RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

What does the *HOMA* index really measure?

Sławomir Ostrowski[a] and Jan Cz. Dobrowolski*[a,b]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

The *HOMA* index was generalized to the function of a property of chemical bonds such as, *e.g.*, characteristics in BCPs. The new *HOMA* indices strongly correlate with the classical one for a variety of cyclic hydrocarbons unsaturated or not. Thus, the indices reflect a more general property then aromaticity: the savoricity.

The first geometrical aromaticity index, *A*, was introduced by Julg and Françoise [1] as the variance of the ring perimeter bond lengths:

$$A = 1 - \frac{225}{n}\sum_{(rs)}\left(1-\frac{d_{rs}}{\bar{d}}\right)^2 \qquad (1)$$

where d_{rs} and $\bar{d}$ for CC distances: particular and averaged, *n* is number of bonds in perimeter cycle, *r* and *s* are running indices over the cycle atoms, and 225 is a normalization factor.

The rationale behind the *A* index was the observation that in the fully aromatic benzene molecule the CC bond lengths are identical whereas in the hypothetical cyclohexatriene they must be strongly alternant. However, the so defined index has exhibited a significant drawback: it recognizes "aromaticity" in definitely non-aromatic, but non-alternant, compounds such as cyclohexane. In consequence, it also does not differentiate aromaticity of different aromatic rings such as central coronene and benzene rings. One can conclude that the sole variance of the bond distances in a ring is not sufficient to adequately express geometrical aspect of aromaticity because it does not differentiate the rings of different aromaticity with satisfactory resolution.

The drawback of the Julg and Françoise index was removed by Krygowski and Kruszewski who defined the *HOMA* (*Harmonic Oscillator Model of Aromaticity*) index (2) widely used to express geometrical aspects of aromaticity:

$$HOMA = 1 - \frac{\alpha}{n}\cdot\sum_{i=1}^{n}\left(R_i - R_{opt}\right)^2 = 1 - \frac{\beta}{n}\cdot\sum_{i=1}^{n}\left(1-\frac{d_i}{d_{opt}}\right)^2 \qquad (2)$$

where R_i and R_{opt} stand for distances of *i*-th ring bond in the analyzed structure and the reference optimal bond in benzene (1.388 A), n is the number of CC bonds in the ring, whereas $\alpha = 257.7$ is a normalization factor which guarantees that *HOMA* of aromatic compound approaches 1 and of its Kekulé non-aromatic structure approaches 0, $\beta = \alpha \cdot R_{opt}^2$.

It is surprising that the *HOMA* index, which seems to be very similar to the Julg and Françoise *A* index, is so successful and well resolves compounds of different aromaticity, while *A* fails in this task. The "mystery" of this difference is in disparity of $\bar{d}$ and d_{opt}. The former refers to mean CC distance in the analyzed compounds, while the latter is the CC distance in the reference benzene molecule. The initial shortcoming of *HOMA*, its limitation to only carbocyclic structures, was removed by introducing parameterizations for most of important heteroatoms.[3] The *HOMA* index was also decomposed into *GEO* and *EN* destabilizing components interpreted as geometrical and energetical factors.[4] The *GEO* component is related to Julg and Françoise *A* (normalized) variance index and Krygowski's *et al. Bond Alternation Coefficient, BAC,*[5] a (normalized) standard deviation. Perfect correlations of *EN* with the CC bond energy and other energy parameters was also demonstrated.[4]

Thus, *GEO* is a function of a bond variance, while *EN* can be interpreted as a (normalized) squared distance between a molecule and the aromaticity standard $\alpha\left(R_{opt}-\bar{R}\right)^2$ where $\bar{R}$ is the mean CC distance in an *n*-membered molecule (n≥3) and R_{opt} can be interpreted as a reference CC distance in an appropriate, hypothetical, fully aromatic planar molecule of a regular *n*-gone skeleton which is benzene if *n=6*. Remark, that the (normalized) squared distance in (2) places a greater weight to more distanced points and therefore is used in optimization problems in which directions are less important then distances. Minimization of an expression very similar to (2) is basic for the *K-mean* clustering method[6] introduced to cluster analysis by Steinhaus.[7]

The above features of the *HOMA* index indicate that the only connection of *HOMA* with aromaticity lies in choosing benzene as a reference molecule and in normalization making *HOMA* equal to 1 for benzene and 0 for its hypothetical Keluléne form. Thus, the index can be formally used for the other (cyclic) molecules without loosing its statistical meaning. Moreover, it seems to be correctly defined also for the non-cyclic molecules if the appropriate meaning is assigned to the reference optimal distances. Here, we are showing that *HOMA* discriminates much larger class of molecules than sole aromatic compounds and are demonstrating that it can be generalized to obtain other valuable (aromaticity) indices.

A neural networks algorithm applied to a set of geometric, energetic, and magnetic aromaticity indices allowed for construction of a unified aromaticity index for series of unsaturated organic compounds.[8] However, comparison of aromaticity indices based on different physical properties leads to the conclusion that aromaticity is a multidimensional phenomenon.[9-11] Nevertheless, some of the indices constructed on different physical bases can give quite concordant evaluation of aromaticity. A very good example is total electron energy *(H)*, its kinetic *(G)* and potential *(V)* components, as well as the electron density *(ρ)* in ring critical point *(RCP)* which are strongly linearly correlated with the geometrical *HOMA* aromaticity index and also show fair linear correlations with the magnetic $NICS(1)_{zz}$ aromaticity index.[12]

As pointed out above, the *HOMA* index can be calculated for every ring irrespectively it is aromatic or not. Therefore, we tested whether the Palusiak's and Krygowski's finding[12] can be extended to saturated and singly and doubly unsaturated hydrocarbon rings. Although, we found some tendencies between *HOMA* and electron density characteristics in RCP, the number of "exceptions" and spread of the points were significant.

Therefore, we introduced a new group of indices in which expression (2) is used as formula defining the function of a variable assigned to the bond. In this approach, the classical *HOMA* is a function of bond lengths, *HOMA=HOMA(R),* but the *HOMA*(•) function can act, for example, on electron density ρ^{BCP} in *BCPs*:

$$HOMA(\rho) = 1 - \frac{\alpha_{BCP}}{n}\cdot\sum_{i=1}^{n}\left(\rho_i^{BCP} - \rho_{opt}^{BCP}\right)^2$$

where *HOMA(•)* is a function of a variable given in parentheses, and ρ stand for electron density in *BCP* of *i*-th bond in the analyzed structure and the optimal value in the reference benzene, *n* is the number of CC bonds in the considered structure, whereas α_{BCP} is a normalization factor. The *HOMA* indices for other bond parameters such as potential and kinetic electron energy in bond critical points *(BCP)*, V^{BCP} and K^{BCP}, respectively, can be determined analogously.

It appeared that for cyclic hydrocarbons,[†,‡] the quadratic correlation between the *HOMA(R)* and *HOMA(ρ)* is excellent (Fig. 1) with the correlation coefficient equal to 0.9997 but the correlation coefficient of a linear fit exceeds 0.999 as well. Fig. 1 shows that the smallest values of *HOMA* and *HOMA(ρ)* are for the saturated 4-, 5-, and 6-membered rings in mono- and bicyclic systems. The next group of compounds, with larger values of *HOMAs*, is constituted of the singly-unsaturated rings fused in various ways to other saturated or unsaturated rings. The following group of cycles is consisted of the doubly-unsaturated rings fused in different manners to the other rings. The group of the highest values of *HOMAs* is formed by the aromatic rings condensed in several ways in polycyclic systems (Fig. 1). The correlation has, however, some exceptions: the cyclopropane, benzyne, and benzdiynes molecules. The data for these molecules were not included to establish the correlations and require further analysis and investigation. Thus, aromatic, non-aromatic, and anti-aromatic compounds served so far to construct correlations between *HOMA* and the other aromaticity indices form a congruent group of molecules constituting only a minor part of a more general class of cyclic molecules (Fig. 1).

If the *HOMA(ρ)* and *HOMA(R)* indices reflect a more general feature than aromaticity itself than what a property is actually measured by these parameters? The plot in Fig. 1 shows an increase in the *HOMA* indices with the increase of the cycle size and the number of the double bonds in a cycle. The relationship is slightly modified by the type of condensation. Thus, the strain of

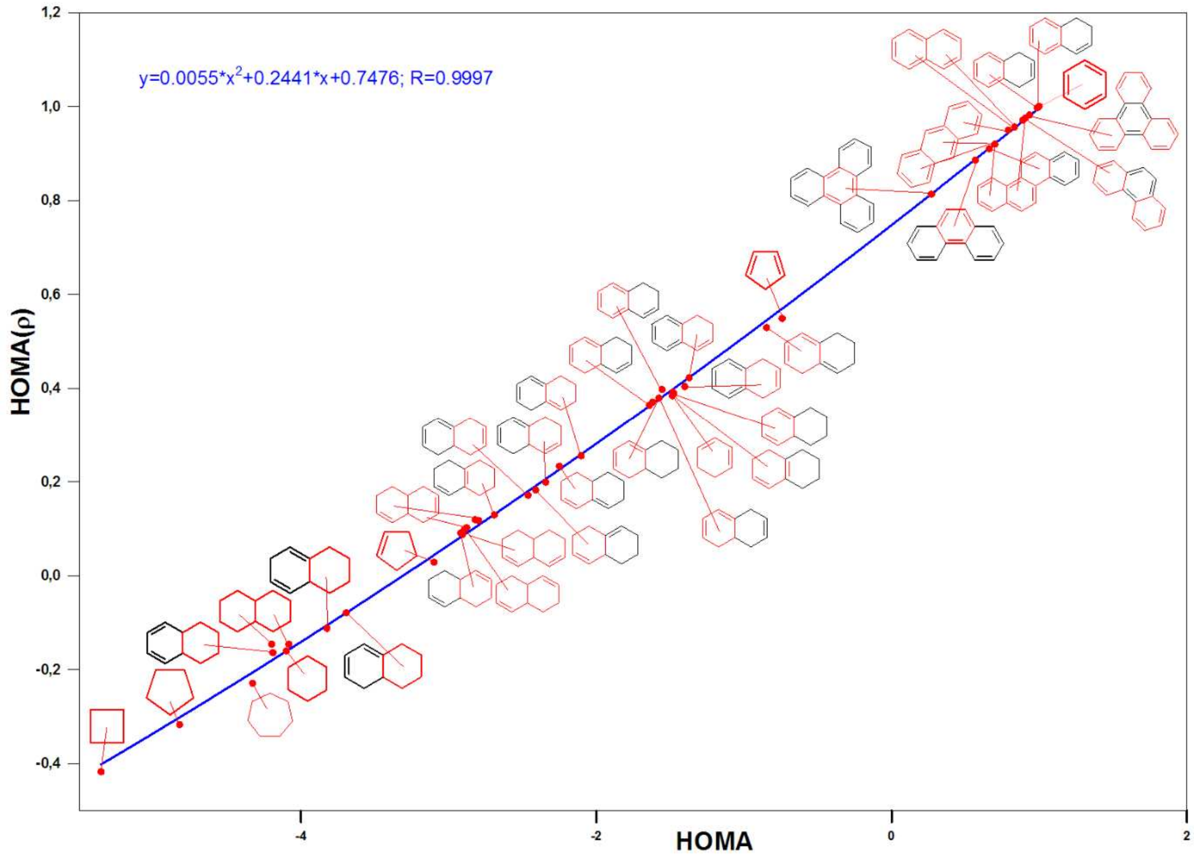

Fig. 1. The linear correlations between the HOMA geometrical index and HOMA(*ρ*) index based on electron densities in bond critical points for series of saturated, unsaturated and aromatic hydrocarbon rings.

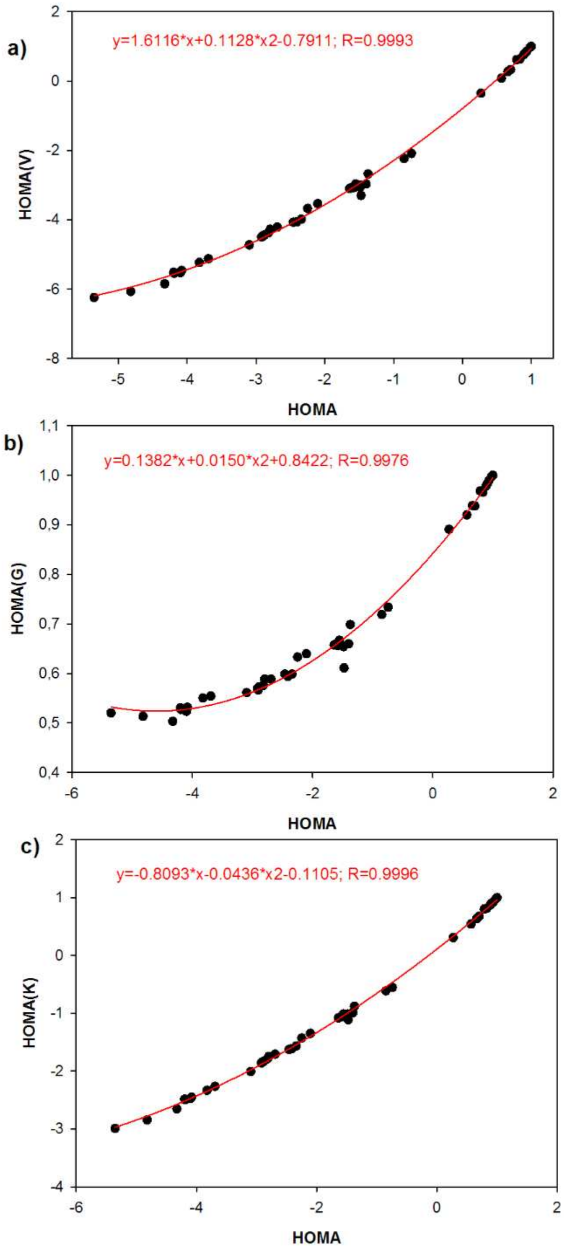

Fig.2 The non-linear correlations between electron density characteristics and *HOMA* index for series of saturated, unsaturated and aromatic hydrocarbon rings: a) $HOMA(V^{BCP})$ potential electron energy, b) $HOMA(G^{BCP})$ - Lagrangian form of kinetic energy, c) $HOMA(K^{BCP})$ - Hamiltonian form of kinetic energy

the cycle, the number of double bonds, the aromaticity and the condensation type play a significant role. The property is more general then aromaticity and neither of the single words constructed by analogy to the term “aromaticity”, *i.e.*, “cyclicity”, “saturaticity” or “unsaturaticity”, and “condensaticity” or “fusicity”, which would describe these properties, seems to be proper or satisfactory to name the feature presented in Fig. 1. Therefore, here, we shall call this property the *savoricity*.

Observe, that also $HOMA(V^{BCP})$ and $HOMA(K^{BCP})$ estimated for the same series of cyclic molecules reflects the savoricity equally well as does the classical *HOMA(R)* index (Fig. 2). The correlations are, however, non-linear (quadratic) but strong with the correlation coefficients exceeding 0.99 (Fig. 2). For the commonly studied aromatic and non-aromatic molecules the correlations can be fitted by a straight line with the correlation coefficients also close to 0.99. Interestingly, very recently Tokatli and Ucun constructed an aromaticity index, which would be called $HOMA(\varepsilon^{BCP})$, where ε^{BCP} is the bond elipticity, if the absolute value function was changed for the second power.[13]

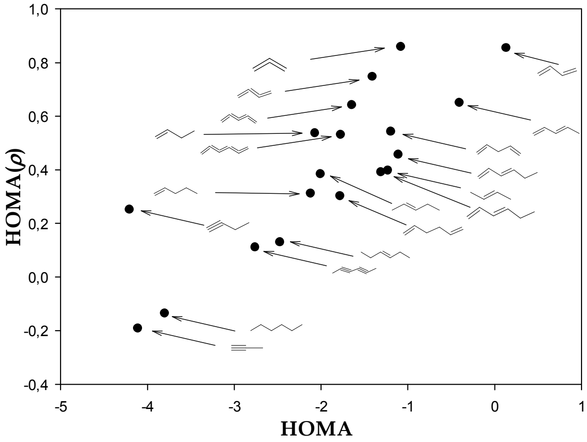

Fig. 3 Plot of the *HOMA(R)* geometrical index against the *HOMA(ρ)* index based on electron densities in bond critical points for series of saturated and unsaturated acyclic hydrocarbons.

Correlations between the empirical structural *HOMA* index and functions defined on the well theoretically grounded electron density characteristics in BCP of a variety of cyclic structures (Figs. 1 and Fig. 2), are an excellent confirmation for the correctness of formulation of the empirical *HOMA* index and for its perfect anchoring in the quantum chemical theory. It is remarkable that the correlations are satisfied not only for aromaticity but also for its generalization, the *savoricity*. Notice, that regular changes of the *HOMA(•)* functions of bond parameters taken in BCPs with *HOMA(R)* are due to good correlations between *R* and bond orders, already observed by Pauling,[14] as well as *R* and parameters in BCP reported by many others.[e.g.,15,16]

Now, taking the advantage of the fact that both *HOMA(R)* and *HOMA(ρ)* can be defined also for acyclic structures we plotted one value against the other for a series of small unbranched alkanes. Fig. 3 shows that the picture for acyclic structures is much more complex. Yet, one can probably select groups of acyclic compounds such as allenes, alkanes, alkadienes, … , etc., that follow regular trends. Nevertheless, a detailed consideration of these and similar tendencies would require a careful consideration of different kind of isomerism that can occur in acyclic compounds, including for example constitutional, geometrical, and conformational isomerism which goes far beyond this study.

Conclusions

The *HOMA* index was generalized by treating it as a function of a parameter assigned to the chemical bonds. The *HOMA* function is constructed practically identically to the classical *HOMA* index, which in the new formulation is the function of bond lengths, *HOMA(R)*. The generalization allows for defining the *HOMA(ρ)* function of electron density in bond critical points. The *HOMA(R)* and *HOMA*(ρ) indices taken for aromatic, unsaturated, and fully saturated cyclic hydrocarbons linearly correlate with regression coefficient of 0.999. This correlation reveals that the both *HOMA* indices are not only aromaticity indices but indices of much more general feature. This feature combines information on cyclicity, degree of cycle unsaturation, and the way of cycle condensation in polycyclic systems. We proposed to call the generalization of the aromaticity property the savoricity. Thus, we demonstrated that *HOMA* index is not an aromaticity index but the index of much more general property of the molecules: their savoricity. Also, *HOMA* functions of the potential and kinetic energy components of electron density in *BCP* are the savoricity indices.

Notes and references

[a] *Institute of Nuclear Chemistry and Technology, 16 Dorodna Street, 03-195 Warsaw, Poland. Fax: +48 (22) 811 1917; Tel: +48 (22) 504 1086; E-mail: j.dobrowolski@ichtj.waw.pl*
[b] *National Medicines Institute, 30/34 Chełmska Street, 00-725 Warsaw, Poland. Fax: +48 (22)841 0652; Tel: +48 (22)841 1888 ext. 355; E-mail: j.dobrowolski@nil.gov.pl*

† Electronic Supplementary Information (ESI) available: schemes of structures and values of *HOMA* indices. See DOI: 10.1039/b000000x/

‡ All calculations were performed using the B3LYP functional,[17,18] standard Pople-type 6-31G** basis set,[19] and Gaussian 09 software.[20] The 6-31G** basis set was shown to perform fairly well in geometry, frequency, energy, and electron density calculations.[21,22] Each minimum was confirmed by checking that all harmonic frequencies were positive. The AIM analysis of electron density parameters was performed using the AIMALL program.[23]

Acknowledgments: This work was supported by the National Science Centre Grant No. 2013/09/B/ST5/03664. The computational Grant G19-4 from the Interdisciplinary Centre of Mathematical and Computer Modelling (ICM) at the University of Warsaw is gratefully acknowledged.

1 A. Julg, and Ph. Françoise, *Theor. Chim. Acta*, 1967, **7**, 249.
2 J. Kruszewski and T. M. Krygowski, *Tetrah. Lett.*, 1972, **13**, 3839.
3 T. M. Krygowski and M. Cyrański, *Tetrah.*, 1996, **52**, 10255.
4 T. M. Krygowski and M. Cyrański, Tetrah., 1996, **52**, 1713.
5 T. M. Krygowski, A. Ciesielski, C. W. Bird and A. Kotschy, *J. Chem. Inf. Comput. Sci.* 1995, **35**, 203-210.
6 A. K. Jain, *Patt. Recog. Lett.*, 2010, **31**, 651.
7 H. Steinhaus, *Bull. Acad. Polon. Sci*. 1956, **4**, 801
8 M. Alonso and B, Herradón, *J. Comput. Chem.* 2010, **31**, 917.
9 A. R. Katritzky, P. Barczynski, G. Musumarra, D. Pisano and M. Szafran, *J. Am. Chem. Soc.*, 1989, **111**, 7.
10 A. R. Katritzky, M. Karelson, S. Sild, T. M. Krygowski and K. Jug, *J. Org. Chem.*, 1998, **63**, 5228.
11 M. K. Cyrański, T. M. Krygowski, A. R. Katritzky and P. von Ragué Schleyer, *J. Org. Chem.*, 2002, **67**, 1333.
12 M. Palusiak and T. M. Krygowski, *Chem. Eur. J.*, 2007, **13**, 7996.
13 A. Tokatli and F. Ucun, *J. Phys. Org. Chem.* 2014, **27**, 380.
14 L. Pauling, L. O. Brockway and J. Y. Beach, *J. Am. Chem. Soc.*, 1935, **57**, 2705.
15 I. Alkorta, L. Barrios, I. Rozas and J. Elguero, *J. Mol. Struct. (Theochem)*, 2000, **496**,131.
16 I. Mata, I. Alkorta, E. Molins and E. Espinosa, *Chem. Eur. J.* 2010, **16**, 2442.
17 A.D. Becke, *J. Chem. Phys*., 1993, **98**, 5648.
18 C. Lee, W. Yang and R.G. Parr, *Phys. Rev. B* 1988, **37**, 785.
19 R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys*., 1971, **54**, 724.
20 M.J. Frisch, *et al*., Gaussian 09, Revision A.1, Gaussian Inc., Wallingford CT, 2009.
21 S. M. Islam, S. D. Huelin, M. Dawe and R. A. Poirier, *J. Chem. Theory Comput.* 2008, **4**, 86.
22 K. Hęclik, B. Dębska and J. Cz. Dobrowolski, *RSC Adv.* 2014, **4**, 17337.
23 AIMAll (Version 14.04.17), Todd A. Keith, TK Gristmill Software, Overland Park KS, USA, 2014 (aim.tkgristmill.com).