Estimation of Properties of Pure Components Using Improved Group Contribution Based and Atom Connectivity Index Based Models and Uncertainty Analysis

Hukkerikar, Amol; Sarup, Bent; Sin, Gürkan; Gani, Rafiqul

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Estimation of Properties of Pure Components Using Improved Group Contribution Based and Atom Connectivity Index Based Models and Uncertainty Analysis

Amol Hukkerikar [a], Bent Sarup [b], Gürkan Sin [a] and Rafiqul Gani [a]

[a] CAPEC, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800, Lyngby, Denmark.
[b] Vegetable Oil Technology Business Unit, Alfa Laval Copenhagen A/S, DK-2860, Søborg, Denmark.

Introduction

Pure component physical and thermodynamic properties are needed to carry out tasks such as product/process design and computer aided molecular design. While use of experimentally measured values for the needed properties are desirable, this is not feasible for many reasons. Therefore, efficient and reliable property prediction tools are necessary. The large gap between the number of compounds of interest and the data available for them, drives the interest for development of better and more versatile property prediction models. Among the group contribution methods for estimation of properties of pure compounds, the Marrero and Gani (MG) method (Marrero and Gani, 2001) is well-known. The MG method (i) allows accurate and reliable estimations; (ii) is able to distinguish among some isomers; (iii) is based on significantly large data-sets; and (iv) is based on large and comprehensive sets of groups. For the case where the molecular structure of a given compound is not completely described by any of the available groups, atom connectivity index based estimation method has been employed together with MG method to create the missing groups and to predict their contributions (Gani et al., 2005). This combined approach has led to the development of group contribution$^+$ method of wider application range than before because the missing groups and their contributions are now easily obtained through the regressed contributions of connectivity indices.

The accuracy of product/process design and molecular design largely depends on the accuracy of the underlying physical and thermodynamic property information. Whiting et al (1993) have illustrated the effect of uncertainty in thermodynamic data and models on the process design through several examples and concluded that uncertainty analysis of

property prediction models is essential for getting accurate and feasible solutions in process design, simulation, and optimization.

The objective of this work is to develop new and improved set of parameters for group contribution based and atom connectivity index based models to estimate properties of pure components together with uncertainty in the estimated property value. This includes: a parameter estimation step to determine the model parameters (group contributions / atom contributions and model constants) and an uncertainty analysis step to establish statistical information about the quality of parameter estimation, such as covariance, standard error and confidence interval in the parameters and in the estimated property value. For parameter estimation, large data-sets of experimentally measured property values of a wide range of pure compounds are taken from the CAPEC database. Classical frequentist approach i.e., least square method was adopted for the estimation of model prediction uncertainty (Sin et al., 2010). Improved group and atom contributions along with the standard error and confidence intervals for MG based models have been obtained via regression for the following properties: normal boiling point, critical constants, standard enthalpy of formation, standard enthalpy of vaporization, standard Gibbs energy, normal melting point, standard enthalpy of fusion, entropy of vaporization, surface tension, viscosity, flash point, auto ignition temperature, Hansen solubility parameters, Hildebrand solubility parameter, aqueous solubility, octanol/water partition coefficient, compressibility factor, molar volume, molar refraction, refractive index and lethal concentration. The performance of predictive models for these properties with the new set of group and atom contributions is highlighted through a set of molecules not used in the regression step. The use of the new set of model parameters and quantification of uncertainty (prediction error) allows one to estimate property values with improved prediction accuracy and obtain rationally the risk/safety factors in product/process design.

References

[1] G. Sin, A. Meyer, and K. Gernaey, Computers and Chemical Engineering, (2010), 34, 1385-1392.

[2] J. Marrero, and R. Gani, Fluid Phase Equilibria, (2001), 183-208.

[3] R. Gani, P. Harper, and M. Hostrup, Industrial and Engineering Chemistry Research, (2005), 44, 7262-7269.

[4] W. Whiting, T. Tong, and M. Reed, Industrial and Engineering Chemistry Research, (1993), 32, 1367-1371.