



Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

journal homepage: www.elsevier.com/locate/saa

Can the electronegativity equalization method predict spectroscopic properties?

T. Verstraelen^{a,*}, P. Bultinck^b^a Center for Molecular Modeling (CMM), Ghent University, 9000 Ghent, Belgium^b Ghent University, Department of Inorganic and Physical Chemistry, Krijgslaan 281 (S3), 9000 Ghent, Belgium

HIGHLIGHTS

- IR spectroscopic quantities derived from the electronegativity equalization method.
- EEM cannot reproduce DFT reference data for these spectroscopic quantities.
- Impact on development of polarizable and reactive force-field models is discussed.

GRAPHICAL ABSTRACT

Quantity		Optimal RRMSE of the EEM model
Atomic charge	$q_{gs} = \bar{q}_{gs} + \chi_{mol} \bar{\eta}^{-1} d$	20% ✓
Molecular dipole	$D_j = \sum_{i=1}^N R_{ij} q_{gs,i}$	30% ✓
Dipole derivatives	$\frac{\partial D_j}{\partial \alpha} = \sum_{i=1}^N \delta_{\alpha, R_{ij}} q_i + R_{ij} \frac{\partial q_{gs,i}}{\partial \alpha}$	72% ✗
Cartesian Hessian	$\frac{\partial^2 E_{gs}}{\partial \alpha \partial \beta} = \frac{\partial q_{gs}^T}{\partial \beta} \frac{\partial \bar{\eta}}{\partial \alpha} q_{gs} + \frac{1}{2} q_{gs}^T \frac{\partial^2 \bar{\eta}}{\partial \alpha \partial \beta} q_{gs}$	99% ✗
Long-range part (>5Å) of the Cartesian Hessian		91% ✗

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Electronegativity equalization
Hirshfeld-I
Dipole moment
Infrared spectroscopy
Dipole derivatives
Molecular mechanics

ABSTRACT

The electronegativity equalization method is classically used as a method allowing the fast generation of atomic charges using a set of calibrated parameters and provided knowledge of the molecular structure. Recently, it has started being used for the calculation of other reactivity descriptors and for the development of polarizable and reactive force fields. For such applications, it is of interest to know whether the method, through the inclusion of the molecular geometry in the Taylor expansion of the energy, would also allow sufficiently accurate predictions of spectroscopic data. In this work, relevant quantities for IR spectroscopy are considered, namely the dipole derivatives and the Cartesian Hessian. Despite careful calibration of parameters for this specific task, it is shown that the current models yield insufficiently accurate results.

© 2013 Elsevier B.V. All rights reserved.

Introduction

Electronegativity equalization is a pervasive concept in chemistry that is fundamental to our understanding of molecular charge distributions [1,2]. One of its modern incarnations is the Electronegativity Equalization Method (EEM), which allows a rapid computation of *ab initio* quality atomic charges from just the molecular geometry and a set of atomic electronegativity and hardness parameters [3–5]. Over the past 27 years, EEM strongly influenced related areas, such as in-silico screening [6], chemical reactivity

descriptors [7–9] and empirical models for molecular potential energy surfaces, hereafter referred to as (polarizable and possibly reactive) force fields [10–16]. Extensive calibration studies have demonstrated the accuracy of atomic charges and the parameter transferability of EEM [17–22,14,23] and its recent generalizations, the Split-Charge Equilibration (SQE) [15,24–26] and Atom-Condensed Kohn-Sham approximated to second order (ACKS2) [16].

EEM is essentially a semi-empirical form of density functional theory (DFT) [27,28], in which the molecular electron density is modeled with a minimal set of variables (the atomic charges) and the energy of a system with a specific molecular geometry is expanded in a second order Taylor series of the charges. From that perspective, one may be tempted to derive other quantities from

* Corresponding author. Tel.: +32 9 264 65 56.

E-mail address: Toon.Verstraelen@UGent.be (T. Verstraelen).

EEM beyond just the atomic charges of the ground state [7–9]. In this work, we explore the ability of EEM to reproduce key quantities in the field of infrared spectroscopy, namely the Cartesian Hessian (i.e. force constants or second order derivative of the energy toward nuclear displacements) and the molecular dipole derivatives (first order derivative of the molecular dipole moment toward nuclear displacements). Such insights are essential for the development of force-field models and their applications in which IR spectra are derived from molecular dynamics simulations. It should be noted that such attempts have been made before by Mortier and co-workers [29,30] although limited to mainly vibrations involving bonded atoms. Here the ambition is to go beyond this and consider the entire vibrational structure. For comparison, similar assessments for traditional fixed-charge models are also considered.

The paper is structured as follows. Analytic expressions for the key spectroscopic quantities are derived in Section 2. Section 3 describes all computational aspects related to the model assessment, i.e. the test set of molecular reference computations, the cost functions used to calibrate EEM parameters and the calibration algorithm. All results are presented in Section 4 and discussed in Section 5. Finally, Section 6 summarizes the main conclusions.

Key spectroscopic quantities in EEM

In this section, we start from the EEM energy expression. Justifications and a detailed description of the underlying physics can be found elsewhere [4,12,31,16]. The molecular energy for a system with N atoms is approximated to second order in the atomic charges,

$$E(\{q\}) = \boldsymbol{\chi}^T \mathbf{q} + \frac{1}{2} \mathbf{q}^T \bar{\boldsymbol{\eta}} \mathbf{q}, \quad (1)$$

where \mathbf{q} is a column vector with N atomic charges, $\boldsymbol{\chi}$ is a column vector with N atomic electronegativity parameters and $\bar{\boldsymbol{\eta}}$ is a symmetric N by N matrix containing all second order coefficients, known as the hardness matrix. The diagonal elements of $\bar{\boldsymbol{\eta}}$ are atomic hardness parameters, while the off-diagonal elements correspond to electrostatic interactions between the atoms. In line with previous work, the electrostatics are based on Gaussian charge distributions, whose widths are derived from covalent radii [24,25]. Note that our atomic hardness parameters do not include the factor 1/2 of the Taylor series as in Parr's operational definition of the atomic hardness [32].

The EEM ground state charges are found by minimizing this energy with a constraint on the total charge q_{mol} ,

$$E_{\text{gs}} = E(\{q\}_{\text{gs}}) = \min_{\{q\}, \mathbf{d}^T \mathbf{q} = q_{\text{mol}}} E(\{q\}), \quad (2)$$

where \mathbf{d} is a column vector filled with N times the number 1. This ground state is defined as the stationary point of the following Lagrangian:

$$L = E - \chi_{\text{mol}} (\mathbf{d}^T \mathbf{q} - q_{\text{mol}}). \quad (3)$$

To facilitate the remainder of the derivations, we introduce the vector with *unconstrained* ground state charges:

$$\tilde{\mathbf{q}}_{\text{gs}} = -\bar{\boldsymbol{\eta}}^{-1} \boldsymbol{\chi} \quad (4)$$

Using this notation, the stationary point of the Lagrangian takes the following form:

$$\chi_{\text{mol}} = \frac{q_{\text{mol}} - \mathbf{d}^T \tilde{\mathbf{q}}_{\text{gs}}}{\mathbf{d}^T \bar{\boldsymbol{\eta}}^{-1} \mathbf{d}} \quad \text{and} \quad \mathbf{q}_{\text{gs}} = \tilde{\mathbf{q}}_{\text{gs}} + \chi_{\text{mol}} \bar{\boldsymbol{\eta}}^{-1} \mathbf{d} \quad (5)$$

We will now treat the derivatives of the energy and the charges towards nuclear displacements. For the sake of compactness, nuclear

Cartesian coordinates are labeled with Greek letters α and β instead of explicitly writing R_{ij} . (At most two are needed as we consider at most second order derivatives. Note that R_{ij} is component j of the Cartesian coordinates of atom i and *not* an interatomic distance between two atoms i and j .) In EEM, the electronegativity parameters are geometry-independent; such derivatives are not considered below. The only geometry-dependence in the model is found in the off-diagonal elements of $\bar{\boldsymbol{\eta}}$.

The forces acting on the atoms consist of first-order derivatives of the energy of the following form:

$$\frac{\partial E_{\text{gs}}}{\partial \alpha} = \underbrace{(\boldsymbol{\chi}^T + \mathbf{q}_{\text{gs}}^T \bar{\boldsymbol{\eta}})}_{=0} \frac{\partial \mathbf{q}_{\text{gs}}}{\partial \alpha} + \frac{1}{2} \mathbf{q}_{\text{gs}}^T \frac{\partial \bar{\boldsymbol{\eta}}}{\partial \alpha} \mathbf{q}_{\text{gs}} \quad (6)$$

In analogy with the Hellmann-Feynman theorem in quantum mechanics, the first term is zero because the charges are obtained from a variational principle. However, in the derivation of the force constants, the response of the charges does appear. Therefore, let us first derive the *charge response* and subsequently consider the force constants. The response of the charges with respect to a change in geometry can be found after careful application of the chain rule to Eq. (5):

$$\frac{\partial \mathbf{q}_{\text{gs}}}{\partial \alpha} = \frac{\partial \tilde{\mathbf{q}}_{\text{gs}}}{\partial \alpha} + \frac{\partial \chi_{\text{mol}}}{\partial \alpha} \bar{\boldsymbol{\eta}}^{-1} \mathbf{d} + \chi_{\text{mol}} \frac{\partial \bar{\boldsymbol{\eta}}^{-1}}{\partial \alpha} \mathbf{d} \quad (7)$$

with

$$\frac{\partial \bar{\boldsymbol{\eta}}^{-1}}{\partial \alpha} = -\bar{\boldsymbol{\eta}}^{-1} \frac{\partial \bar{\boldsymbol{\eta}}}{\partial \alpha} \bar{\boldsymbol{\eta}}^{-1}, \quad (8)$$

$$\frac{\partial \tilde{\mathbf{q}}_{\text{gs}}}{\partial \alpha} = -\bar{\boldsymbol{\eta}}^{-1} \frac{\partial \bar{\boldsymbol{\eta}}}{\partial \alpha} \tilde{\mathbf{q}}_{\text{gs}}, \quad (9)$$

$$\frac{\partial \chi_{\text{mol}}}{\partial \alpha} = \frac{q_{\text{mol}} - \mathbf{d}^T \tilde{\mathbf{q}}_{\text{gs}}}{(\mathbf{d}^T \bar{\boldsymbol{\eta}}^{-1} \mathbf{d})^2} (\mathbf{d}^T \frac{\partial \bar{\boldsymbol{\eta}}^{-1}}{\partial \alpha} \mathbf{d}) - \frac{1}{\mathbf{d}^T \bar{\boldsymbol{\eta}}^{-1} \mathbf{d}} \frac{\partial \mathbf{d}^T \tilde{\mathbf{q}}_{\text{gs}}}{\partial \alpha}. \quad (10)$$

The Cartesian Hessian can be found through the application of the chain rule to Eq. (6):

$$\frac{\partial^2 E_{\text{gs}}}{\partial \alpha \partial \beta} = \frac{\partial \mathbf{q}_{\text{gs}}^T}{\partial \beta} \frac{\partial \bar{\boldsymbol{\eta}}}{\partial \alpha} \mathbf{q}_{\text{gs}} + \frac{1}{2} \mathbf{q}_{\text{gs}}^T \frac{\partial^2 \bar{\boldsymbol{\eta}}}{\partial \alpha \partial \beta} \mathbf{q}_{\text{gs}} \quad (11)$$

In EEM, the molecule is modeled as a superposition of spherical atomic charge distributions. Hence, the Cartesian components of the molecular dipole moment simply become

$$D_j = \sum_{i=1}^N R_{ij} q_{\text{gs},i}, \quad (12)$$

where q_i is the charge of atom i . The dipole derivatives are obtained by differentiating D_j once with respect to a nuclear displacement:

$$\frac{\partial D_j}{\partial \alpha} = \sum_{i=1}^N \delta_{\alpha, R_{ij}} q_i + R_{ij} \frac{\partial q_{\text{gs},i}}{\partial \alpha}. \quad (13)$$

In this paper, the EEM model will also be compared to a model with fixed atomic charges. Such a comparison is relevant as many force-field models treat the atomic charges as geometry-independent parameters. The above results for the Cartesian Hessian and the dipole derivatives are still valid for fixed-charge models, except that the terms containing derivatives of the charges can be dropped.

Assessment protocol

To test the accuracy of EEM for the key spectroscopic quantities of interest, DFT reference computations were carried out for a set

of 166 organic drug-like molecules [19,20,22]. The molecular geometries were optimized at the PBE0/6-311+G(d,p) level of theory with Gaussian09 [33], followed by computation of the dipole moment, the Cartesian Hessian and the dipole derivatives, all at the same level of theory. Finally, the Hirshfeld-I atomic charges [34] are computed with Horton [35]. Previous studies have shown that Hirshfeld-I and Mulliken charges are suitable reference data for the calibration of EEM parameters [22–24]. However, in order to properly describe negatively charged moieties, one needs to include diffuse functions in the basis set, which leads to ill-defined Mulliken charges. In contrast, Hirshfeld-I charges are known to be robust with respect to the choice of basis set and are therefore used as reference data in this work [36]. A text file with all relevant results of the reference computations is provided as Supplementary data.

Five different sets of EEM parameters were calibrated, each to reproduce one class of the DFT reference data:

- *AC*: Hirshfeld-I atomic charges.
- *MD*: Molecular dipole moment.
- *MDD*: Molecular dipole derivatives.
- *FH*: Full Hessian.
- *LRH*: Long-range Hessian, i.e. only considering those matrix elements of Cartesian Hessian that correspond to pairs of atoms that are at least 5 Å apart. At such larger distances, exchange-correlation effects and higher order terms should become negligible in DFT. One may expect that the EEM approximation, which is based classical electrostatics, becomes more reliable in this limit.

Each calibration was carried out by minimizing a corresponding least-squares cost function of the following form:

$$C_Z(\mathbf{p}) = \sum_{m=1}^M \sum_{n=1}^{N_{z,m}} (x_{\text{EEM},n}(\mathbf{p}) - x_{\text{DFT},n})^2 \quad (14)$$

where Z refers to the class of reference data (*AC*, *MD*, *MDD*, *FH*, *LRH*), \mathbf{p} is the vector of independent parameters. The first sum runs over all molecules (M in total) and the second sum runs over all corresponding pairs of EEM predictions and reference data in class Z for molecule m . Each term is a squared error between a reference result and the corresponding EEM prediction. The vector of parameters consists of atomic electronegativity and hardness parameters for each element found in the set of molecules (H, C, N, O, F and Cl). Since the electronegativity parameters can only be calibrated up to an unknown constant, the electronegativity of Hydrogen is used as a reference [25]. Hence, there are 11 independent parameters. The optimal parameters are denoted as \mathbf{p}_Z , where Z can be any of *AC*, *MD*, *MDD*, *FH*, *LRH*. Note that only the calibration *AC* relies on Hirshfeld-I charges while all other calibrations are independent of the method used to compute the charges. The implementation of all spectroscopic quantities was validated numerically with finite-difference approximations.

The covariance matrix adaption evolutionary strategy (CMA-ES) [37] was used to numerically minimize the five cost functions. Our choice for CMA-ES is motivated by several of its advantages. CMA-ES only requires that cost function values can be computed, i.e. no analytic gradients of the cost function must be implemented. Furthermore, CMA-ES has the robustness and (to some extent) the efficiency of a quasi-Newton method, i.e. it constructs a second order model of the function during minimization, without resorting to numerically unstable finite difference approximations. In line with previous work, lower-bounds were imposed on the hardness parameters to guarantee that the hardness matrix remains positive definite [24–26]. All EEM

predictions by the five sets of calibrated parameters are provided as Supplementary data.

To assess EEM's ability to reproduce the five different quantities listed above, relative root-mean-square errors (RRMSEs) are computed for all combinations of optimal parameters and reference data:

$$\text{RRMSE}_Z(\mathbf{p}) = \sqrt{\frac{C_Z(\mathbf{p})}{\sum_{m=1}^M \sum_{n=1}^{N_{z,m}} x_{\text{DFT},n}^2}} \quad (15)$$

An RRMSE of 0% corresponds to an exact reproduction of the reference data, while a value of 100% (or larger) indicates that the errors are equal to (or larger than) the reference data. A more detailed statistical analysis (including absolute RMSEs and parity plots) are provided as Supplementary data.

For the cases *MD*, *MDD*, *FH* and *LRH*, the RRMSEs are also computed for a fixed-charge model. In analogy with the EEM approach, a fixed-charge model still describes the molecule as a superposition of spherical Gaussian functions. However, the charges are now fixed at the Hirshfeld-I values and no geometry dependence of the charges is considered.

Results

All calibrations lead to a converged set of parameters shown in Table 1. Only for the calibrations *AC* and *MD*, the parameters have a reasonable order of magnitude, e.g. compared to previous calibrations [22,24] or experimental properties of isolated atoms [32]. The condition number reported in Table 1 is based on the covariance estimate of the CMA-ES algorithm. A high condition number indicates the presence of a manifold of parameter vectors that approximately minimizes the cost function. In such cases, the parameters are not uniquely defined and the optimization landscape reflects many (near) degenerate solutions [25]. Note that most results do not depend on the choice of the method for the population analysis. More specifically, all parameter sets in Table 1, except for the column \mathbf{p}_{AC} , are independent of the method used to compute the atomic charges.

Table 2 lists the RRMSE computed for each combination of a cost function and a set of optimal parameters. This reveals how dependent the cost function is on the target function used in the optimization. Only the *AC* and *MD* calibrations lead to parameters that are sufficiently accurate for the description of atomic charges or molecular dipoles. This is in line with the previously observed good performance of EEM algorithms for Hirshfeld-I atomic charges [24–26] and it confirms that the CMA algorithm is an effective method for the calibration of EEM parameters. The *AC* and *MD* parameter sets do not give a useful approximation of the dipole derivatives nor the Hessian, including when one only considers the long-range part of the Hessian. Parameters calibrated with the *MDD*, *FH* or *LRH* cost function fail to predict any property, i.e. all RRMSEs are close to or above 100%. Note that, as was to be expected, the diagonal elements in Table 2 are lower than the off-diagonal elements on the same row or column. Again, note that all RRMSEs in Table 2, except for the first row and the first column, are independent of the choice of population analysis.

The accuracy of the fixed-charge model for the classes *MD*, *MDD*, *FH* and *LRH* is shown in Table 3. From these results, it is clear that such a fixed-charge model is only capable of reproducing the molecular dipole moment to some accuracy. The errors on the dipole derivatives and the Cartesian Hessian are close to 100%, also when one only considers the long-range part of the Hessian.

Table 1

Calibrated EEM Parameters for the five classes of reference data. All values of the different parameters are given in eV. The condition number is derived from the covariance matrix in the CMA-ES algorithm.

	P_{AC}	P_{MD}	P_{MDD}	P_{FH}	P_{LRH}
χ_{C-H}	0.87	14.46	0.97	143.73	11.64
χ_{N-H}	1.60	12.88	14.71	143.77	31.49
χ_{O-H}	3.31	12.58	431.61	392.51	631.43
χ_{F-H}	4.89	10.65	367.47	320.13	394.14
χ_{Cl-H}	1.20	10.47	206.52	233.09	303.12
η_H	15.47	37.87	128.79	191.54	182.62
η_C	11.40	13.29	16.19	11.71	18.84
η_N	11.57	16.56	27.38	11.70	22.12
η_O	14.14	15.32	523.84	140.37	685.59
η_F	21.61	21.52	515.27	118.61	284.50
η_{Cl}	8.93	21.25	444.23	40.09	252.03
Condition number	1.5×10^4	5.1	1.2×10^4	73	23

Table 2

RRMSE values (%) for all possible pairs of optimal parameters and classes of reference data.

	P_{AC}	P_{MD}	P_{MDD}	P_{FH}	P_{LRH}
RRMSE _{AC}	20	158	92	326	135
RRMSE _{MD}	49	30	237	657	458
RRMSE _{MDD}	92	133	72	223	84
RRMSE _{FH}	100	100	100	99	100
RRMSE _{LRH}	100	103	93	286	91

Table 3

RRMSE values (%) for the fixed-charge model.

RRMSE _{MD}	19
RRMSE _{MDD}	77
RRMSE _{FH}	99
RRMSE _{LRH}	95

Discussion

The diagonal elements of Table 2 show how well the EEM model can reproduce a class of reference data. As extensively demonstrated in the literature, we observe that it is relatively easy to obtain EEM parameters that reproduce the Hirshfeld-I atomic charges [24–26]. A similar result is found for the calibration of parameters aimed at reproducing molecular dipole moments, albeit with a larger RRMSE. When considering the two EEM calibrations with reasonable parameters (AC and MD), the error in the reproduction of the dipole moment (RRMSE_{MD} in Table 2) is worse than that of the fixed charge model (see Table 3). This was to be expected: the fixed-charge model contains a huge number of non-transferable parameters, i.e. one charge for every atom in the set, as compared to the 11 transferable EEM parameters. The Hirshfeld-I charges are known to reproduce the dipole moment of organic molecules [38,39]. However, this attractive feature of Hirshfeld-I can not be fully reproduced by an EEM model based on Hirshfeld-I charges (or any other class of reference data). This corresponds to a known pathology of EEM. Namely, due to the metallic polarizability scaling in EEM, [40] molecular dipole moments are underestimated on average, especially in the limit of larger molecules. [26] This trend is also confirmed in Fig. S9 (and to lesser extent in Fig. S20) in the Supplementary data, showing that EEM tends to underestimate the molecular dipole moment.

The poor description of the dipole derivatives is somewhat disappointing. Some authors previously suggested a good mapping between EEM and the charge-sensitivity analysis (CSA, i.e. a complete set of second order derivatives of the electronic energy towards AIM charges and nuclear coordinates) [41,29,30,42]. This would imply a reasonable reproduction of dipole derivatives with

EEM, which we do not observe. Compared to a fixed-charge model (see Table 3), EEM is only a marginal improvement. The parity plots of the dipole derivatives for each model (Figs. S10, S21, S32, S43 and S54 in the Supplementary data) show that the large RRMSEs in row 3 of Table 2 are representative, i.e. not caused by a few outliers. These results have two important consequences. First, when EEM or fixed charges are used in the simulation of IR spectra, we expect similarly large errors on peak heights, i.e. IR intensities. Second, the good geometry dependence of atomic charges in EEM [39] is not valid for small displacements of the nuclei. This is reminiscent of the limitations of EEM for the description of the dipole polarizability [40,15,43,24,16] and suggests that EEM fails to describe any type of electronic linear response. An interesting topic of future research would therefore be the reproduction of dipole derivatives with recent generalizations of EEM that fix the dipole polarizability weaknesses, e.g. SQE [15] or ACKS2 [16].

The poor reproduction of the Hessian is in line with expectations. Several force-field models use the EEM energy as one of the contributions to the total energy [13,14]. Obviously, the EEM energy alone is not sufficient to describe the curvature of the PES. Surprisingly, this observation still holds when only the long-range parts of the Hessian are considered. Again, the parity plots for the Hessian matrix elements (Figs. S11, S12, S22, S23, S33, S34, S44, S45, S55, S56 in the Supplementary data) confirm that the large RRMSEs in row 4 and 5 of Table 2 are not due to outliers. Although electrostatic interactions are the main long-range interactions in a force-field model, the EEM long-range contributions to the Hessian do not match the DFT reference data. Again, compared to a fixed-charge model (see Table 3), EEM is only a marginal improvement.

The long range Hessian model performs somewhat better than the Full Hessian model which is not unexpected. As EEM entails an energy expansion in terms of atomic charges only and considers only a rather simplified model for the electrostatic interaction, many aspects of the pair-wise interaction, such as exchange effects, are not taken properly into account. Such exchange effects grow smaller with increasing internuclear distance, as do higher order contributions to the electrostatic interaction. This explains why LRH is better than FH. On the other hand, in most CSA work mentioned above, the bonded, and thus shorter interactions were emphasized. Note, however, that Baekelandt et al. [30] used a force field model that corrects for covalent character of the bond.

Summary

We have shown that the bare EEM model can not describe (let alone predict) the key molecular properties related to IR spectroscopy: the dipole derivatives and the Cartesian Hessian. The potential improvements of EEM over a simple fixed-charge model are negligible. Although EEM (or a fixed charge model) is

often used as one of many contributions in empirical models for the potential energy surface, the poor performance of the method implies that computer simulations of IR spectra require a more advanced model of the response of atomic charges to a displacement of the nuclei. This complements the well-known limitations of EEM for the description of the dipole polarizability tensor and suggests that EEM fails to reproduce (even qualitatively) any linear response property. As an example of such a more advanced method, the ACKS2 model, which also includes a term for the electronic kinetic energy, may improve the accuracy the dipole derivatives.

Acknowledgments

T.V. is a post-doctoral researcher funded by the Foundation of Scientific Research – Flanders (FWO). Funding was also received from the Research Board of Ghent University (BOF), and BELSPO in the frame of IAP/7/05. P.B. acknowledges the Scientific Research Foundation – Flanders (FWO-Vlaanderen) and the Research Board of Ghent University (BOF) for continuous support.

Supplementary material

Supplementary data reflecting the molecular data set and raw EEM results as well as more in-depth statistical analysis associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.saa.2013.10.124>.

References

- [1] R.T. Sanderson, An interpretation of bond lengths and a classification of bonds, *Science* 114 (1951) 670–672.
- [2] R.T. Sanderson, *Polar Covalence*, Academic Press, New York, 1983.
- [3] W.J. Mortier, K. Vangenechten, J. Gasteiger, Electronegativity equalization: application and parametrization, *J. Am. Chem. Soc.* 107 (1985) 829–835.
- [4] W.J. Mortier, S.K. Ghosh, S. Shankar, Electronegativity-equalization method for the calculation of atomic charges in molecules, *J. Am. Chem. Soc.* 108 (1986) 4315–4320.
- [5] W.J. Mortier, Electronegativity equalization and its applications, *Struct. Bond.* 66 (1987) 125–143.
- [6] J.S. Puranen, M.J. Vainio, M.S. Johnson, Accurate conformation-dependent molecular electrostatic potentials for high-throughput in silico drug discovery, *J. Comput. Chem.* 31 (2010) 1722–1732.
- [7] B.G. Baekelandt, W.J. Mortier, J.L. lievens, R.A. Schoonheydt, Probing the reactivity of different sites within a molecule or solid by direct computation of the atomic sensitivities via an extension of the electronegativity equalization method, *J. Am. Chem. Soc.* 113 (1991) 6730–6734.
- [8] P. Bultinck, R. Carbó-Dorca, Algebraic relationships between conceptual DFT quantities and the electronegativity equalization hardness matrix, *Chem. Phys. Lett.* 364 (2002) 357–362.
- [9] P. Bultinck, W. Langenaeker, R. Carbó-Dorca, J.P. Tollenaere, Fast calculation of quantum chemical molecular descriptors from the electronegativity equalization method, *J. Chem. Inf. Comput. Sci.* 43 (2003) 422–428.
- [10] A.K. Rappe, W.A. Goddard III, Charge equilibration for molecular dynamics simulations, *J. Phys. Chem.* 95 (1991) 3358–3363.
- [11] S.W. Rick, S.J. Stuart, B.J. Berne, Dynamical fluctuating charge force fields: application to liquid water, *J. Chem. Phys.* 101 (1994) 6141–6156.
- [12] D.M. York, W. Yang, A chemical potential equalization method for molecular simulations, *J. Chem. Phys.* 104 (1996) 159–172.
- [13] A.C.T. van Duin, A. Strachan, S. Stewman, Q. Zhang, X. Xu, W.A. Goddard III, ReaxFF_{SiO} reactive force field for silicon and silicon oxide systems, *J. Phys. Chem. A* 107 (2003) 3803–3811.
- [14] S. Patel, C.L. Brooks, A nonadditive methanol force field: bulk liquid and liquid-vapor interfacial properties via molecular dynamics simulations using a fluctuating charge model, *J. Chem. Phys.* 122 (2005) 024508.
- [15] R.A. Nistor, J.G. Polihronov, M.H. Müser, N.J. Mosey, A generalization of the charge equilibration method for nonmetallic materials, *J. Chem. Phys.* 125 (2006) 094108.
- [16] T. Verstraelen, P.W. Ayers, V. Van Speybroeck, M. Waroquier, ACKS2: Atom-condensed Kohn-Sham DFT approximated to second order, *J. Chem. Phys.* 138 (2013) 074108.
- [17] K.A. Van Genechten, W.J. Mortier, P. Geerlings, Intrinsic framework electronegativity: a novel concept in solid state chemistry, *J. Chem. Phys.* 86 (1987) 5063–5071.
- [18] S.L. Njo, J. Fan, B. van de Graaf, Extending and simplifying the electronegativity equalization method, *J. Mol. Catal. A* 134 (1998) 79–88.
- [19] P. Bultinck, W. Langenaeker, P. Lahorte, F. De Proft, P. Geerlings, M. Waroquier, J.P. Tollenaere, The electronegativity equalization method I: parametrization and validation for atomic charge calculations, *J. Phys. Chem. A* 106 (2002) 7887–7894.
- [20] P. Bultinck, W. Langenaeker, P. Lahorte, F. De Proft, P. Geerlings, C. Van Alsenoy, J.P. Tollenaere, The electronegativity equalization method II: applicability of different atomic charge schemes, *J. Phys. Chem. A* 106 (2002) 7895–7901.
- [21] K.S. Smirnov, D. Bougeard, Including the polarization in simulations of hydrated aluminosilicates. Model and application to water in silicalite, *Chem. Phys.* 292 (2003) 53–70.
- [22] P. Bultinck, R. Vanholme, P.L.A. Popelier, F. De Proft, P. Geerlings, High-speed calculation of AIM charges through the electronegativity equalization method, *J. Phys. Chem. A* 108 (2004) 10359–10366.
- [23] C.-M. Ionescu, S. Geidl, R. Svobodová Vařeková, J. Koča, Rapid calculation of accurate atomic charges for proteins via the electronegativity equalization method, *J. Chem. Inf. Model.* 53 (2013) 2548–2558.
- [24] T. Verstraelen, V. Van Speybroeck, M. Waroquier, The electronegativity equalization method and the split charge equilibration applied to organic systems: parametrization, validation, and comparison, *J. Chem. Phys.* 131 (2009) 044127.
- [25] T. Verstraelen, P. Bultinck, V. Van Speybroeck, P.W. Ayers, D. Van Neck, M. Waroquier, The significance of parameters in charge equilibration models, *J. Chem. Theory Comput.* 7 (2011) 1750–1764.
- [26] T. Verstraelen, S.V. Sukhomlinov, V. Van Speybroeck, M. Waroquier, K.S. Smirnov, Computation of charge distribution and electrostatic potential in silicates with the use of chemical potential equalization models, *J. Phys. Chem. C* 116 (2012) 490–504.
- [27] R. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford Science Publications, Oxford, UK, 1989.
- [28] W. Kohn, A.D. Becke, R.G. Parr, Density functional theory of electronic structure, *J. Phys. Chem.* 100 (1996) 12974–12980.
- [29] G.O.A. Janssens, B.G. Baekelandt, H. Toufar, W.J. Mortier, R.A. Schoonheydt, On the relation between electronic and nuclear vibrations via the response matrix derived from semiempirical density functional-based sensitivity analysis, *Int. J. Quant. Chem.* 56 (4) (1995) 317–326.
- [30] B.G. Baekelandt, G.O.A. Janssens, H. Toufar, W.J. Mortier, R.A. Schoonheydt, R.F. Nalewajski, Mapping between electron population and vibrational normal modes within the charge sensitivity analysis, *J. Phys. Chem.* 99 (1995) 9784–9794.
- [31] P. Itskowitz, M.L. Berkowitz, Chemical potential equalization principle: direct approach from density functional theory, *J. Phys. Chem. A* 101 (1997) 5687–5691.
- [32] R.G. Parr, R.G. Pearson, Absolute hardness: companion parameter to absolute electronegativity, *J. Am. Chem. Soc.* 105 (1983) 7512–7516.
- [33] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, *Gaussian 09 Revision A.02*, Gaussian Inc.: Wallingford CT, 2009.
- [34] P. Bultinck, C. Van Alsenoy, P.W. Ayers, R. Carbó-Dorca, Critical analysis and extension of the Hirshfeld atoms in molecules, *J. Chem. Phys.* 126 (2007) 144111, <http://dx.doi.org/10.1063/1.2715563>.
- [35] T. Verstraelen, S. Vandenbrande, M. Chan, F.H. Zadeh, C. González, P.A. Limacher, A. Malek; Horton 1.2.0, 2013 <<http://theochem.github.com/horton/>>.
- [36] P. Bultinck, P.W. Ayers, S. Fias, K. Tiels, C. Van Alsenoy, Uniqueness and basis set dependence of iterative Hirshfeld charges, *Chem. Phys. Lett.* 444 (2007) 205–208.
- [37] N. Hansen, The CMA evolution strategy: A comparing review, in: J. Lozano, P. Larra naga, I. n. Inza, E. Bengoetxea (Eds.), *Towards a New Evolutionary Computation, Studies in Fuzziness and Soft Computing*, vol. 192, Springer Berlin, Heidelberg, 2006, pp. 75–102 (Chapter 4).
- [38] S. Van Damme, P. Bultinck, S. Fias, Electrostatic potentials from self-consistent hirshfeld atomic charges, *J. Chem. Theory Comput.* 5 (2009) 334–340.
- [39] T. Verstraelen, E. Pauwels, F. De Proft, V. Van Speybroeck, P. Geerlings, M. Waroquier, Assessment of atomic charge models for gas-phase computations on polypeptides, *J. Chem. Theory Comput.* 8 (2012) 661–676.
- [40] G.L. Warren, J.E. Davis, S. Patel, Origin and control of superlinear polarizability scaling in chemical potential equalization methods, *J. Chem. Phys.* 128 (2008) 144110.
- [41] R.F. Nalewajski, The hardness based molecular charge sensitivities and their use in the theory of chemical reactivity, in: K.D. Sen (Ed.), *Chemical Hardness of Structure and Bonding*, 80, Springer, Berlin, Heidelberg, 1993, pp. 115–186.
- [42] R. Heidler, G.O.A. Janssens, W.J. Mortier, R.A. Schoonheydt, Charge sensitivity analysis of the interaction of pyrrole with basic FAU-type zeolites using the electronegativity equalization method, *Microporous Mater.* 12 (1997) 1–11.
- [43] R.A. Nistor, M.H. Müser, Dielectric properties of solids in the regular and split-charge equilibration formalisms, *Phys. Rev. B* 79 (2009) 104303.