Prediction Solubilities of Medium-Sized Complex Chemicals in Liquid Solvents: Retrieval, Reduction and Application

J. Abildskov\textsuperscript{C, S} and R. Gani

Department of Chemical Engineering, Technical University of Denmark, Lyngby, Denmark
ja@kt.dtu.dk

R. Sass

Dept. Information Systems and Databases, DEHEMA e.V., Frankfurt am Main, Germany

Processing and formulation of pharmaceutical, medical, and agrochemical products usually requires the solubility of an active substance in a liquid. Fast and reliable data can greatly enhance the effectiveness of design engineers. However, due to the structural complexity of active ingredients, property prediction models usually provide inaccurate results. Therefore, trial-and-error experimentation dominates most design work. In increasingly competitive climates, reliance on predicted property data is becoming more often necessary. Before complex (solid) solutes can be adequately studied, and their properties predicted, the systematization of the entire area is necessary, in particular with regards to property data.

Despite its practical importance, solid/liquid equilibrium has not received as much attention in the thermophysical property community as has vapor/liquid equilibrium. Many references to experimental work exist, but many are quite old. Because data generation continues to progress rapidly, new results are continuously accumulating in the literature. The time is therefore appropriate for a comprehensive presentation where new (and older) results are made accessible. We have begun to systematically develop a new database [1, 2] of solid solubility data of binary mixtures, where the solute is a medium-sized organic component. The aims are to establish:

1. A comprehensive standard reference database for storage/retrieval of solid/liquid data with public access;
2. Better data prediction models;
3. Efficient data correlation/evaluation procedures;

The paper provides a report covering the data (5100 solubility data points for 340 solutes and 170 solvents from nearly 200 articles) and developments in data correlation, evaluation, prediction, and application to transport and distribution problems.