



Error estimates in semi-empirical estimation methods of surface reactions

Jonathan E. Sutton, Dionisios G. Vlachos*

Catalysis Center for Energy Innovation and Center for Catalytic Science and Technology, Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716, United States

ARTICLE INFO

Article history:

Received 11 July 2012

Revised 26 September 2012

Accepted 6 October 2012

Available online 9 November 2012

Keywords:

Error estimate

Error propagation

Linear scaling relation

Group additivity

Brønsted–Evans–Polanyi

ABSTRACT

In this paper, we examine the theoretical and numerical error properties of three semi-empirical methods when used separately and in conjunction with each other using oxygenated hydrocarbons as sample species. These methods are group additivity (GA), the linear scaling relations (LSRs), and the Brønsted–Evans–Polanyi (BEP) correlations. GA estimated heat of formation has nearly zero mean error for all species. GA heats of reaction have no systematic error only when GA is used with C₂ or larger species, and DFT is used for all smaller species. Aside from a constant offset, LSR-estimated binding energies and heats of reaction show adequate agreement with the corresponding DFT energies. BEP-estimated activation energies are shown to have systematic errors when the input heats of reaction have systematic errors. Each method produces estimates whose errors are normally distributed with a nonzero mean and which are uncorrelated. Combinations of techniques also possess normally distributed errors.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Understanding reaction pathways has been a long standing goal of heterogeneous catalysis. Recently, computation-driven catalyst design has also attracted attention for identifying more effective catalysts [1–5]. Simplified Langmuir–Hinshelwood rate expressions and microkinetic modeling have been employed toward this latter goal [6], with parameters calculated via theoretical techniques (e.g., density functional theory). However, as the complexity of the chemistry model increases, the exclusive use of first-principles techniques quickly becomes prohibitively expensive [7,8]. One promising approach to reducing the computational cost is to use inexpensive semi-empirical estimation techniques as initial approximations and then refine the model with the more accurate theoretical techniques only where necessary, an approach known as hierarchical multiscale modeling [9–11].

A number of semi-empirical techniques are widely used in the literature. Enthalpies and entropies of formation for novel molecules can be quickly and accurately estimated using group-additivity (GA) techniques [12–14]. The heat of formation and binding energies on different surfaces can be estimated using the linear scaling relations (LSRs) of Nørskov and coworkers [15–18]. Heat of formation can be used to estimate activation energies using Brønsted–Evans–Polanyi (BEP) type correlations [8,19–24]. These techniques can be used singly or in conjunction with each other to rapidly parameterize a kinetic model. Fig. 1 shows the possible

relationship among techniques and the important model parameters.

GA techniques for estimating the thermochemistry of gas-phase molecules were originally developed by Benson and coworkers [25–28]. They have recently been adapted to the estimation of thermochemical properties of surface species. The first such adaptation for proof of concept was by Kua et al. [12] for simple hydrocarbons. This method was extended by Vlachos and coworkers using graph theory [13,14] for more complex oxygenates, including polyols, acids, ethers, and esters (as well as alcohols, ketones, aldehydes, and hydrocarbons). LSRs were originally developed for estimating the binding energy of AH_x type species on transition metals from the binding energy of the heteroatom A [15] as an outgrowth of the Unity Bond Index–Quadratic Exponential Potential (UBI–QEP) formalism (also known as Bond Order Conservation) [29]. They have since been extended to transition metal oxides, nitrides, and sulfides [16], transition metal carbides [30], and to larger molecules [13,17,18,24]. BEP correlations are linear free energy relations that relate the activation energy to the corresponding heat of reaction for a particular type of molecule and reaction. Recent articles [8,31,32] have reviewed these techniques and the available correlations. More recently, Sutton and Vlachos [32] have investigated the assumptions inherent to the BEP correlations. To our knowledge, they also presented the first detailed analysis of the errors inherent to these correlations.

Each of these techniques has the potential to greatly reduce the computational burden associated with parameterizing a microkinetic model. However, their range of applicability is not well understood and this motivates this work. For example, GA is expected to give fairly accurate results when limited to the types of

* Corresponding author.

E-mail address: vlachos@udel.edu (D.G. Vlachos).

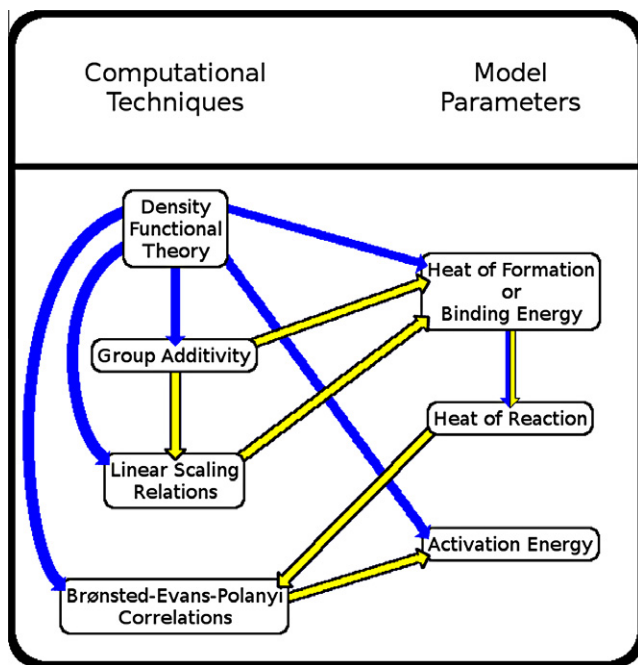


Fig. 1. Graphical depiction of how DFT and the semi-empirical techniques may be used to generate parameters for a microkinetic model. The blue lines indicate DFT is used to generate the thermodynamic properties of groups for GA, the correlation parameters for the linear scaling relations, and Brønsted–Evans–Polanyi correlations or the model parameters (thermochemistry and reaction barriers) directly. The yellow lines show how the semi-empirical techniques may be used to generate the model parameters. GA can be used to estimate heat of formation for novel species on the reference metal directly or on a new metal via the linear scaling relations. The heat of formation is then used to calculate heat of reaction. The activation energies are then estimated from the heat of reaction via the Brønsted–Evans–Polanyi correlations. The equilibrium and rate constants needed in the model can then be calculated from the heat of reaction and activation energies, respectively.

molecules used in deriving the group values. Its accuracy for other molecules and metals is still an open question. Likewise, the LSRs are useful for rapidly estimating the binding energy or heat of formation of a species on a new metal and/or surface facet, but it is still not clear what accuracy can be expected in doing so. Finally, BEP correlations have received more attention, and it is clear that they are suitable for quickly estimating the activation energies for new molecules in a homologous series. Still, the implications of using a published correlation with a set of molecules which do not exactly follow the original distribution have not been fully elucidated.

In this paper, we examine the error distributions of both the heat of formation and the heats of reaction as calculated using GA methods. We also examine the possible error distributions of GA methods on other metal surfaces. For the LSR technique, we first investigate the errors inherent to the simplest form of the LSR model. Then, we extend this understanding to more complex molecules. Finally, we build on prior BEP work to investigate the effect of the correlation parameters, the heats of reaction, and the metal surface on the error distribution.

2. Computational systems

In order to test the inherent accuracy of the various semi-empirical techniques, we compare semi-empirical estimates to DFT-derived values. For testing GA, we use results for ethylene glycol on Pt(111) previously published by Saliccioli et al. [13,33]. These calculations comprise results for all dehydrogenated deriva-

tives as well as their corresponding C–H, O–H, and C–C reactions. For investigating other metal surfaces, we turn to the Ni and surface bimetallic Ni–Pt heat of formation reported by Saliccioli et al. [11,13]. We also have performed additional DFT calculations for the same set of ethylene glycol derivatives on Ru(0001) using the same computational technique as used by Saliccioli et al. In all cases, we assume that the vibrational frequencies are independent of the metal surface involved. A complete list of the species used and the corresponding heat of formation is included in the supporting information.

For the LSRs, we used the SIESTA DFT code [34] to calculate the binding energies of CH_x , OH, and CH_xOH_y species on close-packed Co, Ni, Pd, Pt, Rh, and Ru. In broad terms, we followed the same basic procedure reported by Chen and Vlachos [35]. Spin polarization was included for calculations involving Co, Ni, and any gas-phase species. For these calculations, we used $p(2 \times 2)$ unit cells with four layer slabs and 15 Å of vacuum between each slab. The bottom two layers were frozen in their bulk positions, while the top two layers were allowed to relax. To investigate the application of the LSRs to more complex species, we also used this procedure to calculate the binding energies of dehydrogenated derivatives of ethanol as well as their possible C–C cracking fragments on close-packed surfaces of Co, Ni, Pt, Rh, and Ru. The energies used and their associated species are included in the supporting information. Previous work has shown that $p(2 \times 2)$ cells are sufficiently large for calculating binding energies of ethanol [32].

For the BEP correlations, we again resort to the ethylene glycol calculations of Saliccioli et al. We use their previously published heat of formation at 0 K and at 298 K. We also use the corresponding GA values at 298 K. Besides the ethylene glycol calculations, we have also performed additional DFT calculations to find the transition state energies for a subset of C–C cracking reactions in ethanol on close-packed Co, Ni, Pt, Rh, and Ru. The computational method for these calculations involved a constrained optimization scheme that searches for the geometrical configuration where (1) the forces on each atom vanish and (2) the energy is maximized along the reaction coordinate but is minimized with respect to all other degrees of freedom [36]. All other aspects of these computations were identical to the other ethanol calculations performed in this work. The reactions examined and their transition state energies are given in the supporting information.

3. Statistical definitions

Theoretical analysis of the error distributions was accomplished via standard statistical definitions and procedures [37]. We summarize some introductory methods for unfamiliar readers. We first begin with a brief discussion of errors. Following this, we outline the use of mathematical expectation operators. We conclude with a description of the correlation coefficient.

The error in the random variable X is defined as the difference between the actual value and its corresponding estimate (represented by X and \hat{X} , respectively):

$$\varepsilon_x = X - \hat{X} \quad (1)$$

We wish to emphasize that the error term is ordinarily assumed to be normally distributed with a zero mean when the estimated value is the result of ordinary least squares regression [37]. This assumption is employed here unless otherwise noted. As an extension of this, we assume that the DFT calculations are exact for the purposes of analyzing the errors in estimates made using DFT results in the training data. Errors stemming from other models for the estimated values (e.g., the errors in DFT) may not necessarily follow this distribution. A discussion of such errors is outside the scope of this work.

Download English Version:

<https://daneshyari.com/en/article/6527821>

Download Persian Version:

<https://daneshyari.com/article/6527821>

[Daneshyari.com](https://daneshyari.com)