Contents lists available at ScienceDirect

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Data quality and assessment, validation methods and error propagation through the simulation software: Report from the Round-Table Discussion at the 10th World Congress of Chemical Engineering in Barcelona (October 1–5, 2017)

P.M. Mathias[a], *A. Soto*[b], *L. Fele-Zilnik*[c], *J.-C. de Hemptinne*[d,*], *A. Bazyleva*[e], *J. Abildskov*[f]

[a] *Fluor Corporation, 3 Polaris Way, Aliso Viejo, CA 92698, United States*
[b] *Department of Chemical Engineering, Universidade de Santiago de Compostela, E-15782 Santiago de Compostela, Spain*
[c] *Department of Catalysis and Chemical Reaction Engineering, National Institute of Chemistry, Hajdrihova 19, P.O. Box 660, SI-1001 Ljubljana, Slovenia*
[d] *IFP Energies nouvelles, 1&4 Avenue de Bois Préau, 92852 Rueil-Malmaison Cedex, France*
[e] *Applied Chemicals and Materials Division, National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80305-3337, United States*
[f] *PROSYS, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, 2800 Kgs. Lyngby, Denmark*

A R T I C L E I N F O

Article history:
Received 12 June 2018
Received in revised form 3 August 2018
Accepted 6 August 2018

Keywords:
Data quality and assessment
Uncertainties
Data reconciliation
Error propagation
Process simulation

A B S T R A C T

The issues of data quality and propagation of data uncertainties into process design and plant specifications are of great current interest. Hence, two Working Parties of the European Federation of Chemical Engineers (EFCE) organized a Round Table Discussion on the topic, as part of the World Congress of Chemical Engineering (WCCE10) in Barcelona, in October 2017. The discussion was guided by industrial and academic experts, with the audience as a key part of the discussion, trying to find some answers in three areas: data acquisition and evaluation of experimental uncertainties, tools for data reconciliation to improve their quality, and impact of data uncertainties on the process at the end.

Several concrete stories are presented that demonstrate the importance of considering data quality and all possible contributions to the uncertainty of chemical process design. Difficulties associated with data quality are discussed at various levels: (1) the experimentalists (measurement issues, evaluation of uncertainties, use of consistency analysis tools); (2) model developers (capture of adequate physics, parameter regression strategies, uncertainty propagation), (3) vendors of process simulation software, and (4) process engineers (who are responsible at the end).

Paths for improvements were proposed through better and more efficient communication among different participants, as well as through education.

1. Introduction

Uncertainty assessment and analysis as well as the impact of the uncertainty concept on thermodynamic property modeling have become central issues for scientists and researchers in the field of thermodynamics and chemical engineering lately (Kim et al., 2013; Magee et al., 2015). These uncertainties result from a combination of factors, starting from data, through the use of a thermodynamic model, all the way to models for process units in a process simulator.

Recently, there has been an increased focus on the quality of measured data, on the recognition and quantification of errors, where they

* *Corresponding author.*
E-mail address: j-charles.de-hemptinne@ifpen.fr (J.-C. de Hemptinne).

https://doi.org/10.1016/j.cherd.2018.08.010
0263-8762/

occur, how they are further transmitted and what their impact might be, not only on thermodynamic property calculations, but also on process simulations (Frutiger et al., 2017). The uncertainty concept has not received the deserved attention of process engineers in the past. Several research studies evaluating the impact of uncertainties in property models on process design and plant operability were performed in the past with different outcomes (Hajipour et al., 2014a,b; Watanasiri, 2011; Diky et al., 2013) and references 43–48 in Diky et al. (2013). In some cases, rough approximations of physical properties are good enough for process design, but sometimes extreme sensitivity of design to physical properties is observed. As stressed by Mathias (2014), uncertainty analysis is still not a routine task of today's industrial practice. Mathias (2014) suggested that this is mainly due to the lack of education and awareness of process engineers and due to the fact that the methods are quite difficult to apply. It is concluded in Kim et al. (2013) that two barriers must be overcome in order to implement uncertainty analysis in everyday chemical engineering practice: education on uncertainties should be introduced in undergraduate and graduate courses, and easy to use methods should become available in process simulators.

These issues were highlighted during a Round Table Discussion on the topic 'Data Quality and Assessment, Validation Methods and Error Propagation through the Simulation Software' that was organized during the 10th World Congress of Chemical Engineering (WCCE10) in Barcelona (October 1st–5th, 2017) as a joint action between two Working Parties of the European Federation of Chemical Engineers (EFCE), namely the Working Party on Thermodynamics and Transport Properties and the Working Party on Fluid Separations. The organizers of the Round Table were A. Soto (Spain), L. Fele-Zilnik (Slovenia), J.-C. de Hemptinne (France). Paul M. Mathias (USA) served as the chair of the discussion panel. Six panelists from both Working Parties were asked to contribute to the discussion with some suggested topics given at the beginning of the Round Table. The content of the discussion, the conclusions, and main messages are presented here.

Three questions were in the center of the Round Table Discussion:

- Data acquisition: how are experimental uncertainties evaluated?
- Data reconciliation: what tools can be used to improve the quality of existing data (*e.g.* consistency analysis, use of predictive models)?
- Impact of data on process simulation: what data have the most impact on the process, and how can their effects be quantified?

The purpose of this document is to report on the main lessons learned from the discussion. It is presented in three steps: first with some examples showing the importance of the issue, second by describing the difficulties encountered by the various types of actors (experimentalist, process engineers, *etc.*), and finally by suggesting some paths forward. The conclusion summarizes the outcomes.

2. Some concrete stories

Several examples which demonstrate the importance of considering data quality and all possible contributions to the uncertainty of process design were presented.

An example of a wrong decision made because of a combination of insufficient or wrong phase equilibrium data with improper interpretation of column parameters, was the explosion at a butadiene distillation facility in Texas City in 1969 (Jarvis, 1971; Freeman and McCready, 1971; King, 1990). In order to repair the stripper compressor, the distillation unit was placed on total reflux. However, the column was slowly losing material because of a leaking valve in the overhead line. Vinylacetylene is the most dangerous impurity to be separated in the process, since it becomes explosive above a certain concentration. Based on the fact that it has the highest boiling temperature among all components in the distillation column, its concentration was monitored at the column base and was shown to be below the hazardous level, while aftermath modeling showed non-ideal mixture behavior with the highest concentration expected higher up, between 10th and 15th trays. That is exactly where the explosion happened.

One panelist provided another example, although references were not provided (Peters). This incident related to the release of gas from a high pressure (80 MPa) gas condensate reservoir. Such a release follows the Joule–Thomson effect. The process simulator calculated a temperature drop, while in practice a temperature increase was observed. The main reason for such dramatic disagreement was that parameters of an equation of state (EOS) were fitted to a wrong data set.

Also, when comparing working fluids for thermodynamic cycles, the net power output is the key relevant quantity. When one takes into consideration its 0.95 confidence interval, for each fluid, additional quantitative information becomes vital for the fluid selection. The ranking of working fluids can be significantly different based on whether the mean value of the net power output is used as a criterion, or, alternatively, whether uncertainties (*e.g.*, the lower bound of the 0.95 confidence interval) are incorporated (Frutiger et al., 2016).

A final example mentioned by the panel concerns the use of temperature-independent binary interaction parameters for vapour–liquid equilibrium (VLE) computations in the design of a pressure swing absorption tower. This may lead to significant errors in the estimation of the number of theoretical stages of the column.

The panel concluded that all these examples clearly demonstrate the effect of uncertainty on the quality of the design and the ability to anticipate unsafe plant operation.

3. Difficulties at different levels

3.1. *For experimentalists: property measurements*

In spite of very rich existing databases, there is still a lack of property data for a large number of chemical systems of interest. Databases are typically designed for understanding particular phenomena, to develop theoretical models, or to respond to specific needs for parameter regression. All participants emphasized the need to continue acquisition of property data necessary for chemical process design (such as thermochemical, thermophysical, transport properties, safety parameters, *etc.*).

3.1.1. *Data acquisition and uncertainties*

Two journal editors (from the Fluid Phase Equilibria and the Journal of Chemical and Engineering Data) participated in the panel. They highlighted two crucial observations from their editorial experience:

1. The quality of experimental data submitted for publication does not always meet the desired standard, and few experimentalists provide realistic and justified uncertainty information.
2. These problems have been recognized and at least partially mitigated by the collaboration between the National Institute of Standards and Technology (NIST) and five journals that publish thermodynamic and thermophysical property data. The editors emphasized that experimental papers in the journals have improved following the project's start in 2009 (Cummings, 2009), and are expected to continue to improve as a result of this fruitful collaboration.