A rigorous methodology for development and uncertainty analysis of group contribution based property models

Frutiger, Jerome; Abildskov, Jens; Sin, Gürkan

Publication date:
2015

A Rigorous Methodology for Development and Uncertainty Analysis of Group Contribution Based Property Models

Jerome Frutiger [C, S], Jens Abildskov and Gürkan Sin
Department of Chemical and Biochemical Engineering, Technical University of Denmark DTU, Kgs. Lyngby, Hovedstaden, Denmark
jfru@kt.dtu.dk

Property prediction models are a fundamental tool of process modeling and analysis, especially at the early stage of process development. Furthermore, property prediction models are the fundamental tool for Computer-aided molecular design used for the development of new refrigerants. Group contribution (GC) based prediction methods use structurally dependent parameters in order to determine the property of pure components. The aim of the GC parameter estimation is to find the best possible set of model parameters that fits the experimental data. In that sense, there is often a lack of attention on numerical and statistical challenges associated with model development and analysis. These challenges include for example (i) performance of optimization algorithms used for finding minimum of the objective function for the parameter estimation, (ii) assessment of parameter estimation errors, (iii) assessment of property model prediction errors, (iii) effect of outliers and data pre-treatment, (iv) formulation of parameter estimation problem (e.g. weighted least squares, ordinary least squares, robust regression, etc.) In this study a comprehensive methodology is developed to perform a rigorous and step-by-step assessment and solution of the pitfalls involved in developing models.

The methodology takes into account of the following steps.

1) Experimental data collection and providing structural information of molecules.

2) Choice of the regression model: a) ordinary least square b) robust or c) weighted-least-square regression.

3) Initialization of estimation by use of linear algebra providing a first guess.

4) Sequential parameter and simultaneous GC parameter by using of 4 different minimization algorithms.

5) Thorough uncertainty analysis: a) based on asymptotic approximation of parameter covariance matrix b) based on boot strap method. Providing 95%-confidence intervals of parameters and predicted property.

6) Performance statistics analysis and model application.

The application of the methodology is shown for a new GC model built to predict lower flammability limit (LFL) for refrigerants. The GC model uses the Marrero-Gani (MR) method which considers the group contribution in different levels both functional and structural. The methodology helps improve accuracy and reliability of property modeling and provides a rigorous model quality check and assurance. This is expected to further their credibility and robustness in wider industrial and scientific applications.