



Dehmer M., Varmuza K., Bonchev D.
(eds.):

**Statistical Modelling of Molecular
Descriptors in QSAR/QSPR**

2012, Wiley-VCH, Weinheim, Germany

Sample from the Preface - referring the book cover

...

We would like to sketch the idea of choosing the book cover in brief.

Note that it has been inspired by a short NASA report from April 1995 trying to demonstrate the complexity of QSAR/QSPR in a multivariate setting. The authors of this report, D.A. Noever, R.J. Cronise, and R.A. Relwani, exposed spiders to substances with different toxicity and claimed that the changes in the spider webs reflect the degree of toxicity. For caffeine - the molecule shown on the book cover - the spiders produced only unstructured webs instead of rather symmetrical, radial webs as shown in the background of the cover.

From a statistical point of view, one regrets that no estimations of the reproducibility are given in the report and obviously no further literature exists dealing with this subject; although the original report has been cited frequently.

From a point of view of QSAR one may doubt that the toxic effect on spiders can be easily translated to explain toxic effect on other animals or even humans. Furthermore, the effect is not really surprising considering well-known effects of drugs and ethanol when it comes to humans.

When speculating one may be seduced to look for relationships between the networks describing chemical structures and the networks of distorted spider webs.

A different approach is the crucial idea on which the book and its contributions is based: Starting from a molecular structure, a set of descriptors is calculated, for example, information-theoretic indices by using Shannon's entropy as indicated by the cover figure. Hence, a set of chemical structures can be thereby represented by a matrix where each row corresponds to a structure. Typically, multivariate data analysis methods can be applied to such data to generate empirical models that relate a property of substances to the molecular descriptors derived from the chemical structures. Essential for such empirical models is a careful and cautious evaluation of the performance - otherwise one might quickly run into speculation and circular reasoning. In this context, we hope that the book may help to avoid this and also might be stimulating for understanding the mentioned problems more deeply.

...

Matthias Dehmer and Kurt Varmuza thank the Austrian Science Fund (FWF) for supporting this work (project P22029-N13).

120315