Catalysis Science & Technology

Accepted Manuscript

Volume 1 | Number 1 | Jan 2013 | Pages 1–100
Catalysis Science & Technology
www.rsc.org/catalysis
ROYAL SOCIETY OF CHEMISTRY

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/catalysis

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Identifying Systematic DFT Errors in Catalytic Reactions[†]

Rune Christensen, Heine A. Hansen, and Tejs Vegge*

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Using CO_2 reduction reactions as an example, we present a widely applicable method for identifying the main source of errors in density functional theory (DFT) calculations. The method has broad applications for error correction in DFT calculations in general, as it relies on the dependence of the applied exchange-correlation functional to the reaction energies rather than on errors versus experiment. As a result, improved energy corrections can now be determined for both gas phase and adsorbed reaction species, particularly interesting within heterogenious catalysis. We show that for CO_2 reduction reactions, the main source of error is associated with the C=O bonds and not the typically energy corrected OCO backbone.

Electroreduction of CO_2 using electricity from renewable sources has the potential to supply carbon neutral transportation fuels, but improved electrocatalysts are needed. Although research within this area has been conducted for decades,[1,2] current challenges include low efficiency, product selectivity and stability of the catalysts. In recent years, significant progress in understanding catalytic activity from fundamental reaction mechanisms has been made for a variety of different heterogeneous electrocatalysts using density functional theory (DFT).[3,4] Within DFT, several levels of calculational complexity in describing the exchange-correlation functional exists. Although higher level methods can be used in some cases, computational cost will often limit treatment of the exchange energy to be performed using the Generalized Gradient Approximation (GGA) for extensive studies of heterogeneous catalysis.[5] For this reason, this work focuses solely on improving the accuracy of calculations relying on functionals with GGA type exchange.

The RPBE[6] and BEEF-vdW[7] exchange-correlation functionals have been developed for catalysis studies and shown to be well suited for determining chemisorption energies.[8,9] However, experimental gas phase reaction enthalpies for CO_2 reduction to products of interest can in many cases not be reproduced within an error of 0.5 eV per CO_2, due to significant systematic errors.[10–12] These inaccuracies can prevent, e.g., accurate determination of product selectivity. To remain at a sufficiently low level of computational cost while improving accuracy, energy corrections have previously been applied in a fitting procedure, where the mean absolute error (MAE) versus experimental data has been minimized for a given set of reactions.[10–12] As a result, energy corrections were previously applied to molecules containing an oxygen-carbon-oxygen (OCO) backbone structure, e.g. CO_2 and HCOOH, for the RPBE functional[10] with an additional correction on the H_2 molecule for the BEEF-vdW functional[11,12]. This correction scheme has subsequently been widely accepted and applied in a large number of high impact papers, e.g. Refs[10,13,14].

The minimized MAE can vary by as little as 0.01 eV for different choices of corrections,[11] making the choice of corrections based solely on the minimized MAE vulnerable to both minor calculational and experimental inaccuracies. Here, we present a new and robust approach for identification of the specific molecules or molecular components requiring an energy correction to obtain the needed accuracy. Here, we demonstrate it for CO_2 reduction reactions, but the approach is generally applicable, e.g. in studies of heterogeneous catalysis of organic compounds or oxygen

---

Department of Energy Conversion and Storage, Technical University of Denmark, Fysikvej bld. 309, DK-2800 Kgs. Lyngby, Denmark. Fax: +45 46 77 57 58; Tel: +45 46 77 58 18; E-mail: teve@dtu.dk

[†] Electronic Supplementary Information (ESI) available: Computational details, extended ensemble correlation analysis, and atomic structures. See DOI: 10.1039/b000000x/

reduction.

Although inferior to RPBE and BEEF-vdW for determination of chemisorption energies,[8,9] we find the PBE functional[15], which is also frequently used within heterogeneous catalysis, able to reproduce most experimental gas phase enthalpies of reaction for CO_2 reduction reactions with sufficient accuracy. The PBE and RPBE functionals differ only in enhancement factor in the GGA exchange energy.[6] If the calculated energy of one specific molecule or molecular component is particularly sensitive towards changes in enhancement factor, it will dominate the ability to reproduce the experimental enthalpies of reaction. Sensitivity towards changes in the enhancement factor can be probed by calculating enthalpies of reaction with GGA type functionals with different enhancement factors. The enhancement factor in the BEEF-vdW functional is composed of a sum of Legendre polynomials with expansion coefficients determined in a machine learning process to obtain the best GGA functional for a range of databases containing different data such as enthalpies of formation, chemisorption energies, reaction barriers and van der Waals interactions. Intrinsic to the functional is an ensemble of functionals with perturbed expansion coefficients. Perturbations are included into the ensemble of functionals based on how well the perturbed functional performs and a "temperature" parameter such that one standard deviation in a quantity calculated with the generated ensemble of functionals corresponds to the predicted calculational error.[7] Ensemble functionals are a computationally very efficient method for calculating enthalpies of formation with a large number of functionals with different enhancement factors. Here, 2000 different ensemble functionals have been examined. The vdW-DF[16] and vdW-DF2[17] functionals are also examined for comparison as they contain vdW correlation similar to that in BEEF-vdW.

Here, the reaction enthalpy is calculated for a range of different CO_2 reduction reactions listed in Table 1. The gas phase reactions in the 'Primary Set' are identical to those examined previously for establishing general energy corrections.[†] [10–12] In addition to the Primary Set, a 'Verification Set' is also introduced. With the exception of reduction to dimethyl ether, i.e. reaction (15), the Verification Set consists of reactions with product molecules containing carbon-oxygen double bonds (C=O). The stoichiometry of all gas phase reactions are normalized to one CO_2 reactant molecule. This is not required but simplifies data treatment.

In addition to gas phase molecules, functional dependent errors can in contrast to previously, also be examined for surface adsorbates. A highly important example; carboxyl (COOH*) adsorption on a Cu (111) surface is presented. By comparing the functional dependence of a surface reaction with a similar gas phase reaction, it can be determined whether a similar correction should be applied for the two. To obtain functional dependence comparable to the gas phase, adsorbed methyl (CH_3*) has been used as reactant as both COOH* and CH_3* bond to a single copper atom through a carbon atom with 3 additional covalent bonds. The compared reactions, (*1a) and (*1b), can be seen in Table 1.

Table 1 Reactions examined

		Primary Set	Reac
H_2+CO_2	$\rightarrow$	$CO+H_2O$	(0)
$4\ H_2+CO_2$	$\rightarrow$	$CH_4+2\ H_2O$	(1)
H_2+CO_2	$\rightarrow$	HCOOH	(2)
$3\ H_2+CO_2$	$\rightarrow$	CH_3OH+H_2O	(3)
$3\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ C_2H_5OH+\frac{3}{2}\ H_2O$	(4)
$\frac{10}{3}\ H_2+CO_2$	$\rightarrow$	$\frac{1}{3}\ C_3H_8+2\ H_2O$	(5)
$\frac{7}{2}\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ C_2H_6+2\ H_2O$	(6)
$3\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ C_2H_4+2\ H_2O$	(7)
$\frac{11}{4}\ H_2+CO_2$	$\rightarrow$	$\frac{1}{4}\ C_4H_6{}^a+2\ H_2O$	(8)
$2\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ CH_3COOH+H_2O$	(9)
$2\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ HCOOCH_3+H_2O$	(10)
		Verification Set	
$2\ H_2+CO_2$	$\rightarrow$	CH_2O+H_2O	(11)
$\frac{3}{2}\ H_2+CO_2$	$\rightarrow$	$OCHCHO+H_2O$	(12)
$\frac{7}{3}\ H_2+CO_2$	$\rightarrow$	$\frac{1}{3}\ C_2H_5COOH^b+\frac{4}{3}\ H_2O$	(13)
$\frac{7}{3}\ H_2+CO_2$	$\rightarrow$	$\frac{1}{3}\ CH_3COOCH_3+\frac{4}{3}\ H_2O$	(14)
$3\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ CH_3OCH_3+\frac{3}{2}\ H_2O$	(15)
$\frac{5}{2}\ H_2+CO_2$	$\rightarrow$	$\frac{1}{2}\ CH_3CHO+\frac{3}{2}\ H_2O$	(16)
		Adsorbate Set	
$2\ H_2O+CH_4$	$\rightarrow$	$HCOOH\ +\ 3\ H_2$	(*1a)
$2\ H_2O+CH_3{}^*$	$\rightarrow$	$COOH^*\ +\ 3\ H_2$	(*1b)

[a] 1,3-Butadiene; [b] Propanoic acid; *Adsorbate on Cu(111).

All calculations have been performed using the Vienna Abinitio Simulation Package (VASP)[18–21] and the Atomic Simulation Environment (ASE),[22] which has been used to generate the ensembles.[†]

Examples of correlated reaction enthalpies can be seen in Figure 1. The observed linear correlations indicate that the functional dependence is dominated by a single molecule or molecular component, or by linearly dependent molecules or molecular components. By assuming a specific molecule or molecular component to dominate the functional dependence, it is possible to predict the slope by dividing the change in number of occurrences in the y-axis reaction with the change in the x-axis reaction as exemplified below. The predicted slope can subsequently be validated against the observed slope. Assuming molecules with an oxygen-carbon-oxygen (OCO) backbone structure to be the major source of functional dependence and thus error,[10–12] the slope for reactions in Figure 1a is predicted to be 1.0 (-1/-1) since in both reactions an OCO backbone is present in the CO_2 reactant, but not in the product giving a change of -1 in both reactions. This fits well with the observed slope.

Figure 1(b-c) show functional dependence approximately following lines with slope 0.5. This disagrees with the general as-

This journal is © The Royal Society of Chemistry [year]

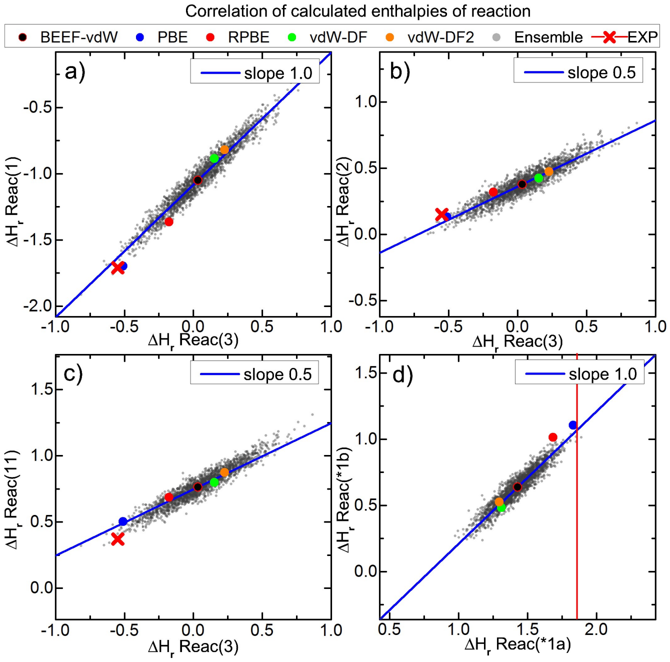


Fig. 1 Correlations in calculted enthalpies of reaction (eV) for various reactions. Functional dependence on energies are observed to correlate linearly for different reactions including surface reactions. Blue lines are drawn with predicted slopes equal to the ratio of broken/formed C=O bonds in the compared reactions. Larger points are self-consistent calculations using different functionals, crosses (red line in **d**) experimental reference values,[23] and the smaller grey semi-transparent points 2000 BEEF-ensemble functionals.

sumption that the OCO backbone dominates functional dependence, as this assumption predicts slopes of 0 for Figure 1b, since the OCO backbone is present in both reactant and product (HCOOH) giving 0 change in y-axis reaction (0/-1), and slope 1.0 for Figure 1c, which is similar to Figure 1a in terms of changes in OCO backbones. The observed slopes do, however, fit with C=O bonds dominating the functional dependence. In both reactions plotted on the y-axes, one of the double bonds in the CO_2 reactant is preserved in the product (HCOOH and H_2CO) and both broken in reaction (3) plotted on the x-axes resulting in predicted slopes of 0.5 (-1/-2). A more in-depth quantitative analysis of the total reaction set finds slopes obtained through linear regression to agree very well with slopes predicted under the assumption that the C=O bonds dominate the functional dependence.†

Figure 1d compares the two reactions in the Adsorbate Set. They are observed to show the same functional dependence, and COOH* (carboxyl) should thus be corrected with the same energy correction as HCOOH (formic acid). From the geometry and bond lengths this is expected as a C=O bond appears to be present in COOH*.† Dispersion forces included in the vdW functionals have an effect for the adsorbate reactions as the vdW functionals appear to be on a lower but parallel line to the non-vdW functionals, although too few non-vdW data points are present for definite conclusions. The offset does not impact the analysis of the functional dependence.

Having identified the C=O bond to be dominating the functional dependence, energy corrections are applied based on the number of C=O bonds in a molecule rather than to molecules with an OCO backbone. The magnitude of the energy correction is then determined by minimizing the MAE versus experiments for the Primary Set of reactions shown in Table 1 in a procedure identical to the one previously applied for determining corrections using the same reference data.[10–12] To make a direct comparison with the previously used corrections for the BEEF-vdW functional, corrections are also applied to H_2 molecules. A comparison of corrections can be seen in Figure 2. For both types of corrections (C=O and OCO), the optimal magnitude of corrections is 0.10 eV for H_2 molecules and 0.29 eV for CO_2 (0.15 eV per C=O bond). This is similar although not identical to what has been found previously (0.09 eV and 0.33 eV,[11] and 0.09 eV and 0.41 eV[12]). The magnitude of corrections is the same within 0.01 eV per bond if reactions in the Verification Set is included in minimization of the MAE. By applying a correction to C=O instead of OCO, a significant reduction in post-correction errors are observed for reactions, where the main product either has the OCO backbone, but only one C=O bond or C=O bonds without the OCO backbone structure (See Figure 2). This is seen to be the case in both the Primary and the Verification Set. This trend is also observed for the RPBE functional. In the case of reduction to HCOOH (reaction(3)), the post-correction error in the calculations with BEEF-vdW changes from 0.13 eV to -0.02 eV. Previously, errors of 0.15 eV[11] and 0.17 eV[12] were obtained for this reaction with the OCO backbone and H_2 corrected BEEF-vdW functional. For $HCOOCH_3$ (reaction (10)) and CH_2O (reaction(11)), significantly different experimental gas phase enthalpies of formation are available in the NIST database, which is used as source for all reference data.[23] The previously used experimental value for $HCOOCH_3$ is by far the highest in the database and extrapolated from the liquid phase enthalpy of formation.[23] Using one of the three alternative experimental references in the NIST database, the error for reaction (10) follows the trend of the other reactions and can be decreased to a few meV after the C=O correction.

The effect of changing the correction scheme and applying corrections to the adsorbate can be very significant as exemplified by the reduction of COOH* to HCOOH. Computed with the BEEF-vdW functional, the reaction energy changes significantly from having a free energy change of -0.82 eV with OCO corrections, to -1.11 eV with C=O corrections.

Optimal magnitude of corrections have been determined for all ensemble functionals. The standard deviation is then calculated for the ensemble of corrected functionals and plotted as error bars

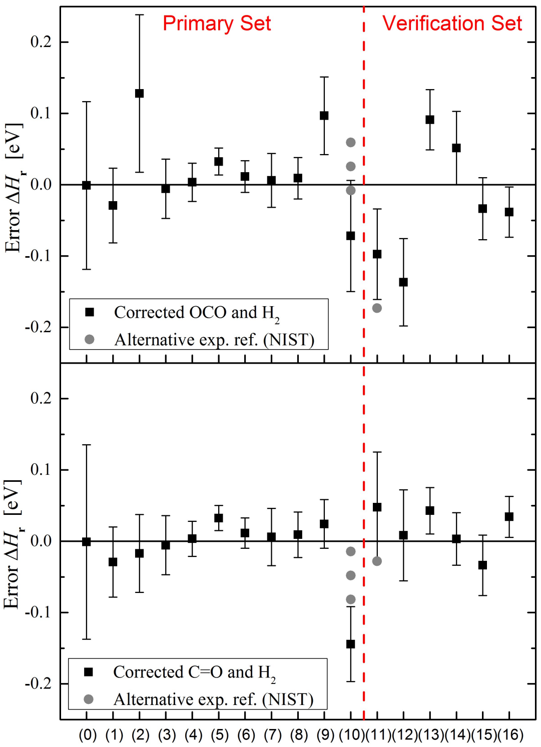


Fig. 2 Comparison of remaining errors after correction of OCO backbone (above) or C=O bonds (below). Error bars show one standard deviation for the corrected ensemble. The grey points mark the error with alternative experimental references present in the NIST database[23].

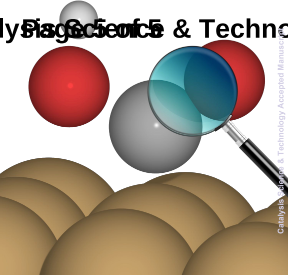
in Figure 2. This can be used as a measure of how well the correction performs for the class of functionals. The ensemble standard deviation is generally decreased with the C=O correction for reactions where it applies differently than the OCO correction, e.g. for reaction (2); the reduction to formic acid. With the C=O corrections, the standard deviation for reaction (0) is significantly larger than for the other reactions. As this is the only depicted reaction including CO, the relatively large standard deviation suggests that the energy of CO could also be functional dependent and require correction on the order of 0.2 eV with certain functionals. CO has previously been found to require corrections using the PBE functional.[9,24] A functional dependent error is also found for carbon-carbon (C=C) double bonds, as described in the ESI.† It is, however, of minor importance for the reactions considered here.

The demonstrated method is not limited to the presented case or catalytic reactions and can be used to identify a dominating error causing structure, molecule or molecular component in other cases, where functional dependence is observed.

For different GGA functionals, the difference in enhancement factor will be small for low density gradients and grow as the density gradient increases.[7] This can explain why functional dependence is most notable for C=O bonds, C=C bonds, and the CO molecule, as they are probably giving rise to the largest density gradients for the species in the reactions. In the future, the reduced density gradient could potentially be used directly for qualitative identification of molecules or molecular components with large functional dependence.

The authors acknowledge support from the Catalysis for Sustainable Energy (CASE) initiative funded by The Danish Agency for Science, Technology and Innovation.

References

1 Y. Hori, *Modern Aspects of Electrochemistry*, Springer New York, 2008, vol. 42, pp. 89–189.
2 J. Qiao, Y. Liu, F. Hong and J. Zhang, *Chem. Soc. Rev.*, 2014, **43**, 631–675.
3 Y. Li, S. H. Chan and Q. Sun, *Nanoscale*, 2015, **7**, 8663–8683.
4 S. Lysgaard, J. S. G. Myrdal, H. A. Hansen and T. Vegge, *Phys. Chem. Chem. Phys.*, 2015, –.
5 M. K. Sabbe, M.-F. Reyniers and K. Reuter, *Catal. Sci. Technol.*, 2012, **2**, 2010–2024.
6 B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B*, 1999, **59**, 7413–7421.
7 J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard and K. W. Jacobsen, *Phys. Rev. B*, 2012, **85**, 235149.
8 B.-T. Teng, X.-D. Wen, M. Fan, F.-M. Wu and Y. Zhang, *Phys. Chem. Chem. Phys.*, 2014, **16**, 18563–18569.
9 J. Wellendorff, T. L. Silbaugh, D. G. Pintos, J. K. Nørskov, T. Bligaard, F. Studt and C. T. Campbell, *Surf. Sci.*, 2015, doi:10.1016/j.susc.2015.03.023.
10 A. A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl and J. K. Nørskov, *Energy Environ. Sci.*, 2010, **3**, 1311–1315.
11 F. Studt, F. Abild-Pedersen, J. B. Varley and J. K. Nørskov, *Catal. Lett.*, 2013, **143**, 71–73.
12 F. Studt, M. Behrens, E. L. Kunkes, N. Thomas, S. Zander, A. Tarasov, J. Schumann, E. Frei, J. B. Varley, F. Abild-Pedersen, J. K. Nørskov and R. Schlögl, *ChemCatChem*, 2015, **7**, 1105–1111.
13 A. A. Peterson and J. K. Nørskov, *J. Phys. Chem. Lett.*, 2012, **3**, 251–258.
14 M. Behrens, F. Studt, I. Kasatkin, S. Kühl, M. Hävecker, F. Abild-Pedersen, S. Zander, F. Girgsdies, P. Kurr, B.-L. Kniep, M. Tovar, R. W. Fischer, J. K. Nørskov and R. Schlögl, *Science*, 2012, **336**, 893–897.
15 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
16 M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 2004, **92**, 246401.
17 K. Lee, E. D. Murray, L. Kong, B. I. Lundqvist and D. C. Langreth, *Phys. Rev. B*, 2010, **82**, 081101.
18 G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
19 G. Kresse and D. Joubert, *Phys. Rev. B*, 1999, **59**, 1758–1775.
20 P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953–17979.
21 J. c. v. Klimeš, D. R. Bowler and A. Michaelides, *Phys. Rev. B*, 2011, **83**, 195131.
22 S. Bahn and K. W. Jacobsen, *Comput. Sci. Eng.*, 2002, **4**, 56–66.
23 *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, ed. P. J. Linstrom and W. G. Mallard, National Institute of Standards and Technology, 2005.
24 F. Calle-Vallejo and M. T. M. Koper, *Angew. Chem., Int. Ed.*, 2013, **52**, 7282–7285.