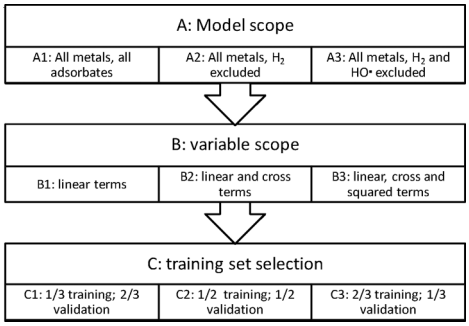


Fig. 2 Summarized modelling strategy for the 27 scenarios studied. Each model follows a three-step course. First, the initial data set is selected (A). Then, the initial variable space is selected (B). Finally, the size of the training set is chosen (C). The variations are tested combinatorially, resulting in 3 × 3 × 3 = 27 models.

and C – the required training set size. For each of these, we explored three permutations, giving a total of 3 × 3 × 3 = 27 models (from A1B1C1 to A3B3C3).

This evolutionary approach evaluates models in ''generations''. In each generation, subsets of variables are tested for their impact on the model's performance. If including a given variable consistently improves performance, the next generation will ''inherit'' it. Similarly, if a variable has a negative effect, it will be eliminated. As an example, Fig. 3 shows the evolution of scenario A1B2C1 (full data set, linear and cross terms included, 1/3 training set size).

The modelling method we used is Partial Least Squares (PLS) regression.[33] Here the variable matrix x is decomposed into a lower dimensional scores matrix t, where each score is a linear combination of the original variables. These principal components (also called latent variables) are then correlated with the response of interest y. For easy interpretation and application, the final PLS model can be transformed into a polynomial relating the original variables matrix x to the response matrix y.

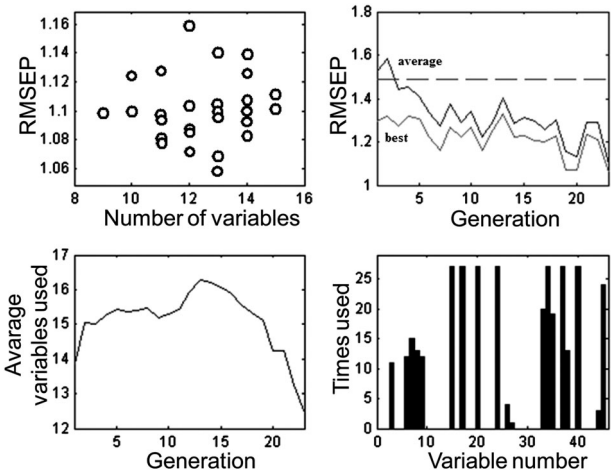


Fig. 3 Example showing the ''evolution'' of model performance during the variable selection procedure for scenario A1B2C1. The top left graph shows the model performance *vs.* the number of variables. The top right graph shows the average (blue) and the best (green) model performance for each generation. The bottom left graph shows the average number of variables used in the models for each generation. The bottom right histogram represents the number of models using each variable for the final generation (generation #24 in this case). The performance indicators for the best model are given in Table 2. Typically, each generation consists of 20–30 models.

Adsorption measurements

The catalysts used to perform adsorption measurements were prepared using wet impregnation. We synthesized three monometallic catalysts (Ni, Pt, Rh) on titania (commercial sample with a surface area of 41 $m^2\ g^{-1}$, pre-dried at 105 °C). Nitrate salts were used as precursors. All catalysts had a metal loading of 1 wt% relative to the support and were prepared by wet impregnation in water (18 h impregnation at ambient temperature). After decanting the excess liquid, the catalysts were dried in air at 105 °C for 18 h, followed by calcination in air at 320 °C for 4 h.

Example: Pt on titania. A stock solution of tetraaminoplatinum(II) nitrate was prepared by dissolving 3.1 grams of $(NH_3)_4Pt(NO_3)_2$ in 25 mL of demineralized water. To five grams of pre-dried support (overnight at 105 °C in air), 800 μL of Pt-stock solution was added, after which the total volume of the system was adjusted to 10.0 mL. The resulting slurry was homogenized for 24 h under ambient conditions. Then, the homogenized sample was dried at 105 °C for 18 h and then calcined in air at 320 °C for 4 h (with ramp rates of 5 °C min^{-1} for both the drying and the calcination).

Equipment

Adsorption isotherms and differential heats of adsorption were assessed directly by a manometric set-up combined with a micro-calorimeter (Calvet type of Setaram C80).

Gas or vapor can be fed into the system by a piston, which can introduce a full, 1/2 or 1/4 stroke. The introduction pressure cannot be higher than 100 kPa. Two pressure transducers with different sensitivities allow an accurate measurement of pressure, of the gas phase in contact with the sample, from 0 kPa to 10 kPa and from 0 kPa to 1000 kPa. The system has two independent data acquisition systems, one for the manometric (isotherm) data and the other for the calorimetric data. A detailed description of this system is published elsewhere.[34] In this work we focused on the isotherm data. This decision was based on greater accuracy for initial adsorption trends (prior to monolayer coverage) compared to the calorimetric data.

Measurement method

For the measurements on the support about 4 grams were used, while for the samples we used about 3 grams. Prior to each measurement the samples were evacuated, put in a H_2 flow and the sample holder was placed in a furnace. Under the hydrogen flow the temperature was raised with 10 °C min^{-1} to 400 °C. After one hour at this temperature the sample holder was evacuated till a better vacuum than 2 × 10^{-6} mbar. After cooling down the sample holder was placed in the calorimeter that was set on 40 °C. When the calorimeter signal was stable the measurement was started. Equilibrium was reached when the change in pressure was below 0.1 kPa over a 40 minute period.

Results and discussion

Modelling the DFT data set

Adsorption is essentially an electronic interaction process. Charge is donated from the surface into the adsorbate LUMO and from the adsorbate HOMO to the surface. The orbital energies and the number of available electrons are thus the key parameters, and the size of the adsorptive also matters. We therefore described each metal by six parameters: the number of d-electrons, the surface energy, the first ionization potential, and the atomic radius, volume, and mass (see Table 1). Similarly, each adsorptive was described by three parameters: the HOMO–LUMO energy gap, the molecular volume, and the mass. These parameters describe both the electronic and the steric characteristics of the adsorptive and the metal surface. Note that adsorptives and metals are treated independently.

Table 2 gives the performance indicators for the final models in each of the 27 scenarios. In each case, we evaluate the performance by comparing the correlation (R^2 and Q^2) between the calculated and observed heats of adsorption, as well as the root mean squared prediction errors (RMSEE and RMSEP). In this evaluation, the performance measures for the validation set have more weight than the measures derived from the training set alone. In general, low prediction errors are considered more important than high correlations, and models with less variables are preferred (because they are simpler).[35] We then selected the best models for each of the three data sets. In all three cases the best model is obtained using linear and cross terms as the initial parameter space, with 1/3 of the available observations used as a training set (models 4, 13 and 22, italicised in Table 2).

Fig. 4 compares the adsorption energies obtained *via* the DFT calculations of Bligaard *et al.*[15] with those predicted by our best descriptor-derived models. The similarity is remarkable. Using only 1/3 of the available data points, we can predict with good accuracy the adsorption behaviour of different molecules on different metals. These models thus open an effective and accessible route for quantifying substrate–surface interactions. Parity plots and model coefficients are given in Fig. 5.

Examining the outliers in Fig. 4 and 5 shows that they correspond to non-linear molecules (CH_4, NH_3 and H_2O). These molecules have a complex adsorption behaviour, and describing their isosteric heat of adsorption with just one number is insufficient.[36] Here, methane is the most difficult adsorptive to explain (a detailed description of the residuals of prediction split per metal and per adsorptive is included in the ESI†). We also see that H_2 adsorption is not explained well by the model. One reason for this is that H_2 adsorption spans only a narrow range of heats of adsorption, of the same order of magnitude as the estimation error of the model. However, a separate model for the heat of adsorption of H_2 using the same descriptors is easily obtained (see ESI†).

Interestingly, excluding H_2 from the training set improved the prediction quality for the other adsorptives. This indicates that within the set of adsorptives considered here, H_2 is an outlier (also confirmed by the diagnostic statistics obtained

Table 2 Description and performance parameters for all 27 modelling scenarios

Entry	Scenario[a]	Latent variables[b] #	Included variables[c] #	Training set size #	R^2 —	RMSEE eV	Validation set size #	Q^2 —	RMSEP eV
1	A1B1C1	1	4	43	0.91	1.52	87	0.86	1.41
2	A1B1C2	1	4	65	0.86	1.73	65	0.92	1.29
3	A1B1C3	1	5	87	0.85	1.62	43	0.86	1.74
4	A1B2C1	3	10	43	0.95	1.06	87	0.91	1.16
5	A1B2C2	3	12	65	0.94	1.15	65	0.89	1.11
6	A1B2C3	3	11	87	0.94	1.17	43	0.93	1.26
7	A1B3C1	2	12	43	0.94	1.10	87	0.90	1.27
8	A1B3C2	2	14	65	0.93	1.24	65	0.89	1.14
9	A1B3C3	3	11	87	0.93	1.17	43	0.92	1.22
10	A2B1C1	2	4	39	0.90	1.58	78	0.91	1.41
11	A2B1C2	2	5	59	0.92	1.48	58	0.91	1.14
12	A2B1C3	3	8	78	0.93	1.36	39	0.91	1.21
13	A2B2C1	3	13	39	0.96	1.01	78	0.93	0.99
14	A2B2C2	3	18	59	0.96	1.04	58	0.92	1.03
15	A2B2C3	3	11	78	0.96	1.03	39	0.90	1.08
16	A2B3C1	3	10	39	0.96	0.96	78	0.93	1.04
17	A2B3C2	3	14	59	0.96	1.02	58	0.92	1.02
18	A2B3C3	3	12	78	0.95	1.08	39	0.89	1.14
19	A3B1C1	2	3	34	0.94	1.31	70	0.93	1.16
20	A3B1C2	2	6	52	0.93	1.23	52	0.95	0.89
21	A3B1C3	3	4	70	0.93	1.17	34	0.94	1.23
22	A3B2C1	3	13	34	0.97	0.87	70	0.95	0.94
23	A3B2C2	2	14	52	0.97	0.88	52	0.94	0.93
24	A3B2C3	3	11	70	0.96	0.89	34	0.95	1.03
25	A3B3C1	3	18	34	0.96	0.90	70	0.95	0.96
26	A3B3C2	3	16	52	0.96	0.90	52	0.95	0.88
27	A3B3C3	3	13	70	0.96	0.93	34	0.96	0.95

[a] Scenario corresponding to data set, variable and training set selection (see Fig. 2). [b] Number of latent variables (principal components) in the final PLS model. [c] Number of original variables used to construct the final PLS model.

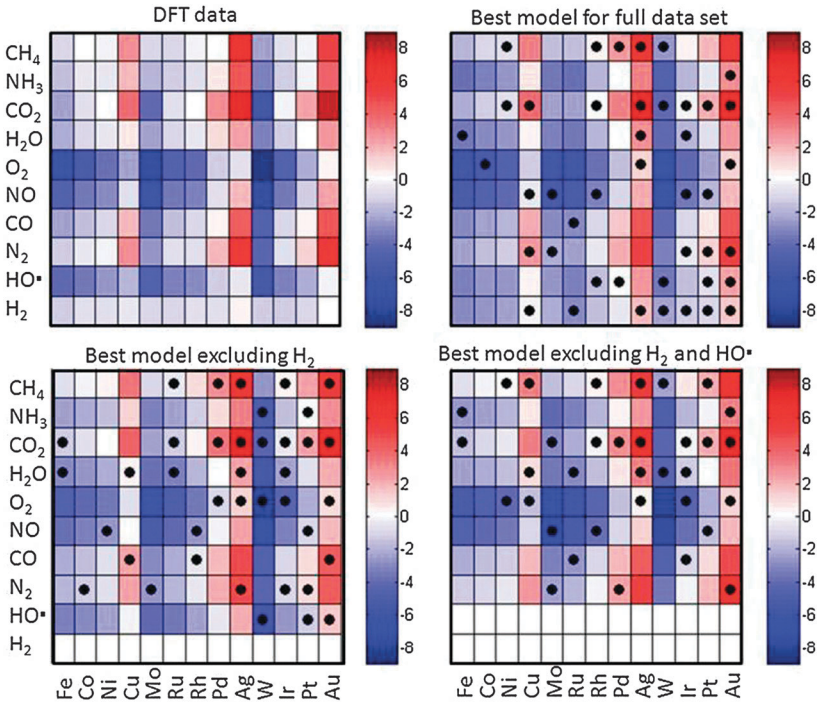


Fig. 4 Color matrices comparing the original DFT data for dissociative chemisorption calculated by Bligaard *et al.* (top left) with the predicted results from the best models for each data set from Table 2 (the black dots denote those observations used for training each model). The plots correspond to model 4 (upper right), model 13 (bottom left) and model 22 (bottom right). All values are in eV. Since the dissociative chemisorptions involve both making and breaking of bonds, both positive and negative values are possible (see Scheme 1).

from the PLS model evaluation). A similar reasoning can be applied to the HO• radical, though to a lesser extent. That said, the same descriptors can still be used to produce a model just for H_2 or HO•. These models are included in the ESI.†

Traditionally, both physisorption and chemisorption are thought of in terms of bond formation – non-covalent or covalent – between the adsorbate and the surface. The fact that our model fits well shows that it is also strongly related to the electronic and steric descriptors used. Interestingly, the model fits well even though we did not include energy-levels of the dissociated species. Apparently, the electronic state of the non-dissociated molecules suffices also for predicting dissociative chemisorption.

What does this mean in simple chemical terms? To predict the chemisorption properties of a new metal or adsorptive, one can simply calculate the descriptor values for it following Table 1, and then feed these to the model (using the coefficient values from the ESI†). As long as the descriptor values are within the population limits,[37] this should give a good idea on how this new adsorptive or metal would react with all metals or adsorptives already described by the model. All this can be done without running any experiments (do note that the DFT derived energies pertain to 0 K).

When applying the model as such, one should keep in mind that it is based on DFT data without van der Waals corrections. However, as chemisorption is more a charge-transfer than a van der Waals interaction, and as Bligaard *et al.* used the rPBE functional, which is well suited for describing adsorption, good results can be expected.[38] Moreover, these DFT data do not suffer from accuracy problems that actual adsorption experiments suffer from, as we show in our experimental example below.

Finally, note that the isosteric heat of adsorption used here is only valid until a certain threshold coverage.[39] Beyond this, adsorbate–adsorbate interactions will influence the measured (apparent) heat of adsorption. Threshold values differ for each adsorptive–surface combination, but some general rules apply. For example, hydrogen bonding between adsorbed molecules (as in water) will lower the threshold. Conversely, small inert species, such as helium, will have a fairly high coverage threshold. This is one reason why the model deviates with non-linear adsorptives. Such molecules are larger, with more adsorbate–adsorbate interactions. The coverage used in the DFT calculations lies above the threshold for isosteric adsorption for these molecules. Consequently, the values calculated for these compounds do not fit well in the same model with the other compounds calculated at the same isosteric coverage.

Applying the method to experimental adsorption data

The real test for the models is of course in describing actual experimental adsorption behaviour. A good model should account both for differences in adsorption capacity as well as for different types of adsorption isotherms. Thus, we measured the initial adsorption isotherms of six common adsorptives (CO, CO_2, CH_4, H_2, N_2 and O_2) on three representative metals (1 wt% Ni, Pt or Rh) supported on titania. These were compared with bare titania, which underwent the same pretreatment.

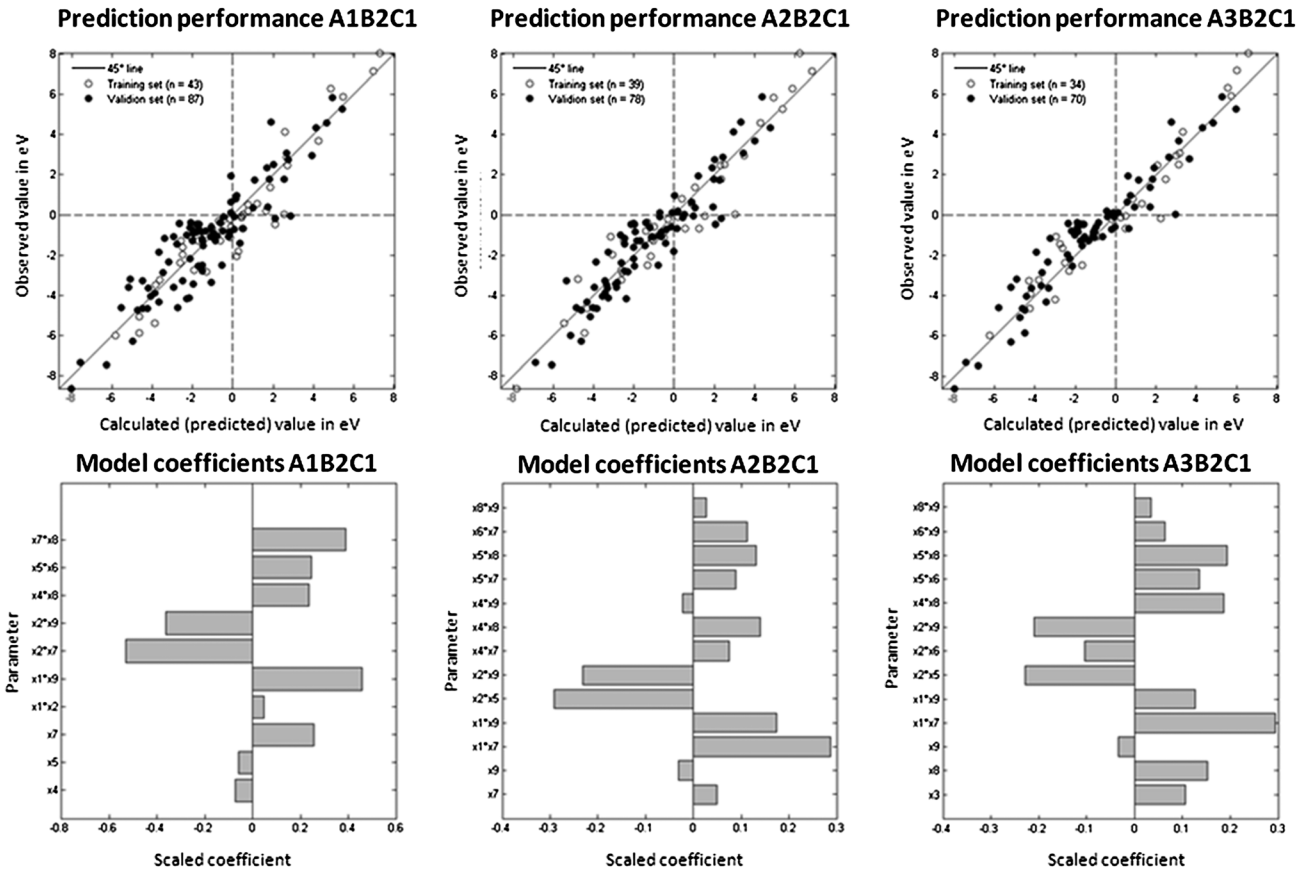
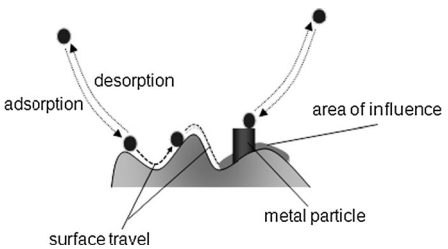


Fig. 5 Correlation plots (observed *vs.* predicted) between the DFT results and our models' predictions (top) and scaled model coefficients (bottom) for the three best models. Model 4 (left), model 13 (middle) and model 22 (right). The improvement of the correlation going from left to right confirms that OH• and H_2 are essentially outliers in the total data set, and thus should be treated separately.

Even on similar catalysts, comparing measured adsorption isotherms is complicated.[40] For starters, the absolute adsorption capacity can vary over several orders of magnitude for different adsorptives. Moreover, various types of adsorption behaviour are common. In the most simple case, adsorption first progresses to a stage where the surface is covered by a monolayer (the Langmuir isotherm). But complex cases, where adsorbate–adsorbate interactions are equally strong as surface–adsorbate interactions, are frequent.

Furthermore, we cannot ignore support effects, since the majority of a catalyst's surface is support material.[41] Consequently, a large part of the adsorption is determined by the support's affinity for the adsorptive (Scheme 2). Here we kept the support constant, so our models do not include explicit support effects. As an example, Fig. 6 shows the experimental adsorption isotherms for CO_2 on bare titania and on the three metal/titania catalysts (full isotherms for all samples are given in ESI†). We see that Ni enhances the affinity for CO_2 compared to bare TiO_2, while Rh decreases this affinity. Conversely, Pt has a much less pronounced effect as indicated by the similar shape and magnitude of the isotherm compared to the bare titania. These results show that a small amount of metal can have a large impact on the adsorption behaviour. Apparently the metal's electronic properties influence the surrounding area of the support. Therefore, as long as the support is the same, trends in adsorption behaviour can be described by the metal properties. Were this not the case, all four materials would give near-identical isotherms.



Scheme 2 Most of the initial adsorption/desorption events occur on the support surface, whereafter the adsorbed species travel to the active site. The fact that a small amount of metal can influence the entire adsorption isotherm shows that there is an area of influence surrounding the metal particles. The size of this region, denoted in light green, is correlated with the density of the metal's valence electrons. Note that the quantities we explain using modelling are equilibrium values. This applies to both the DFT and experimental data.

When modelling this experimentally measured data, we made two generalisations. First, to avoid confusion with regard to the various discrete thermal events occurring during initial adsorption, we focussed on the adsorbed volume rather than

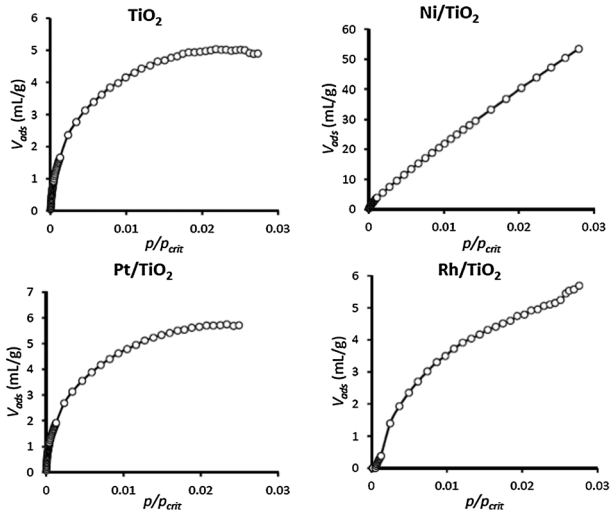


Fig. 6 Experimental adsorption isotherms for CO_2 on the catalyst systems studied. The horizontal axis denotes the reduced pressure of CO_2, the vertical axis denotes the volume of adsorbed CO_2 in cubic centimeters per gram of catalyst material. All metal loadings are constant at 1 wt% relative to the mass of TiO_2. For comparison, the isotherm for bare TiO_2 is also given.

Gas	Metal	V_{ads} (mL/g)	Gas	Metal	V_{ads} (mL/g)
CH_4	Ni	n.a.	H_2	Ni	3.81E-02
CH_4	Pt	2.53E-02	H_2	Pt	5.01E-03
CH_4	Rh	n.a.	H_2	Rh	1.52E-02
CO	Ni	3.72E-01	N_2	Ni	7.82E-03
CO	Pt	1.41E+00	N_2	Pt	2.82E-03
CO	Rh	1.58E+00	N_2	Rh	1.52E-02
CO_2	Ni	3.66E+00	O_2	Ni	n.a.
CO_2	Pt	1.72E+00	O_2	Pt	n.a.
CO_2	Rh	2.79E-01	O_2	Rh	1.03E-01

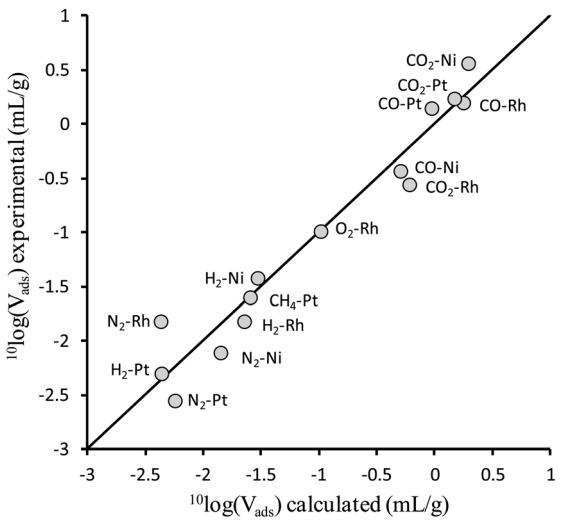


Fig. 7 Parity plot for the descriptor model based on experimental adsorption data on supported metals. The horizontal axis represents the adsorbed volume as calculated by the model, the vertical axis represents the experimentally obtained adsorption value. In the table, the experimentally obtained adsorbed volume at p/p_{crit} = 0.001 is given.

on the heat of adsorption. Since the two are correlated linearly, one can expect the same descriptors to be valid for both. Second, to allow a comparison of the values obtained for the various adsorptives and metals in a single model, one must define a comparable measure for adsorption. Here we selected the adsorbed volume per unit mass and the reduced pressure p/p_{crit} equal to 0.001 for all samples. The rationale is that at this reduced pressure, the coverage is sufficiently low, minimizing adsorptive–adsorptive interactions.

For modelling the experimental adsorption data we used the set of descriptors that explained best the DFT data (model 4 in Table 2). Fig. 7 shows the observed *vs.* predicted plot (parity plot) for the new model along with the measured values. Again, the quality of the model is remarkable (R^2 = 0.94, Q^2 = 0.86). This confirms that the descriptors identified above as the important ones for modelling adsorption based on the DFT data are equally valid for experimental data.

Conclusions

In this work we demonstrate that with simple, easily accessible descriptors a complex phenomenon like dissociative chemisorption can be described using empirical models. The first demonstration relates to *in silico* generated adsorption data. Our models can predict these data with Q^2 values ranging from 0.91–0.95 and RMSEP values ranging from 0.94–1.16 eV. In a second demonstration, we show that the same descriptors applied to experimental data also do an excellent job, explaining the adsorption volumes at constant pressure for a diverse set of gases and real life catalysts, with a Q^2 of 0.86.

Despite the inherent complexity of catalysis at solid surfaces, there is much to be gained from simple heuristic models. This is certainly true in the case of the adsorption of molecules on metal surfaces. By choosing simple descriptors that tally with chemical intuition, one can obtain fast and effective predictions of adsorption energies. We have shown this by modelling and predicting adsorption for a wide range of metals and adsorptives. Our models perform well both for DFT data and actual experimental data. Combined with the ease of adding new elements and molecules, this creates a powerful and practicable tool for finding new catalysts and optimizing existing ones. We strongly believe that combining "quick and dirty" descriptor models with high-level models and screening experiments will open exciting opportunities in catalyst discovery and optimisation in the coming decades.

Definitions

Training set — Subset of the data set that is used to construct a model.

Validation or prediction set — Subset of the data that is not used to construct the model but is used as an unseen test set to assess model performance.

R^2 — Squared correlation coefficient between the observed values and the values calculated by the model for the training set.

Q^2	Squared correlation coefficient between the observed values and the values predicted by the model for the validation set.
RMSEE	Root mean square error of estimate for the observations in the training set.
RMSEP	Root mean square error of estimate for the observations in the validation set.

Acknowledgements

We thank the Netherlands Organization for Scientific Research (NWO) for funding received under the Casimir program, and the Netherlands Research School Combination-Catalysis (NRSC-C) for a fundamental research project grant.

Notes and references

1 G. Rothenberg, *Catalysis: Concepts & Green Applications*, Wiley-VCH, 2008.
2 R. E. Oesper, *J. Chem. Educ.*, 1948, **25**, 531–536.
3 G. Occhipinti, H. R. Bjørsvik and V. R. Jensen, *J. Am. Chem. Soc.*, 2006, **128**, 6952–6964.
4 E. Burello and G. Rothenberg, *Int. J. Mol. Sci.*, 2006, **7**, 375–404.
5 J. A. Hageman, J. A. Westerhuis, H.-W. Frühauf and G. Rothenberg, *Adv. Synth. Catal.*, 2006, **348**, 361–369.
6 C. Klanner, D. Farrusseng, L. Baumes, M. Lengliz, C. Mirodatos and F. Schüth, *Angew. Chem., Int. Ed.*, 2004, **43**, 5347–5349.
7 M. A. Bànares and G. Mestl, *Adv. Catal.*, 2009, **52**, 43–128.
8 S. Chilukoti, E. Widjaja, F. Gao, H. Zhang, J. W. H. Niemantsverdriet and M. Garland, *Phys. Chem. Chem. Phys.*, 2008, **10**, 3535–3547.
9 C. H. Christensen and J. K. Nørskov, *J. Chem. Phys.*, 2008, **128**, 182503.
10 J. D. Grunwaldt, B. Kimmerle, A. Baiker, P. Boye, C. G. Schroer, P. Glatzel, C. N. Borca and F. Beckmann, *Catal. Today*, 2009, **145**, 267–278.
11 B. Kimmerle, J. D. Grunwaldt, A. Baiker, P. Glatzel, P. Boye, S. Stephan and C. G. Schroer, *J. Phys. Chem. A*, 2009, **113**, 3037–3040.
12 A. N. Parvulescu, D. Mores, E. Stavitski, C. M. Teodorescu, P. C. A. Bruijnincx, R. J. M. K. Gebbink and B. M. Weckhuysen, *J. Am. Chem. Soc.*, 2010, **132**, 10429–10439.
13 M. P. Andersson, T. Bligaard, A. Kustov, K. E. Larsen and J. Greeley, *J. Catal.*, 2006, **239**, 501–506.
14 T. Bligaard and J. K. Nørskov, *Phys. Rev. Lett.*, 2007, **016105**, 4–7.
15 T. Bligaard, J. K. Norskov, S. Dahl, J. Matthiesen, C. H. Christensen and J. Sehested, *J. Catal.*, 2004, **224**, 206–217.
16 M. T. M. Koper, T. E. Shubina and R. A. v. Santen, *J. Phys. Chem. B*, 2002, **106**, 686–692.
17 D. Loffreda, F. Delbecq, D. Simon and P. Sautet, *J. Chem. Phys.*, 2001, **115**, 8101–8111.
18 M. Neurock, S. A. Wasileski and D. Mei, *Chem. Eng. Sci.*, 2004, **59**, 4703–4714.
19 J. K. Nørskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, *Nat. Chem.*, 2009, **1**, 37–46.
20 F. J. Keil, *Top. Curr. Chem.*, 2012, **307**, 69–107.
21 E.-J. Ras, M. J. Louwerse and G. Rothenberg, *Catal. Sci. Technol.*, 2012, **2**, 2456–2464.
22 J. N. Brønsted, *Chem. Rev.*, 1928, 231–338.
23 M. G. Evans and M. Polanyi, *Trans. Faraday Soc.*, 1938, **34**, 11–24.
24 J. K. Nørskov, T. Bligaard, B. Hvolbaek, F. Abild-Pedersen, I. Chorkendorff and C. H. Christensen, *Chem. Soc. Rev.*, 2008, 2163–2171.
25 J. K. Nørskov, T. Bligaard, A. Logadottir, S. Bahn, L. B. Hansen, M. Bollinger, H. Bengaard, B. Hammer, Z. Sljivancanin, M. Mavrikakis, Y. Xu, S. Dahl and C. J. H. Jacobsen, *J. Catal.*, 2002, **209**, 275–278.
26 D. Loffreda, F. Delbecq, F. Vigné and P. Sautet, *Angew. Chem., Int. Ed.*, 2009, 8978–8990.
27 P. Sabatier, *Ber. Dtsch. Chem. Ges.*, 1911, **44**, 1984–2001.
28 J. K. van der Waal, *Catal. Today*, 2011, **171**, 207–210.
29 G. C. Demitras and E. L. Muetterties, *J. Am. Chem. Soc.*, 1977, **99**, 2796–2797.
30 A. Griboval-Constant, J.-M. Giraudon, G. Leclercq and L. Leclercq, *Appl. Catal., A*, 2004, **260**, 35–45.
31 L. Vitosa, A. V. Rubana, H. L. Skrivera and J. Kollar, *Surf. Sci.*, 1998, **411**, 186–202.
32 *PLS Toolbox v4.2 for Matlab*, Eigenvector Research Inc.
33 P. Geladi and B. R. Kowalski, *Anal. Chim. Acta*, 1986, **185**, 1–17.
34 A. F. P. Ferreira, M. C. Mittelmeijer-Hazeleger and A. Bliek, *Microporous Mesoporous Mater.*, 2006, **91**, 47–52.
35 G. Rothenberg, *Catal. Today*, 2008, **137**, 2–10.
36 A. Boutin, R. J. M. Pellenq and D. Nicholson, *Chem. Phys. Lett.*, 1994, **219**, 484–490.
37 Most implementations of the PLS algorithm provide more elaborate means (DModX, PModX, Hotelling's *T* squared *etc.*) to evaluate whether a new data point falls within the scope of the model.
38 B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**, 7413–7421.
39 O. Talu and R. L. Kabel, *AIChE J.*, 1987, **33**, 510–514.
40 K. S. W. Sing, *Pure Appl. Chem.*, 1982, **54**, 2201–2218.
41 G. B. Raupp and J. A. Dumesic, *J. Phys. Chem.*, 1985, **89**, 5240–5246.