

CrossMark  
click for updates

## Identifying systematic DFT errors in catalytic reactions†

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

Cite this: *Catal. Sci. Technol.*, 2015, 5, 4946Received 14th August 2015,  
Accepted 14th September 2015

DOI: 10.1039/c5cy01332a

www.rsc.org/catalysis

Using CO<sub>2</sub> reduction reactions as examples, 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 on the reaction energies rather than on errors *versus* the experimental data. As a result, improved energy corrections can now be determined for both gas phase and adsorbed reaction species, particularly interesting within heterogeneous catalysis. We show that for the CO<sub>2</sub> 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<sub>2</sub> 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,<sup>1,2</sup> 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).<sup>3,4</sup> Within DFT, several levels of calculational complexity in describing the exchange–correlation functional exist. 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.<sup>5</sup> For this reason, this work focuses solely on improving the accuracy of calculations relying on functionals with GGA type exchange.

The RPBE<sup>6</sup> and BEEF-vdW<sup>7</sup> exchange–correlation functionals have been developed for catalysis studies and showed to be well-suited for determining chemisorption energies.<sup>8,9</sup> However, experimental gas phase reaction enthalpies for CO<sub>2</sub> reduction to products of interest can, in many cases, not be reproduced within an error of 0.5 eV per CO<sub>2</sub>, due to significant systematic errors.<sup>10–12</sup> 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.<sup>10–12</sup> As a result, energy corrections were previously applied to molecules containing an oxygen–carbon–oxygen (OCO) backbone structure, *e.g.* CO<sub>2</sub> and HCOOH, for the RPBE functional<sup>10</sup> with an additional correction on the H<sub>2</sub> molecule for the BEEF-vdW functional.<sup>11,12</sup> This correction scheme has subsequently been widely accepted and applied in a large number of high impact papers.<sup>10,13,14</sup>

The minimized MAE can vary by as little as 0.01 eV for different choices of corrections,<sup>11</sup> 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<sub>2</sub> reduction reactions, but the approach is generally applicable, *e.g.* in studies of heterogeneous catalysis of organic compounds or oxygen reduction.

Although inferior to RPBE and BEEF-vdW for determination of chemisorption energies,<sup>8,9</sup> we find the PBE functional,<sup>15</sup> which is also frequently used within heterogeneous catalysis, able to reproduce most experimental gas phase enthalpies of reaction for CO<sub>2</sub> reduction reactions with sufficient accuracy. The PBE and RPBE functionals differ only in enhancement factor in the GGA exchange energy.<sup>6</sup> If the calculated energy of one specific molecule or molecular component is particularly sensitive towards changes in the

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

† Electronic supplementary information (ESI) available: Computational details, extended ensemble correlation analysis, and atomic structures. See DOI: 10.1039/c5cy01332a



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 the 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 in 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 using the generated ensemble of functionals corresponds to the predicted calculational error.<sup>7</sup> The use of ensemble functionals is 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<sup>16</sup> and vdW-DF<sup>217</sup> functionals are also examined for comparison as they contain vdW correlation similar to that of BEEF-vdW.

Here, the reaction enthalpy is calculated for a range of different CO<sub>2</sub> 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.<sup>†10–12</sup> In addition to the primary set, a ‘verification set’ is also introduced. With the exception of the 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 is normalized to one CO<sub>2</sub> reactant molecule. This is not required but this simplifies data treatment.

In addition to gas phase molecules, functional dependent errors can, in contrast to previously used correctional approaches, 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<sub>3</sub>\*) has been used as the reactant as both COOH\* and CH<sub>3</sub>\* 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.

All calculations have been performed using the Vienna *ab initio* Simulation Package (VASP)<sup>18–21</sup> and the Atomic Simulation Environment (ASE),<sup>22</sup> which has been used to generate the ensembles.<sup>†</sup>

Examples of correlated reaction enthalpies can be seen in Fig. 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

Table 1 Reactions examined

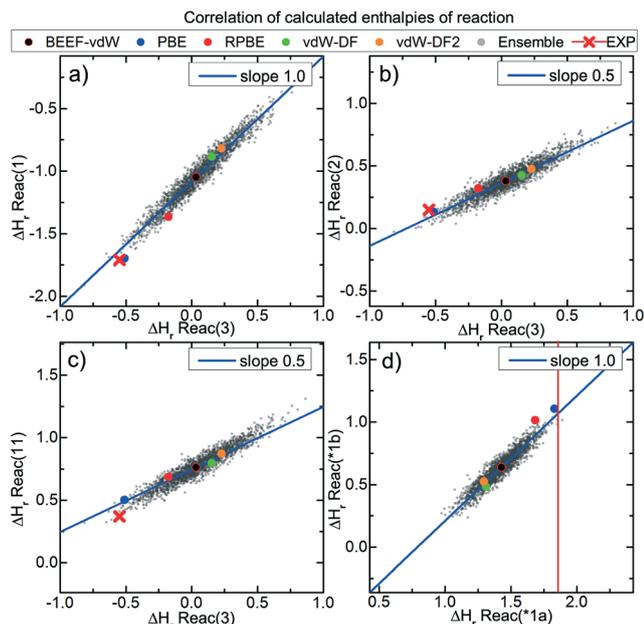
Primary set		Reac
H <sub>2</sub> + CO <sub>2</sub>	→	CO + H <sub>2</sub> O (0)
4H <sub>2</sub> + CO <sub>2</sub>	→	CH <sub>4</sub> + 2H <sub>2</sub> O (1)
H <sub>2</sub> + CO <sub>2</sub>	→	HCOOH (2)
3H <sub>2</sub> + CO <sub>2</sub>	→	CH <sub>3</sub> OH + H <sub>2</sub> O (3)
3H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ C <sub>2</sub> H <sub>5</sub> OH + $\frac{3}{2}$ H <sub>2</sub> O (4)
$\frac{10}{3}$ H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{3}$ C <sub>3</sub> H <sub>8</sub> + 2H <sub>2</sub> O (5)
$\frac{7}{2}$ H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ C <sub>2</sub> H <sub>6</sub> + 2H <sub>2</sub> O (6)
3H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ C <sub>2</sub> H <sub>4</sub> + 2H <sub>2</sub> O (7)
$\frac{11}{4}$ H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{4}$ C <sub>4</sub> H <sub>6</sub> <sup>a</sup> + 2H <sub>2</sub> O (8)
2H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ CH <sub>3</sub> COOH + H <sub>2</sub> O (9)
2H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ HCOOCH <sub>3</sub> + H <sub>2</sub> O (10)
Verification set		
2H <sub>2</sub> + CO <sub>2</sub>	→	CH <sub>2</sub> O + H <sub>2</sub> O (11)
$\frac{3}{2}$ H <sub>2</sub> + CO <sub>2</sub>	→	OCHCHO + H <sub>2</sub> O (12)
$\frac{7}{2}$ H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{3}$ C <sub>2</sub> H <sub>5</sub> COOH <sup>b</sup> + $\frac{4}{3}$ H <sub>2</sub> O (13)
$\frac{7}{3}$ H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{3}$ CH <sub>3</sub> COOCH <sub>3</sub> + $\frac{4}{3}$ H <sub>2</sub> O (14)
3H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ CH <sub>3</sub> OCH <sub>3</sub> + $\frac{3}{2}$ H <sub>2</sub> O (15)
$\frac{5}{2}$ H <sub>2</sub> + CO <sub>2</sub>	→	$\frac{1}{2}$ CH <sub>3</sub> CHO + $\frac{3}{2}$ H <sub>2</sub> O (16)
Adsorbate set		
2H <sub>2</sub> O + CH <sub>4</sub>	→	HCOOH + 3H <sub>2</sub> (*1a)
2H <sub>2</sub> O + CH <sub>3</sub> *	→	COOH* + 3H <sub>2</sub> (*1b)

<sup>a</sup> 1,3-Butadiene. <sup>b</sup> Propanoic acid. \* Adsorbate on Cu(111).

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 the 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,<sup>10–12</sup> the slope for the reactions in Fig. 1a is predicted to be 1.0 (–1/–1) since in both reactions an OCO backbone is present in the CO<sub>2</sub> reactant, but not in the product, giving a change of –1 in both reactions. This fits well with the observed slope.

Fig. 1b and c show the functional dependence approximately following lines with a slope of 0.5. This disagrees with the general assumption that the OCO backbone dominates the functional dependence, as this assumption predicts a slope of 0 for Fig. 1b, since the OCO backbone is present in both reactant and product (HCOOH), giving a change of 0 in the y-axis reaction (0/–1), and a slope of 1.0 for Fig. 1c, which is similar to Fig. 1a in terms of changes in the OCO backbones. The observed slopes do, however, fit with the C=O bonds dominating the functional dependence. In both reactions plotted on the y-axis, one of the double bonds in the CO<sub>2</sub> reactant is preserved in the product (HCOOH and H<sub>2</sub>CO) and both are broken in reaction (3) plotted on the x-axis, resulting in a predicted slope of 0.5 (–1/–2). In a more in-depth quantitative analysis of the total reaction set, one can find the slopes obtained through linear regression to agree





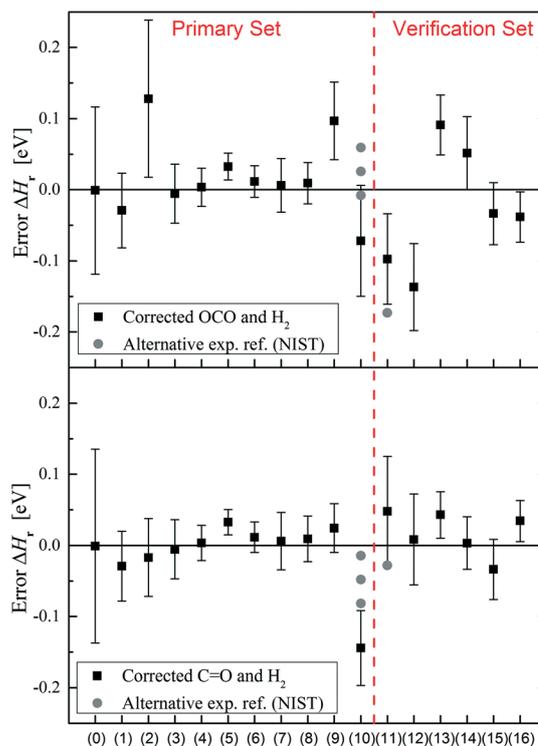
**Fig. 1** Correlations in the calculated enthalpies of reaction (eV) for various reactions (a–d). Functional dependence on energies is observed to correlate linearly for different reactions including surface reactions (d). Blue lines are drawn with predicted slopes equal to the ratio of broken/formed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

very well with the slopes predicted under the assumption that the C=O bonds dominate the functional dependence.<sup>†</sup>

Fig. 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 the C=O bond appears to be present in COOH\*.<sup>†</sup> Dispersion forces included in the vdW functionals have an effect on 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 the experimental data 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.<sup>10–12</sup> To make a direct comparison with the previously used corrections for the BEEF-vdW functional, corrections are also applied to the H<sub>2</sub> molecules. A comparison of corrections can be seen in Fig. 2. For both types of

corrections (C=O and OCO), the optimal magnitudes of corrections are 0.10 eV for the H<sub>2</sub> molecules and 0.29 eV for CO<sub>2</sub> (0.15 eV per C=O bond). These are similar although not identical to what have been found previously (0.09 eV and 0.33 eV,<sup>11</sup> and 0.09 eV and 0.41 eV (ref. 12)). The magnitude of corrections is the same within 0.01 eV per bond if the reactions in the verification set are included in the minimization of the MAE. By applying a correction to C=O instead of OCO, a significant reduction in post-correction errors are observed for the reactions, where the main product either has the OCO backbone but with only one C=O bond or has the C=O bond without the OCO backbone structure (see Fig. 2). This is seen to be the case in both the primary and the verification sets. This trend is also observed for the RPBE functional. In the case of the 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<sup>11</sup> and 0.17 eV<sup>12</sup> were obtained for this reaction using the OCO backbone and H<sub>2</sub> corrected BEEF-vdW functional. For HCOOCH<sub>3</sub> (reaction (10)) and CH<sub>2</sub>O (reaction (11)), significantly different experimental gas phase enthalpies of formation are available in the NIST database, which is used as the source for all reference data.<sup>23</sup> The previously used experimental value for HCOOCH<sub>3</sub> is by far the highest in the database and is extrapolated from the liquid phase enthalpy of formation.<sup>23</sup> Using one of the three alternative experimental



**Fig. 2** Comparison of remaining errors after correction of the OCO backbone (above) or the 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.<sup>23</sup>



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 using 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 magnitudes 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 in Fig. 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.<sup>9,24</sup> 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 the different GGA functionals, the difference in the enhancement factor will be small for low density gradients and increase as the density gradient increases.<sup>7</sup> This can explain why functional dependence is most notable for the C=O bond, the C=C bond, and the CO molecule, as they probably give 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.

## Acknowledgements

The authors acknowledge the 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, DOI: 10.1039/C5CP00298B.
- 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: Condens. Matter Mater. Phys.*, 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: Condens. Matter Mater. Phys.*, 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. Garcia-Pintos, J. K. Nørskov, T. Bligaard, F. Studt and C. T. Campbell, *Surf. Sci.*, 2015, **640**, 36–44.
- 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. Kühn, M. Hävecker, F. Abild-Pedersen, S. Zander, F. Girgsdies, P. Kurr, B.-L. Knief, 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: Condens. Matter Mater. Phys.*, 2010, **82**, 081101.
- 18 G. Kresse and J. Furthmüller, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 11169–11186.
- 19 G. Kresse and D. Joubert, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**, 1758–1775.
- 20 P. E. Blöchl, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**, 17953–17979.
- 21 J. C. V. Klimeš, D. R. Bowler and A. Michaelides, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 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.

