MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **69** (2013) 845-846

ISSN 0340 - 6253

A Noteworthy Referee Report

on the paper

M. Dehmer, M. Grabner, The discrimination power of molecular identification numbers revisited, *MATCH Commun. Math. Comput. Chem.* **69** (2013) 785-794

This paper, published in the same issue of *MATCH Communications in Mathematical and in Computer Chemistry,* was earlier submitted for publication to the *Current Computer-Aided Drug Design*. In May 2012 the paper was rejected, based on the following referee report.

Comments for the Authors (continue on another sheet, if necessary):

The paper investigates the degeneracy of several topological indices, namely a number of molecular ID numbers based on a TI proposed by Randic, and a number of entropies calculated from graph spectra.

It should be mentioned clearly from the beginning that the TI degeneracy problem has no application in QSAR, QSPR, molecular design, drug development, molecular modeling, or any other related research topic.

The study of degeneracy of graph invariants was initiated by Harary and Balaban, and it is an important topic in graph theory and mathematical chemistry.

As such, the topic of this paper is not fit for CCADD. In addition, the paper adds too little compared with previous studies in the same topic, and the paper is of no interest to the readers of CCADD. It must be mentioned that despite decades of research in TI degeneracy, there is not a single application of TI degeneracy in QSAR or related fields.

As to the subject of this paper, it is a simple exercise which does not push the limits of knowledge. In 1985 it was reported the TI degeneracy for molecular ID numbers representing chemical trees with up to 20 carbon atoms. After more than 25 years and extraordinary progress in computing power, this paper reports results for up to 22 carbon atoms, which by now is a simple exercise. One would expect after 25 years to push the limit to much higher numbers, at least hundreds.

The number of TI investigated for degeneracy is also small, and the selection is at least strange, because there is no connection between molecular ID numbers and spectral entropy. Important categories of TI are ignored without explanation or without a scientific justification of the selection limited to these indices.

As to the TI based on spectra, it must be mentioned that spectral entropy is known from the 1950's and it is by no means a "new" development, as a formula. All previous studies are ignored in this manuscript.

Regarding the degeneracy of spectral entropy, it must be kept in mind that it is a useless investigation, because all isospectral graph matrices have identical spectra and as a consequence they have also identical spectral entropy (or any other TI computed only from spectra), and this is a result that can be obtained without computing the spectral entropy. However, this simple property was not even considered, which resulted in this unprofessional approach. Bottom line, the correct approach is to compare the spectra and not to compute the spectral entropy.

Randic and Trinajstic studied extensively the graph invariant degeneracy, and they provided important insight into the structural origin of such degeneracy.

In this manuscript such investigation is missing, which explains its low grade in review.

Even as a graph theory or mathematical chemistry investigation, there is too little new contribution in this manuscript.

Matthias Dehmer's comments: The review is in parts erroneous, contradicting and shows that there is some confusion when dealing with the degeneracy problem of topological graph measures. As a matter of fact, the degeneracy problem in terms of graphs has been studied extensively by Bonchev (1981), Szymanski (1985), Ihlenfeldt (1994) and many others. As an application, Ihlenfeldt and Gasteiger (1994) developed unique hashcodes for the indentification and classification of molecular graphs. Also, Carter et al. (1988) employed highly discriminating ID numbers for solving problems in structure-oriented drug design. In my opinion, it is valuable to evaluate the uniqueness of existing topological graph measures by using different graph classes, even then, they are claimed to be highly unique. As the present paper shows, the ID numbers are not applicable to large data sets (e.g., exhaustively generated graphs) and already known indices, which are much easier to compute, give very similar results for the uniqueness.