MATCH
Communications in Mathematical
and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **56** (2006) 331-344

ISSN 0340 - 6253

On the Dimensionality of Aromaticity Criteria

Ernesto Estrada*

Complex System Research Group, X-Rays Unit, RIAIDT, Edificio CACTUS, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain.

(Received May 3, 2006)

A plausible explanation for the contradictory results published by several groups on the dimensionality of the concept of aromaticity is found. We demonstrate that the good linear relationships between the molar diamagnetic exaltation (Λ) and the resonance energy (RE) of heteroaromatic compounds is not fundamental, but only a consequence of the strong correlation of both properties with the molecular weight. Our findings are based on the fact that molecular weight appears as a "hidden variable" in the resonance energy and as an explicit parameter in the diamagnetic exaltation. Consequently, depending on the data set of compounds selected linear correlations or orthogonality between both properties can be found. New results are obtained mainly when the data set contains great molecular weight variability. This linearity can decrease or even disappear when a data set with smaller molecular weight variability is used.

---

* Fax: +34 981 547 077. E-mail: estrada66@yahoo.com

INTRODUCTION

Aromaticity appears as a not too readily demarcated concept in chemistry.[1] However, it is of inestimable importance for the understanding of organic chemistry as a whole and there is a great amount of both theoretical and practical results that have been inspired by this concept. Due to the incidence of aromaticity on the reactivity, stability, physical, biological and environmental properties of organic compounds, which are in general quantitative properties, much attention has been paid to the study of aromaticity as a quantitative concept.[2, 3]

The first serious attempt to rationalize the quantitative criteria of aromaticity started in 1989 with a series of papers of Katritzky *et al.*[4-11] The first of these papers demonstrated, by using principal component analysis, that there are two independent scales of aromaticity that were called "classical" and "magnetic" criteria of aromaticity.[3] Subsequent papers of this series refined this treatment, but basically the results were maintained. That is, the demonstration that the aromaticity of organic compounds as a quantitative concept is a multidimensional phenomenon.[5-9]

Contrary to these results of Katritzky *et al.*, which concluded that: "the classical and magnetic concepts of aromaticity are almost completely orthogonal", some reports claiming excellent linear correlations between these two criteria of aromaticity appeared in the literature. First, Schleyer's group in 1995 reported that linear relationships exist among the energetic, geometric and magnetic criteria of aromaticity.[12-14] A year later Bird[15] "demonstrated" that "good linear relationships are shown to exist between experimental diamagnetic susceptibility enhancement for some 50 aromatic and heteroaromatic ring systems and the corresponding resonance energies and/or aromaticity indices."

In a 1998 paper, Katritzky *et al.*[11] reaffirmed the multidimensional character of the aromaticity concept. According to them “recent claims that linear relationships exist between energetic, geometric, and magnetic criteria of aromaticity are shown to be invalid for any representative set of heteroaromatics in which the number of heteroatoms varies.” A joint effort of these groups to “present an authoritative assessment” of the dimensionality of aromaticity was published in 2002.[16] In this paper the authors show the existence of “statistically significant correlations among the various aromaticity criteria”, provided the whole set of 75 compounds is used. However, when comparisons are restricted to regions or groups of compounds the correlations can deteriorate or even vanish.

The solution of this controversy is of central importance for the whole experimental and theoretical organic chemistry. There is no doubt from these studies that correlations exist between aromaticity criteria in some instance as well as that these correlations disappear in other situations. In this sense the works of Klein using partial ordering (poset) can be one possible direction.[17,18] Accordingly, “aromaticity is not a single numerical index but rather a collection of posetically interrelated property deviations for which nevertheless useful interrelations and correlations occur”.[18] However, we think that the main question here is to clarify why this apparently contradictory results can arise in dependence of the dataset selected. If you have a general physical law expressed by means of a linear correlation it should be fulfilled whatever the dataset you use.

CORRELATIONS WITH A “HIDDEN” VARIABLE

In a paper published in 1997 Otto Exner analyzed “How to get wrong results from good experimental data” giving a survey of incorrect applications of regression analysis.[19] In this work Exner analyzed a case of correlation with a hidden variable published in a classical textbook on quantum chemistry, which involves calculated delocalization energies *DE* (in β units) and experimental resonance energies *RE*. There is a good linear correlation between these two variables with $R = 0.994$. However, as Exner explains there is no a real cause-and-effect relation between these two variables: “it is not fair that calculations of *DE* give a true picture of the strength of conjugation in the given molecules”.[19] The only cause of this correlation is the existence of a hidden variable: the molecular weight (*M*). Both variables, *DE* and *RE*, depend on the molecular weight — *RE* is derived from the molar enthalpy of combustion, which depends explicitly on *M*, and *DE* depends strongly on the number of π-electrons. When both variables are normalized by dividing them by *M* the correlation drops up to $R = 0.718$, which leads to the conclusion that “*DE* is unable to predict the experimental specific resonance energy”. In this sense it is evident that the molecular weight is a “hidden variable” in the definition of the resonance energy, which is the main cause of the correlation observed with the delocalization energy. However, there are other cases in which the molecular weight appears in the very definition of the variable, as in the case of the “molar quantities”, which is the case of the diamagnetic susceptibility.

DIAMAGNETIC SUSCEPTIBILITY EXALTATION

The diamagnetic susceptibility exaltation or enhancement Λ is defined as the difference between the diamagnetic susceptibility of an organic compound (χ_M) and the corresponding magnitude for a hypothetical bond-localized structure (χ'_M):

$$\Lambda = \chi_M - \chi'_M \qquad (1)$$

The magnitude generally used as the diamagnetic susceptibility is the so-called molar diamagnetic susceptibility, defined as follows:[20]

$$\chi_M = \kappa M / d \qquad (2)$$

where κ is the magnetic susceptibility per unit of volume, or simply the volume susceptibility, M is the molecular weight and d is the density of the substance. Consequently, the diamagnetic susceptibility exaltations are, as the diamagnetic susceptibility, directly related to the molecular weight, i.e., they are size extensive:

$$\Lambda = M\left[(\kappa / d) - (\kappa' / d')\right] \tag{3}$$

where κ' and d' are the volume susceptibility and the density of the hypothetical bond-localized structure. As it is evident from expression (3) the diamagnetic susceptibility exaltation Λ explicitly depends on the molecular weight. If we joint together this observation with the fact that the resonance energy contain molecular weight as a "hidden variable" we can find a plausible explanation to the controversy about the dimensionality of aromaticity criteria.

CORRELATIONS BETWEEN AROMATICITY CRITERIA

In the 1998 paper of Katritzky *et al.*[11] reaffirming the multidimensional character of the aromaticty concept it was clear that the good linear correlation claimed by Schleyer *et al.*[12] was only a consequence of the use of a reduced data set of monocyclic heteroaromatic compounds. When, these compounds are included into a more general data set of compounds (monocyclic heteroaromatics) the correlation completely disappears. However, the data set used by Bird has a great structural variability and Katritzky *et al.*[11] found that the good linear correlation found by this author only disappears when monocycles and compounds with two and three fused cycles are considered separately. These authors also found a good linear relationship between the resonance energy and the number of rings in the compounds. However, it is not clear if the effect of the ring size is the real cause of the relations existing between both criteria of aromaticity. It is possible that the "ring size effect" is only a consequence of another more basic factor that should be investigated.

With the objective to investigate the possible cause of contradiction concerning the dimensionality of the aromaticity concept we select the series of 19 heterocyclic compounds for which Bird[15] reported the diamagnetic susceptibility exaltations and the resonance energies. Compounds for which any of these two values were not reported in the mentioned study were excluded. The data set is shown in Table 1.

Table 1. Molar diamagnetic exaltations (Λ), resonance energies (RE), Bird index of aromaticity (I_A) and molecular weights (MW) for the heteroaromatic compounds. Data taken from ref. 15.

No.	compound	Λ-10^{-6}cm^{3}mol	RE kcal/mole	I_A	MW
1	pyrrole	10.2	34.8	85	67.09
2	imidazole	12.3	40.0	79	68.08
3	furan	8.9	27.2	53	68.08
4	pyrazole	11.4	40.4	90	68.08
5	isoxazole	13.1	34	52	69.06
6	oxazole	10.4	26.2	47	69.06
7	pyridine	13.5	43.36	86	79.10
8	pyridazine	13.1	32.7	99	80.09
9	pyrimidine	12.7	40.6	84	80.09
10	1,3,5-triazine	19.5	44.9	100	81.08
11	thiophene	13.0	43	81.5	84.14
12	thiazole	11.6	42.0	79	85.13
13	indole	24.2	73.8	146	117.15
14	benzofurane	22.2	55.4	94	118.14
15	benzimidazole	16.6	78.9	148	118.14
16	quinoline	26.85	81.0	134	129.16
17	isoquinoline	24.75	81.0	133	129.16
18	carbazole	33.2	111.7	207	167.21
19	acridine	40.75	108.3	198	179.22

A correlation between the diamagnetic exaltations and the classical criteria of aromaticity for these 19 heterocyclic compounds produced a regression coefficient of 0.929 for the resonance energies ($\Lambda = 0.951 + 0.308RE$) and of 0.895 for the Bird aromaticity index[18] ($\Lambda = -0.389 + 0.173I_A$). This data set has a great variability in the molecular weight of the compounds considered. For instance, the standard deviation in the molecular weight of the 19 compounds is 34.53, and the biggest difference in molecular weight expressed in percentage is 62.56 %. However, the data set originally used by Katritzky[4] in 1989 had a standard deviation in the molecular weight of only 7.41 and the biggest difference of this magnitude was only 20 %. Similar values are obtained for the data set recently used by Katritzky,[11] a standard deviation of 7.30 and a difference of 21.2 %.

If we analyze in detail the data set of Bird,[15] we see that the compounds form three groups according to their molecular weights (MW). The first group is comprised from compounds with MW < 70, the second is comprised from compounds with 70 < MW < 90, and the third from compounds with 100 < MW < 130 (carbazole and acridine are not included in any of these groups). The standard deviation in the MW for compounds in these groups are 0.74, 2.45 and 6.23 for the first, second and third groups, respectively. When the resonance energies are correlated to the diamagnetic susceptibility exaltations for compounds in these groups the following correlation coefficients are obtained: 0.620, 0.390 and 0.133 for first, second and third group, respectively. Similar results are obtained with the Bird index for the three groups with correlation coefficients of 0.138, 0.712 and 0.136, respectively.

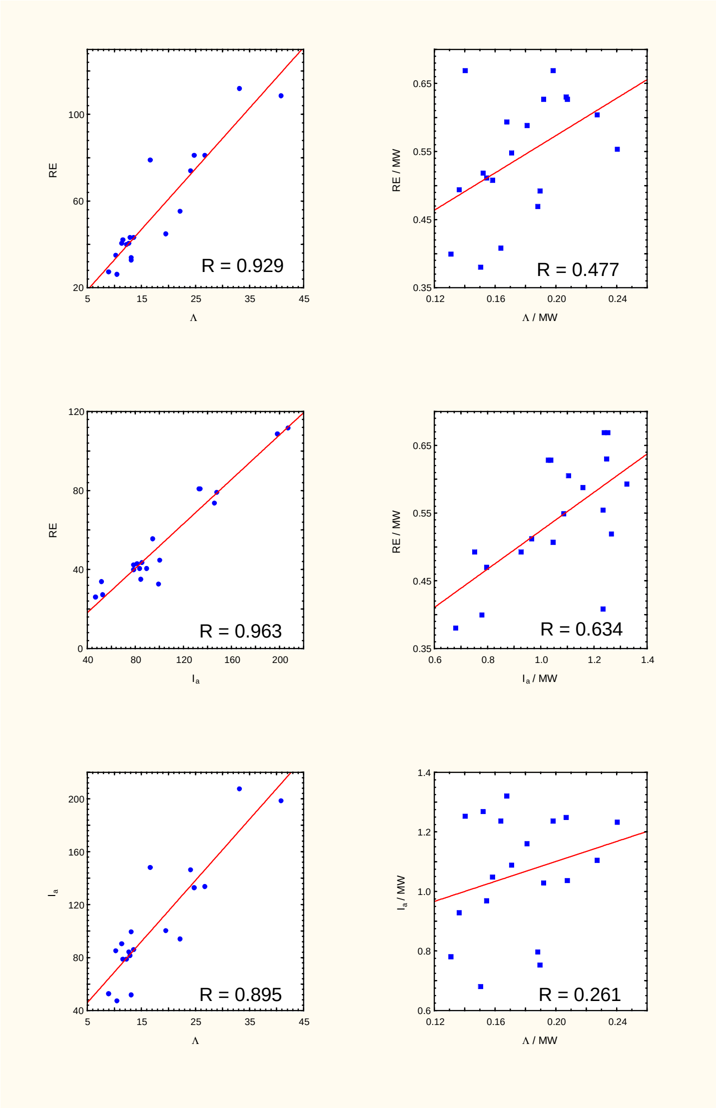
These results point to the conclusion that the good linear correlation observed by Bird and probably by others is not fundamental, but only a consequence of the linear dependence of both criteria of aromaticity with the molecular weight. The correlation between the diamagnetic exaltations and the molecular weights of the 19 heteroaromatic compounds is 0.963 ($\Lambda = -5.866 + 0.242MW$), which is not surprising considering that the molecular weight appears explicitly in the definition of this magnitude (see Eq. (3)). While the correlation between the resonance energy and the molecular weight for these compounds gives an excellent coefficient of 0.971 ($RE = -17.66 + 0.740MW$). Similarly, the Bird index of aromaticity[21] correlates well with the molecular weight, regression coefficient 0.928 ($I_A = -13.47 + 1.212MW$). As can be seen both criteria of aromaticity are strongly correlated with the molecular weight and it is probably the main cause of the good linear correlations found between them. However, in data sets where the molecular weight plays an unimportant role this good linear correlation disappears. This is, for instance, the case of the three groups formed by us from the data set of Bird.[15]

If these conclusions are true, the good linear correlation between both criteria of aromaticity in Bird's data set should disappear by normalizing the properties correlated by the molecular weight. Here we want to point out a very important question. This normalization of the diamagnetic susceptibility exaltation and the resonance energy do not pretend to obtain new variables that can be considered as criteria of aromaticity. Our unique objective is to find a way to demonstrate that the principal cause of the observation of good linear correlations between both criteria of aromaticity is the influence of the molecular weight in both properties. A better approach to obtaining an "absolute" aromaticity criterion is the one used by Aihara in a recent paper who demonstrates that the use of "circuit resonance energy" accounts for both the energetic and magnetic criteria or aromaticity.[22] This criterion, which is essentially the same as the one introduced by Bosanac and Gutman in 1977,[23] is based on chemical graph theory and can be considered as a "structure-explicit" criterion of aromaticity instead of "structure-implicit" or "structure-cryptic" criteria frequently used, according to the phraseology introduced by Trinajstić *et al*.[24]

In Fig. 1 we illustrate the relationship existing between the resonance energy and the Bird index of aromaticity with the diamagnetic susceptibility exaltation, all normalized by dividing by the molecular weight. For the sake of comparison we also illustrate the correlations of the non-normalized variables as well as the correlations between the magnetic criteria of aromaticity. As can be seen there is a lack of linear correlation between both criteria of aromaticty for the 19 heterocyclic compounds when these criteria are normalized by the molecular weight. The poor correlation coefficients obtained when the normalized variables are used are a clear evidence that the good linear correlation previously reported was caused by the strong dependence of both properties on the molecular weight.

These results confirm our hypothesis about the role of the molecular weight in the explanation of the contradictory results reported in the literature on the orthogonality of classical and magnetic criteria of aromaticity. They also explain the observations of Katritzky[11] that better results are obtained when the data set of diamagnetic exaltations reported by Bird[15] are related to the number of rings instead of to the resonance energies. In this case the number of rings acts as a molecular descriptor codifying information on the molecular weight but not as a fundamental variable.

Figure 1. Illustration of the regression plots between different non-normalized (left) and normalized (right) aromaticity criteria for the compounds studied in this work.



In order to analyze the relationship between the different criteria of aromaticity, the molecular weight and the criteria of aromaticity normalized, we use principal component analysis (PCA).[25] By considering Λ, RE, I_A, M, Λ/M, RE/M and I_A/M in a PCA experiment we obtain three eigenvalues. The first of them (5.004) explains 71.48 %, the second (1.096) explains 15.67 % and the third (0.629) explains 8.98 % of the variance of the studied variables for the 19 compounds. In general, they explain 96.12 % of the total variance. The factor loadings shown in Table 2 obtained after Varimax rotation[25] clearly identify the term Λ / M as completely orthogonal to the rest of the variables including the criteria of aromaticity normalized by the molecular weight. This characteristic feature of Λ / M reflexes the uniqueness of this aromaticity criterion. Exner´s idea of considering M in the context of delocalization and resonance energies could be misguided due to the fact that changing inert side-chains, or isotopes, should certainly have little to do with either DE or RE. Consequently, introducing the effects of such substitutions into RE/M or DE/M would just introduce spurious variations. However, this is not the case for Λ, because according to Eq. (3) even an isotopic substitution will alter the value of this property, which explicitly depends on M.

Table 2. Factor loadings after a Varimax rotation for the different criteria of aromaticity and molecular weight studied in the present work.

variable	factor 1	factor 2	Factor 3
Λ	**0.863**	0.069	0.489
RE	**0.941**	0.251	0.216
I_A	**0.873**	0.410	0.210
MW	**0.967**	0.065	0.231

Λ / MW	0.323	0.149	**0.934**
RE/MW	0.575	**0.681**	0.192
I_A/MW	0.095	**0.962**	0.099

The concept of diamagnetic susceptibility exaltation has an important role in the modern theory of aromaticity. For instance, Schleyer[26] has defined aromatic compounds on the basis of this physical magnitude as follows: “Compounds which exhibit significantly exalted diamagnetic susceptibility are aromatic.” The currently used criteria of magnetic aromaticity, that is the molar diamagnetic susceptibility exaltation, is a multidimensional physicochemical magnitude that strongly depends, in an explicit way, on the molecular weight. Thence we indicate an alert as the possible failures which can occur when this magnitude is correlated to other extensive properties such as the resonance energy. The existence or not of good linear correlations between both criteria of aromaticity are influenced by the nature of the data set selected. If there is a great variability in the molecular weight it is possible to find a high degree of linearity between both magnitudes.

By way of conclusion we would like to call the attention of users to regression analysis for deriving cause-and-effect relations between variables as well as physical, chemical and biological reasoning. As the review commentary of Exner[19] shows — and the current work also illustrates — the use of regression analysis might suggests results completely at variance with the original data. We need to take into account the existence of hidden variables or explicit dependence with other variables, such as in the case of molar quantities, such as the molar refraction which is widely used in QSAR and QSPR. In closing, we have to take care in the use of regression analysis to extract a conclusion to avoid a proof the saying — usually attributed to

Mark Twain or Benjamin Disraeli — that “there are three kinds of lies: lies, damned lies, and statistics”.

REFERENCES

(1) Lloyd, D. *J. Chem. Inf. Comput Sci*. **1996**, 36, 442.

(2) Minkin, V. J.; Glukhovtsev, M. N.; Simkin, B. Ya. *Aromaticity and Antiaromaticity: Electronic and Structural Aspects*; John Wiley & Sons: New York, 1994.

(3) Schleyer, P. v. R. Guest Ed. Special issue on aromaticity. *Chem. Rev*. **2001**, 101(5). Randić, M. *Chem. Rev*. **2003**, 103, 3449.

(4) Katritzky, A. R.; Barczynski, P.; Musumarra, G.; Pisano, D.; Szafran, M. *J. Am. Chem. Soc*. **1989**, 111, 7.

(5) Katritzky, A. R.; Feygelman, V.; Musumarra, G.; Barczynski, P.; Szafran, M. *J. Prakt. Chem*. **1990**, 332, 853.

(6) Katritzky, A. R.; Feygelman, V.; Musumarra, G.; Barczynski, P.; Szafran, M. *J. Prakt. Chem*. **1990**, 332, 870.

(7) Katritzky, A. R.; Barczynski, P. *J. Prakt. Chem*. **1990**, 332, 853.

(8) Katritzky, A. R.; Szafran, M.; Malhotra, N.; Ullah Chaudry, S.; Anders, E. *Tetrahedron Comput. Methodol*. **1990**, 3, 247.

(9) Katritzky, A. R.; Karelson, M.; Malhotra, N. *Heterocycles* **1991**, 32, 127.

(10) Katritzky, A. R.; Karelson, M.; Wells, A. P. *J. Org. Chem*. **1996**, 61, 1619.

(11) Katritzky, A. R.; Karelson, M.; Sild, S.; Krygowski, T. M.; Jug. K. *J. Org. Chem*. **1998**, 63, 5228.

(12) Schleyer, P. von R.; Freeman, P. K.; Jiao, H.; Goldfuβ, B. *Angew. Chem., Int. Ed. Engl.* **1995**, 34, 337.

(13) Schleyer, P. von R.; Jiao, H. *Pure Appl. Chem*. **1996**, 68, 209.

(14) Subramanian, G.; Schleyer, P. von R.; Jiao, H. *Angew. Chem., Int. Ed. Engl.* **1996**, 35, 2638.

(15) Bird, C. W. *Tetrahedron* **1996**, 52, 9945.

(16) Cyrañski, M. K.; Krygowski, T. M.; Katrizky, A. R.; Schleyer, P. v. R. *J. Org. Chem.* **2002,** 67, 1333.

(17) Klein, D. J. *J. Chem. Ed.* **1992**, 691.

(18) Klein, D. J.; Babić, D. *J. Chem. Inf. Comput. Sci*. **1997**, 37, 656.

(19) Exner, O. *J. Phys. Org. Chem*. **1997**, 10, 797.

(20) Dorfman, Y. G. *Diamagnetism and the Chemical Bond*; Edmond Arnold Pub. Ltd.: London, 1965.

(21) Bird, C. W. *Tetrahedron* **1992**, 48, 335.

(22) Aihara, J-I. *J. Am. Chem. Soc.* **2006**, 128, 2873.

(23) Bosanac, S.; Gutman, I. *Z. Naturforsch*. **1977,** 32a, 10.

(24) Trinajstić, N.; Randić, M.; Klein, D. J. *Acta Pharm. Jugosl*. **1986**, 36, 267.

(25) Malinowski, E. R.; Hoowery, D. G. *Factor Analysis in Chemistry*; Wiley: New York, 1980.

(26) Schleyer, P. von R. 8th International Symposium on Novel Aromatic Compounds (ISNA 8); Braunschweig, 1995. Cited in ref. 1.