CROATICA CHEMICA ACTA
CCACAA, ISSN 0011-1643, e-ISSN 1334-417X
Croat. Chem. Acta **83** (3) (2010) 349–351.
CCA-3426

Note

Bond-additive Modeling. 3. Comparison between the Product-connectivity Index and Sum-connectivity Index

Damir Vukičević[a,*] and Nenad Trinajstić[b]

[a]*Faculty of Natural Sciences and Mathematics, University of Split, Nikole Tesle 12, HR-21000 Split, Croatia*
[b]*The Ruđer Bošković Institute, HR-10002 Zagreb, Croatia*

RECEIVED MAY 25, 2009; REVISED JANUARY 7, 2010; ACCEPTED JANUARY 14, 2010

Abstract. Using the framework of a previously defined procedure (Refs. 16 and 17), we compared the Randić (product-)connectivity index and the sum-connectivity index for the benchmark sets of molecules, that is, 18 octanes, 82 polycyclic aromatic hydrocarbons, 209 polychlorobiphenyls and 22 phenethylamines.

Keywords: Randić (product-)connectivity index, sum-connectivity index, bond-additive modeling

INTRODUCTION

Randić proposed in 1975 a structural descriptor, called branching index,[1] that later became the well-known Randić connectivity index, which is the most used molecular descriptor in structure-property-activity modeling, *e.g.*[2–8] The name *connectivity index* that replaced the original Randić term the *branching index* was suggested by Kier as stated by Randić in his review on the development of the connectivity index.[9] The first paper in which the Randić connectivity index was used in modeling molecular properties also appeared in 1975.[10] Mathematical properties of the Randić connectivity index have also been continuously studied, *e.g.*[11–13]

The Randić connectivity index, denoted by χ, is defined as:[1]

$$\chi = \sum_{ij \in E(G)} \left(d_i d_j \right)^{-1/2}.$$

where G denotes a (molecular) graph, E the set of edges (bonds) in a graph (molecule), d_i the valence (degree) of a vertex (atom) i and the product $d_i d_j$ the weight of the i-j edge (bond). We call for convenience the Randić connectivity index the *product*-connectivity index.

A recently introduced variant of the connectivity index,[14,15] denoted as ${}^{s}\chi$ which we call the *sum*-connectivity index, is defined as

$${}^{s}\chi = \sum_{ij \in E(G)} \left(d_i + d_j \right)^{-1/2}.$$

The aim of this report is to compare the product- and sum-connectivity indices and to determine predictive properties of the sum-connectivity index using the procedure already outlined in two previous reports in this series.[16,17] Stimulus for this and related reports is the following suggestion:[18] *In order to promote the comparison of new and old molecular descriptors, evaluate their predictive ability and better understand their meaning, the International Academy of Mathematical Chemistry*[19] *suggests the use of some benchmark data sets*. In order to perform such analyses, here we used the modification of the software that has already been used in the analyses of Adriatic indices.[16,17]

The following four sets of chemical compounds are proposed as the benchmak sets:[16] the set of 18 octane isomers, the set of 82 polycyclic aromatic hydrocarbons (PAH), the set of 209 polychlorobiphenyls (PCB) and the set of 22 phenethylamines (Phenet). 16 properties and 102 descriptors are given for the set of octane isomers; 3 properties and 112 descriptors are given for PAHs; 8 properties and 106 descriptors for PCBs; and one property and 110 descriptors for phenethylamines. We present in Table 1 the best correlations of one-parameter linear models based on descriptors in these benchmark sets and correlations of product-connectivity index and sum-connectivity index. We analyzed all proposed properties, except the melting point since it is known that its features cannot be captured by descriptors that solely depend on degrees of vertices.

* Author to whom correspondence should be addressed. (E-mail: vukicevi@pmfst.hr)

Table 1. Analyses of the predictive properties of the product-connectivity index and the sum-connectivity index

Set of chemical compounds	Property	Highest correlation coefficient R^2 in the benchmark set	Correlation coefficient R^2 of the product-connectivity index	Correlation coefficient R^2 of the sum-connectivity index
Octane isomers	Boiling point	0.7762	0.6737	0.6438
	Heat capacity at T constant	0.5048	0.2488	0.2067
	Heat capacity at p constant	0.5938	0.0199	0.0375
	Entropy	0.9163	0.8205	0.8519
	Density	0.5937	0.3121	0.3318
	Enthalpy of vaporization	0.8861	0.8763	0.8683
	Standard enthalpy ofvaporisation	0.9201	0.9181	0.9239
	Enthalpy of formation	0.8324	0.7233	0.6915
	Standard enthalpy of formation	0.6692	0.0076	0.0158
	Motor octane number	0.9274	0.6037	0.6707
	Molar refraction	0.9794	0.0072	0.0247
	Acentric factor	0.9923	0.8176	0.8647
	Total surface area	0.7169	0.0628	0.0894
	Octanol-water partition coefficient	0.2941	0.0294	0.0259
	Molar volume	0.5483	0.29588	0.3184
PAH	Boiling point	0.97978	0.9764	0.9777
	Octanol-water partition coefficient	0.9413	0.8993	0.8871
	Relative retention time	0.9600	0.8984	0.8992
	Octanol-water partition coefficient	0.9305	0.8420	0.8431
	Total surface area	0.9965	0.9926	0.9930
PCB	log Henry constant	0.7057	0.0147	0.0146
	log water solubility	0.9440	0.9209	0.9215
	log water activity coefficient	0.8260	0.8041	0.8049
	Relative enthalpy of formation	0.6713	0.4566	0.4564
Phenet	Biological activity: log(1/C)	0.5405	0.36589	0.3664

We learn from results in Table I the following: (i) both connectivity indices produce similar values of the correlation coefficients in most cases considered; (ii) the sum-connectivity index has slightly better predictive properties for the standard enthalpy of vaporisation of octane isomers than any of indices in the benchmark set and the product connectivity index; (iii) in three cases (the boiling points of PAH, the total surface areas of PCB and the water solubilties of PCB) the connectivity indices produce models with high values of the correlation coefficients that are also comparable to the models produced by the benchmark sets, (iv) in several cases the connectivity indices produced decent models (the correlation coefficients being between 0.8 and 0.9), but inferior to the benchmark models and finally (v) among the 11 models based on the connectivity indices

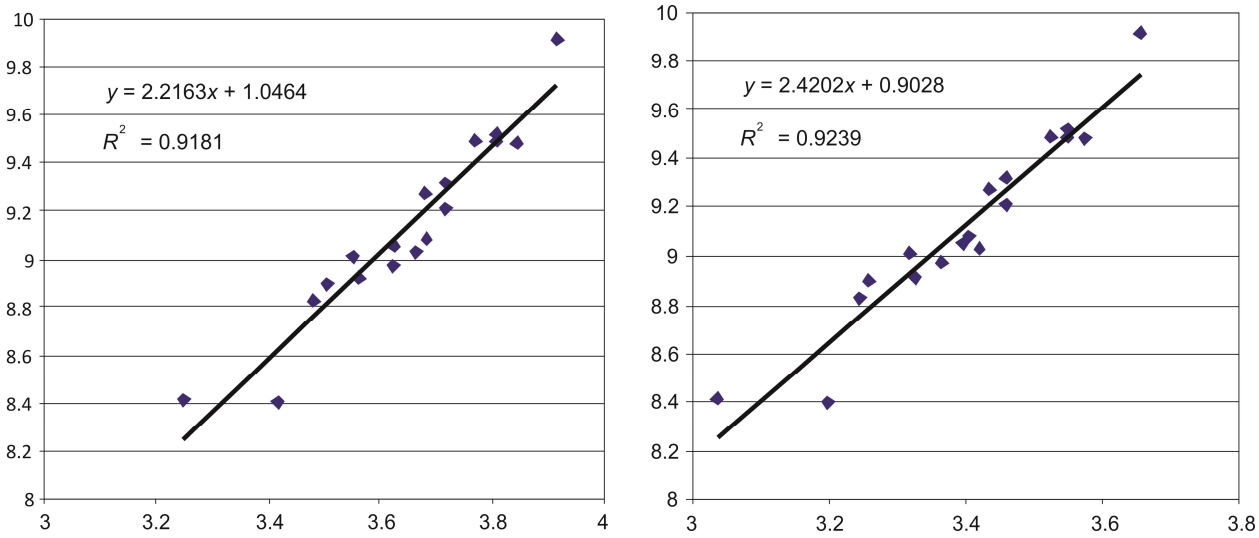

Figure 1. Linear models for predicting the standard enthalpy of vaporisation of octane isomers by product-connectivity index (left hand diagram) and connectivity-sum index (right hand diagram).

with values of the correlation coefficient higher than 0.8, models based on the sum-connectivity indices have only in 2 cases higher values of the correlation coefficient than models based on the product-connectivity indices, and in the case of the remaining 9 models the differences are rather negligible.

In Figure 1, we present linear models for prediction of standard enthalpy of vaporisation of octane isomers by the product-connectivity index (left hand-side) and connectivity-sum index (right hand-side).

Acknowledgements. This report was supported in part by the Ministry of Science, Education and Sport of Croatia.

REFERENCES

1. M. Randić, *J. Am. Chem. Soc*. **97** (1975) 6609−6615.
2. L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
3. L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Research Studies Press/Wiley, Letchworth/New York, 1986.
4. J. Devillers and A. T. Balaban (Eds.), *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon and Breach, The Netherlands, 1999.
5. M. Karelson, *Molecular Descriptors in QSAR/QSPR*, Wiley-Interscience, New York, 2000.
6. R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000.
7. L. Pogliani, *Chem. Rev*. **100** (2000) 3827−3858.
8. R. Garcia-Domenech, J. Galvez, J. V. de Julian-Ortiz, and L. Pogliani, *Chem Rev.* **108** (2008) 1127−1169.
9. M. Randić, *MATCH Commun. Math. Comput. Chem*. **59** (2008) 5−124.
10. L. B. Kier, L. H. Hall, W. J. Murray, and M. Randić, *J. Pharm. Sci*. **64** (1975) 1971−1974.
11. X. Li and I. Gutman, *Mathematical Aspects of Randić-Type Molecular Structure Descriptors*, University of Kragujevac, Kragujevac, 2006.
12. I. Gutman and B. Furtula (Eds.), *Recent Results in the Theory of Randić Index*, University of Kragujevac, Kragujevac, 2008.
13. Z. B. Du, B. Zhou, and N. Trinajstić, *MATCH Commun. Math. Comput. Chem*. **62** (2009) 131−142.
14. B. Zhou and N. Trinajstić, *J. Math. Chem*., in press. Available on request.
15. B. Lučić, N. Trinajstić and B. Zhou, *Chem. Phys. Lett*. **475** (2009) 146−148.
16. D. Vukičević and M. Gašperov: Bond Additive Modeling 1. Adriatic Indices, *Croat. Chem. Acta*, in press.
17. D. Vukičević: Bond Additive Modeling 2. Mathematical properties of Max-min rodeg index, *Croat. Chem. Acta*, in press.
18. http://www.moleculardescriptors.eu/dataset/dataset.htm
19. http://www.iamc-online.org/

SAŽETAK

Vezno-aditivno modeliranje. 3. Usporedba između produktnoga indeksa povezanosti i zbrojitbenoga indeksa povezanosti

Damir Vukičević[a] i Nenad Trinajstić[b]

[a]*Prirodoslovno-matematički fakultet, Sveučilište u Splitu, Nikole Tesle 12, HR-21000 Split, Hrvatska*
[b]*Institut Ruđer Bošković, HR-10002 Zagreb, Hrvatska*

Uporabivši već opisani postupak u ref. 16 i 17, autori su usporedili Randićev (produktni) indeks povezanosti s zbrojidbenim indeksom povezanosti na sljedećim skupovima molekula: 18 oktana, 82 policiklička aromatička ugljikovodika 209 poliklorobifenila i 22 fenetilamina, koje je *Međunarodna akademija za matematičku kemiju* predložila kao temeljne skupove molekula za testiranje prediktivnosti molekularnih deskriptora.