

## Systematic Validation and Improvement of Quantum Chemistry Methods for the Prediction of Physical and Chemical Properties

C.A. Gonzalez, R. D. Johnson, and K.K. Irikura (838)

The use of modern computational chemistry methodologies in the prediction of molecular properties has become increasingly popular mainly due to significant improvements in the algorithms, the accuracy of the methods, and the advent of powerful computer resources. This is particularly true in the area of thermochemistry, where researchers in industry and academia perform quantum chemistry calculations on a routine basis. Recently, there has been an increased interest in the use of theoretical predictions of physical and chemical properties made by state-of-the-art quantum chemistry methodologies to fill gaps in the available experimental data. This approach could potentially eliminate the cost of experimental measurement and will be particularly useful in cases where extrapolations based on available experimental data are not possible or are unreliable. However, despite the aforementioned progress, the predictive power of most of the quantum chemistry methodologies has not been established on solid ground due to the lack of knowledge of their relative uncertainties. To understand these uncertainties an exhaustive and systematic validation of these methodologies involving robust strategies to compare theory and experiment is necessary. In addition, the results of such validations must be widely available to the scientific community so new scientific discoveries could leverage from this knowledge.

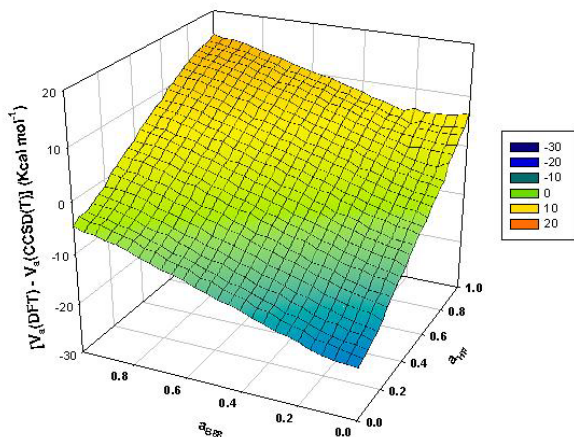
**A team of NIST researchers is working on an integrated approach that will lead to the systematic validation as well as improvements of popular and new theoretical methods using high quality experimental data.**

The NIST approach involves the development of state-of-the-art databases in order to disseminate the accuracy, reliability and transferability of the most popular quantum chemistry methodologies in the prediction of physical and chemical properties of a large variety of chemical systems ranging from isolated atoms and molecules to condensed phase. The theoretical methodologies being studied include semi-empirical Hamiltonians, wavefunction-based *ab initio* molecular orbital theory, and Density Functional Theory (DFT) while the experimental data used in the validation include measurements already available in the literature and also made by researchers at NIST.

*Assessing the Accuracy and Reliability of Density Functional Theory.* Density Functional Theory (DFT) has become one of the most widely used quantum chemistry methods mainly due to its relatively low computational expense. The electronic energy within the DFT formalism is a functional of the electron density that consists of an “exchange” energy term and a “correlation” energy term. Although in principle there exists a universal DFT functional that can describe the electronic problem exactly, the discovery of such a functional has eluded scientists for years. In practice, arbitrary forms for the DFT functionals are parameterized in order to reproduce the experimental energetics of a relatively small set of molecular systems.

CSTL researchers have performed a systematic validation of a particular family of DFT functionals, the so-called “hybrid-GGA”, in order to assess their accuracy and transferability to different chemical properties. The performance of these functionals in the prediction of properties such as singlet-triplet gaps, reaction barriers, dipole moments, and polarizabilities showed a marked erratic behavior leading to the conclusion that these functionals are not transferable to the computation of different chemical properties. In addition, the results of this work indicate that the major source of error can be traced back to the “exchange” functional, which exhibits significant spurious and uncontrollable electron “correlation” energy due to the empirical parameterization used in the development of these functionals.

DFT Classical Barrier Error (in kcal mol<sup>-1</sup>) Relative to CCSD(T)  
for CH<sub>3</sub> + H<sub>2</sub> → CH<sub>4</sub> + H<sub>2</sub>



of closed forms for the derivatives of these functionals with respect to the electron density. These derivatives (or potentials as they are known) are critical in the calculation of molecular properties in a “variational” manner and are difficult if not impossible to implement for reliable XC functionals such as the Meta-GGA DFT due to their complicated mathematical forms. In this work, we have developed a general and simple methodology called the “Approximate Self-Consistent Potential” (ASCP) that allows scientists to compute the corresponding derivatives without the knowledge of a closed form. The method has been implemented in two of the most popular quantum chemistry packages and tested in the calculations of absolute energies and atomization energies of 20 different molecules. The results indicate that the method is robust and sufficiently general.

**I. Gonzalez, C. Gonzalez, V.V. Karasiev, E. Ludena, A. Hernandez, "Basis Set Dependent SC Alpha Exchange-Only and Exchange-Correlation Calculations" J. Chem. Phys. 118, 8161, 2003.**

*Quantifying Uncertainty in Prediction of Enthalpies of Formation and Vibrational Frequencies.* It is well recognized that a complete expression of experimental measurement includes both the central value and its associated uncertainty. The central value alone is incomplete. Throughout computational science, the results of models are usually reported without their associated uncertainties, making them incomplete. The term *virtual measurement* refers to a prediction from a computational model together with its associated uncertainty, and emphasizes the analogy with experimental (physical) measurement. Our investigation of the uncertainties associated with predictions from quantum chemistry models are intended to help “virtual measurements” supplant “calculated results.” The results from *ab initio* calculations differ from experiment by systematic biases, not random errors. For a given *ab initio* calculation corrections can be applied to compensate for this bias and an uncertainty can be determined from the distribution of corrections for a class of molecules. We use the experimental and *ab initio* data contained in the NIST Computational Chemistry Comparison and Benchmark Database (CCCBDB), which contains data for over 600 molecules, over 100



<http://srdata.nist.gov/cccbdb/>

calculations for each molecule, and over 4000 vibrational frequencies. We combine this data with analysis based upon the *Guide to the Expression of Uncertainty in Measurement*, published by the International Organization for Standardization (ISO), to determine biases and uncertainties. We have applied this to enthalpies of formation and vibrational frequencies. This allows us to upgrade these “calculated values” into “virtual measurements”. For the enthalpies of formation we found the class of molecules to be critical in determining the bias and uncertainties. For the vibrational frequencies we found the uncertainties to be orders of magnitude larger than previously believed. We plan to continue extracting such information from the CCCBDB to determine the uncertainties for other calculated properties as well.

calculations for each molecule, and over 4000 vibrational frequencies. We combine this data with analysis based upon the *Guide to the Expression of Uncertainty in Measurement*, published by the International Organization for Standardization (ISO), to determine biases and uncertainties. We have applied

**K.K. Irikura, R.D. Johnson III, R.N. Kacker, "Uncertainty Associated with Virtual Measurements from Computational Quantum Chemistry Models," Metrologia 41, 369-375, 2004.**

*Improvements to the Accuracy of Density Functional Theory.* In order to solve the problems exhibited by the “exchange” functionals we proceeded to develop a rigorous and efficient method (the SC- $\alpha$  method) for the calculation of the exact “exchange” functional. This method has been implemented in various popular quantum chemistry software packages and it has been tested in more than 40 molecules; the results have shown the possibility of computing the exact “exchange” functional in complicated polyatomic systems at the computational cost comparable to the inexpensive Hartree-Fock formalism. Encouraged by these results, we focused in procedures that could improve the “exchange-correlation” functional as a whole. One of the most difficult steps in the generation of “exchange-correlation” functionals (XC) that do not require empirical parameters (and therefore, making it more general that the semi-empirical “hybrid-GGA” functionals) is the development