H. Senderowitz [1]

STATISTICAL MODELING IN MATERIAL SCIENCES

[1] Department of Chemistry, Bar Ilan University, Ramat-Gan 52900, Israel.

Statistical modeling (also termed Quantitative Structure Activity Relationship, QSAR or Quantitative Structure Property Relationship, QSPR) is a general name for a host of methods that attempt to correlate a specific activity for a set of compounds with their structure-derived descriptors by means of a mathematical model. Statistical modeling has been widely applied in many fields including chemistry, biology, and environmental sciences. In particular, the role of QSAR models in the identification of new compounds and in their subsequent optimization has been constantly growing and is now recognized by many practitioners of computer aided drug design and computer aided material design methodologies.

This lecture will focus on the application of statistical modeling techniques in material sciences discussing methods typically used in this field (e.g., Principle Component Analysis, Cluster analysis and linear and non-linear regression), material descriptors (e.g., material composition, material spectra) and the challenges in obtaining them and highlighting selected applications.

Special emphasis will be put on the application of statistical modeling in the newly emerging field of photovoltaic cells entirely made of metal oxides (MO). Such cells have the potential to provide clean and affordable energy if their power conversion efficiencies are improved. Such improvements require the development of new MOs which in turn could benefit from combining combinatorial material sciences for producing solar cells libraries with statistical tools to direct synthesis efforts. With this in mind we developed a QSAR workflow and applied it to the analysis of several solar cell libraries. Our results demonstrate that QSAR models with good prediction statistics for multiple solar cells properties could be developed and that these models highlight important factors affecting these properties in accord with experimental findings. The resulting models are therefore suitable for designing better solar cells. We further demonstrate that the similar property principle commonly invoked in pharmaceuticals design could be extended to PV cells.

1. Yosipof A., Nahum O.E., Anderson A.Y., Barad H., Zaban A., Senderowitz H. *Molecular Informatics,* 2015, accepted for publication.
2. Yosipof A., Senderowitz H. *J. Comput. Chem.* 2014, accepted for publication
3. Yosipof A., Senderowitz H. *J Chem. Inf. Model.* 2014, **54:** 1567-77.