MATCH
Communications in Mathematical and in Computer Chemistry

MATCH Commun. Math. Comput. Chem. **68** (2012) 959-960

ISSN 0340 - 6253

BOOK REVIEW

Statistical Modelling of Molecular Descriptors in QSAR/QSPR

edited by

Matthias Dehmer, Kurt Varmuza & Danail Bonchev

Wiley–Blackwell, Chichester, 2012, XIX+436 pp. ISBN 978-3-527-32434-7

The book "*Statistical Modelling of Molecular Descriptors in QSAR/QSPR*" consists of a short Preface (written by Dehmer, Varmuza and Bonchev), followed by 15 chapters, written by a total of 40 authors. In the Preface it is pointed out that the book presents theoretical and practical results towards the statistical analysis and modeling of data generated by using molecular descriptors; the major goal of the book is to advocate and promote a combination of statistical, information–theoretic, and data analysis techniques in QSAR/QSPR.

The first chapter, entitled *Current Modeling Methods Used in QSAR/QSPR* (pp. 1–31), written by L. C. Yee and Y. C. Wei, provides a brief, informative, and palatable introduction to the methods used in QSAR/QSPR studies, avoiding complicated technical details. On the first page of Chapter 1 we read of a remarkable and most important detail: QSAR studies may save to the pharmaceutical industry up to US$ 100 million in developing costs of each new drug.

The following 12 chapters deal with particular aspects of these methods. Their titles are self–explanatory: **2.** *Developing Best Practices for Descriptor–Based Property Prediction: Appropriate Matching of Databases, Descriptors, Methods, and Expectations* (pp. 33–64), **3.** *Mold*^{2} *Molecular Descriptors for QSAR* (pp. 65–109), **4.** *Multivariate Analysis of Molecular Descriptors* (pp. 111–147), **5.** *Partial–Order*

Ranking and Linear Modeling: Their Use in Predictive QSAR/QSPR Studies (pp. 149–174), **6.** *Graph–Theoretical Descriptors for Branched Polymers* (pp. 175–199), **7.** *Structural–Similarity–Based Approaches for the Development of Clustering and QSPR/QSAR Models in Chemical Databases* (pp. 201–228), **8.** *Statistical Methods for Predicting Compound Recovery Rates for Ligand–Based Virtual Screening and Assessing the Probability of Activity* (pp. 229–243), **9.** *Molecular Descriptors and the Electronic Structure* (pp. 245–292), **10.** *New Types of Descriptors and Models in QSAR/QSPR* (pp. 293–305), **11.** *Consensus Models of Activity Landscapes* (pp. 307–326), **12.** *Reverse Engineering Chemical Reaction Networks from Time Series Data* (pp. 327–348), **13.** *Reduction of Dimensionality, Order, and Classification in Spaces of Theoretical Descriptions of Molecules: An Approach Based on Metrics, Pattern Recognition Techniques, and Graph Theoretic Considerations* (pp. 349–364).

The last two chapters are concerned with topics unrelated to QSAR/QSPR: **14.** *The Analysis of Organic Reaction Pathways by Brownian Processing* (pp. 365–391), **15.** *Generation of Chemical Transformations: Reaction Pathways Prediction and Synthesis Design* (pp. 393–425). Leaving these chapters out of the book would have been an asset rather than a loss.

Each chapter contains an exhaustive list of references. The book ends with a detailed Index (pp. 427–436).

With the exception of the last two chapters, the book provides a valuable survey of the many (but certainly not all) techniques used in contemporary QSAR/QSPR studies. Each chapter outlines a particular method, and illustrates it by carefully chosen and reasonably limited real–world examples.

A neophyte in the field should first read Chapter 1 and then jump to Chapters 9 and 3. In the beginning of Chapter 3 is a section on the history of QSAR/QSPR, pointing out its nineteenth century origins.

Those already mastering the fundamentals of QSAR/QSPR, may especially enjoy reading the Chapters 4, 5, and 10, authored, respectively, by Roberto Todeschini, Eduardo A. Castro, and Timothy Clark, with coauthors.

In summary, "*Statistical Modelling of Molecular Descriptors in QSAR/QSPR*" is a valuable treatise, aimed at practitioners, useful both for beginners and experts. It should be a must for any decent science library.

Ivan Gutman