

Development of on-demand critically evaluated thermophysical properties data in process simulation*

Suphat Watanasiri

Aspen Technology, Inc., 200 Wheeler Road, Burlington, MA 01803, USA

Abstract: Accurate thermophysical properties are essential to the development of high-quality process simulation models of chemical processes. Therefore, process-modeling software (simulator) must provide accurate, reliable, and easily accessible property data and models to enable efficient and robust process design. Property data and parameters for components of interest are generally available in the databases of the simulator. For components that are not in the databases, their property data must be supplied by the user. The number of components available in a typical simulator is about 1700. The number and types of components available in the simulator limit the scope and accuracy of process models that can be developed.

In this paper, we review past practice in obtaining the necessary property data required in developing a process model and describe a new methodology that can be used to overcome the shortcomings of the current method. The new method is based on the dynamic data evaluation concept that combines the experimental data obtained from a comprehensive electronic database with structure-based property estimation system and data analysis and regression programs to generate critically evaluated property data. The concept and necessary software have been implemented in a process simulator, resulting in a new workflow that enables high-fidelity process models to be developed more easily and efficiently.

Keywords: critically evaluated property data; dynamic data evaluation; process simulation; ThermoData Engine; thermophysical properties.

INTRODUCTION

The 21st century is proving difficult for companies in the continuous processing industries (CPIs) due to wild fluctuations in energy, raw material and capital costs, globalization of markets, and the challenging economic environment. In response, chemical manufacturers are increasing their focus on production of higher-value specialty chemicals and intermediates. Pharmaceutical companies are developing synthesis routes for a wide slate of products while reducing production costs for their drugs. New processes are being developed to manufacture biofuel and to synthesize liquid fuel from natural gas and synthesis gas from coal, coke, oil sands, and biomass. Oil companies are working to improve yields through better characterization and “molecule management” in their facilities.

In addressing these challenges, the industries are increasingly relying on process modeling tools to design, operate, and optimize chemical processes to maximize yields, minimize energy and capital costs, and ensure environmental compliance. These software tools combine basic engineering princi-

*Paper based on a presentation made at the 21st International Conference on Chemical Thermodynamics (ICCT-2010), Tsukuba, Japan, 1–6 August 2010. Other presentations are published in this issue, pp. 1217–1281.

ples (e.g., mass and energy balances, phase and chemical equilibria, and reaction kinetics) with sophisticated computer algorithms to enable engineers to predict or simulate the behavior of chemical processes. One such tool is a process simulator such as Aspen Plus® [1].

Accurate thermophysical properties are keys to generating simulation results that can be used with confidence. Inaccurate or incorrect physical properties lead to poor design or plants that do not operate as expected resulting in enormous economic loss and environmental harm. Therefore, process simulator must provide accurate, reliable, and easily accessible property methods and data to enable efficient and robust process design. The problem is more challenging when dealing with new chemicals where no or little property information is known.

Property data and parameters for components of interest in the process model are generally available in the databases that are supplied with the simulator. The number of components available can be relatively small compared to the tens of thousands of compounds that may be of interest to the process industries. For components that are not available, their property data must be supplied by user. This can be accomplished using a combination of experimental data (from the literature or from their own measurements), estimation from group-contribution methods [2] and data regression [3]. However, this exercise requires significant know-how and can incur significant costs. It is evident that the number and types of components available in the simulator limit the scope and accuracy of process models that can be developed. Therefore, it is important to increase the scope of the components and mixtures available in the process simulator.

A property database is traditionally created by physical property experts who developed the data over an extended period of time, and the data are then deployed in a static database. The database may be updated periodically as resources allow. To significantly increase the scope of the components and mixtures available for process design and development, this static approach is not practical due to a number of serious shortcomings. The time and effort required to collect, analyze, fill gaps, and evaluate the data is very significant. The expertise required is limited to a small number of groups (universities and research institutions, large industrial companies, government laboratories). There is also a significant lag between the time the static database is developed and when it is available in the simulator.

In this work, a new method is developed to help overcome these shortcomings. It is based on the dynamic data evaluation concept [4] developed at the Thermodynamic Research Center (TRC), which is part of the U.S. National Institute of Standards and Technology (NIST). The concept combines the experimental data obtained from a comprehensive, evergreen electronic database with structure-based property estimation system and data analysis and regression programs to generate critically evaluated property data. The concept and necessary software have been implemented in Aspen Plus [1] and are described in detail in this paper. The new technology enables creation of critically evaluated properties required in simulation based on the current state of the data and process requirements. This technology is a result of a government/industry cooperation that leveraged more than a decade of research and development work performed at NIST-TRC. Combined with many years of development efforts at the author's company, the technology is available to a very large number of process simulation users worldwide.

THERMOPHYSICAL PROPERTY DATA AND PROCESS SIMULATION

Process modeling uses the laws of thermodynamics and chemistry and chemical engineering principles to carry out mass and energy balances for each unit of equipment in a plant. The fundamental equations used to predict phase equilibrium or to size and rate equipment require reliable thermophysical property data of the chemical components and mixtures that are involved. For example, phase equilibrium, which dictates the distribution/separation of species in distillation columns, depends on the component's vapor pressures and activity coefficients (or fugacity coefficients), while heat transfer depends on the component's thermal conductivity and heat capacity. The accuracy (or fidelity) of the process simulation models is directly related to the accuracy of the underlying thermophysical property informa-

tion. Poor data, poorly estimated property parameters, or inappropriate model can lead to very large errors in the calculated capital costs of the plant, which in turn can lead to poor investment decisions. For example, a 20 % error in thermal conductivity can lead to a 13 % error in capital cost, while 20 % error in density can lead to 16 % error in capital cost [5]. For phase equilibrium property, such as activity coefficients, the impact can be even larger, depending on the characteristic of the system involved, for example, the relative volatility of the key components. Consider a close boiling system, with relative volatility of 1.10. In order to separate the two components via distillation, a large number of stages are required. Assuming a total reflux and a purity specification, the minimum number of stages can be determined. For this system, a column consisting of 140 stages is required with an estimated capital cost of USD 18 million. Figure 1 shows a sensitivity study relating the error in the relative volatility to the minimum number of stages required to achieve the same purity specification. An error of -5% can result in a doubling of the number of stages required [6]. This very simplified example illustrates the potential high costs of inattention to the accuracy of the physical property model used to design any process.

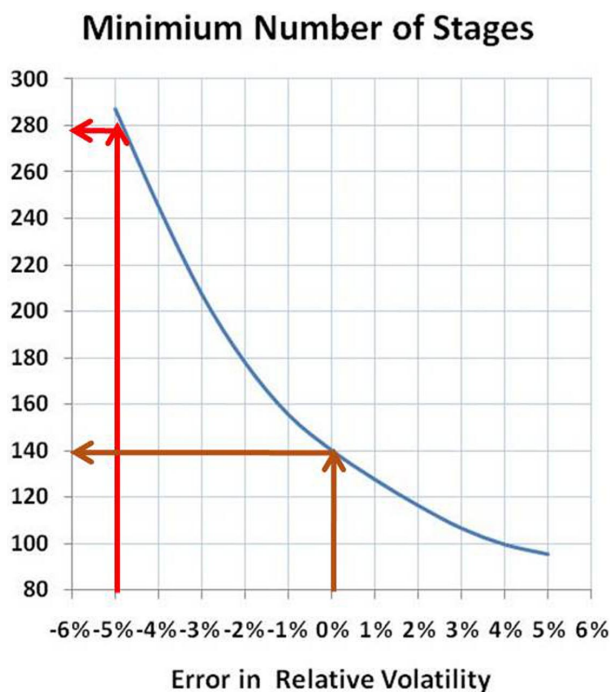


Fig. 1 Effect of error in relative volatility on the minimum number of stages for a distillation column required to achieve a desired separation of a close boiling mixture (with baseline relative volatility = 1.10). Adapted from [6].

In addition to the direct effect on capital costs, a model based on inaccurate or incorrect physical properties could lead to infeasible process designs, poor plant operability, or unsafe plants. Thus, it is absolutely essential that process simulation models are based on accurate thermophysical property data. Finally, chemical industry trends toward new chemicals and new chemical families present an enormous challenge to providing the necessary property information in a timely and cost-effective manner.

Figure 2 shows schematically the relationships between the thermophysical property database, the physical property calculation engine, and the process simulator engine. These are highly complex components, therefore, only the essential elements will be briefly described here to provide context.

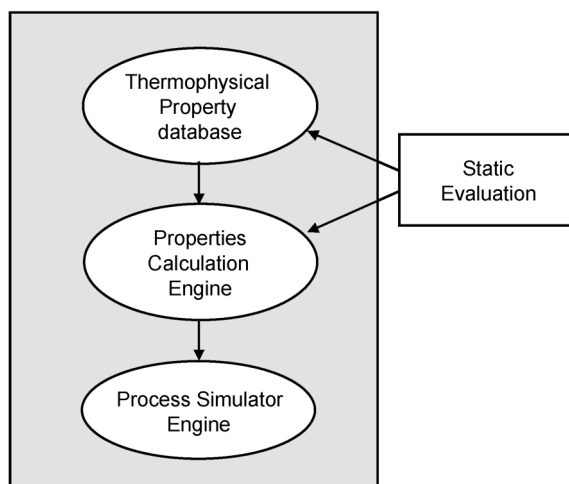


Fig. 2 Simplified diagram showing relationships among property database, property calculation engine, and process simulator engine, which constitute the process simulation domain (shaded area). The static data evaluation process resides outside the process simulation domain, resulting in inefficiency in providing the thermophysical property data needed in process simulation.

The process simulator engine comprises algorithms to solve the distillation column; perform compressor, heat exchanger, and other unit operations calculations; tear and converge recycle streams; and optimize the process subject to constraints, including economic (e.g., equipment costs) constraints. The simulator requires thermophysical properties, which are supplied by the physical property calculation engine. The simulator provides the process conditions (temperature, pressure, flow rates, and components involved), property requests (properties and phase of interest), and the methods and models (correlations) to be used, such as a specific equation-of-state model to the property calculation engine. The property calculation engine in turn calculates and returns the requested properties.

The property calculation engine calculates the requested properties using the specified methods and models. The key properties needed by the simulator are fugacity coefficients, enthalpy (and heat capacity), entropy, Gibbs free energy, molar volume, viscosity, thermal conductivity, surface tension, and diffusion coefficients. The property engine must also calculate derivatives of these properties with respect to temperature, pressure, and composition in order to support Newton-based algorithms. For an equation-based simulator (such as [1]), analytical derivatives of properties are essential. In order to serve a wide range of industries (oil and gas, chemicals, petrochemicals, pharmaceuticals, specialty chemicals, polymers, etc.), a wide variety of models capable of representing the components and complex mixtures present in these diverse processes must also be available. The models include a large number of equations of state, activity coefficients, corresponding-states, and empirical models. The property engine also includes flash (phase equilibrium) calculations that can handle vapor, liquid, and solid phases, including occurrence of multiple liquid phases and solid solution. In many applications, such as electrolytes or reactive distillations, both phase and chemical equilibria must be considered simultaneously.

Property models require property data and/or parameters as input. For example, the Peng–Robinson equation of state [7] requires critical temperature, critical pressure, and acentric factor of pure components and a binary interaction parameter for each component pair. Pure-component property such as vapor pressure depends only on temperature and is calculated using correlations, such as the Antoine [8] or Wagner [9] equations. The Antoine equation depends on three compound-specific

constants, while the Wagner equation depends on the critical temperature and critical pressure as well as four compound-specific constants.

These pure-component critical properties and correlation constants, and binary interaction parameters are stored in the thermophysical property database of the simulator. It is important to note that in general, this type of database does not contain the “raw” experimental data—just the evaluated correlation constants and model parameters.

A partial list of properties required in process calculations is given in Table 1. For each property, examples of equations or models are given. Only a small number of models are shown for illustrative purposes as the actual number of models available for each property can be quite large. Interested readers can refer to the on-line help feature of the simulator [1] for the complete list and detailed descriptions of available models.

Table 1 Partial list of properties data required in process simulation [1].

Pure-component	
Single-valued properties	
Critical temperature	
Critical pressure	
Critical volume	
Critical compressibility factor	
Triple-point temperature	
Enthalpy of fusion	
Normal boiling temperature	
Pitzer acentric factor	
Temperature-dependent properties	Model equations
Vapor pressure	Antoine [8], Wagner [9]
Saturated liquid density (or molar volume)	Rackett [11]
Enthalpy of vaporization	Watson [12]
Ideal gas heat capacity	Aly–Lee [13]
Liquid viscosity	Andrade [14]
Liquid thermal conductivity	TDE polynomials [1]
Surface tension	TDE expansion [1]
Mixtures	Mixture models
Fugacity or fugacity coefficients	Peng–Robinson [7], PSRK [15], NRTL [16], UNIQUAC [17]
Molar enthalpy (and heat capacity)	Compute from EoS and ideal gas Cp
Molar entropy	Compute from EoS and ideal gas Cp
Molar Gibbs free energy	Compute from EoS and ideal gas Cp
Molar volume	Compute from EoS, Rackett [11]
Viscosity	Compute from pure-component property using mixing rules
Thermal conductivity	Compute from pure-component property using mixing rules
Surface tension	Compute from pure-component property using mixing rules
Diffusion coefficients	Wilke–Chang [18]

The number of components available in a typical simulator is relatively small (~1700 [10]) compared to the tens of thousands of compounds that may be of interest to the process industries. Aspen Plus version 2006 [1] (before the implementation of the technology described in this work) contains a relatively larger set of compounds as shown in Table 2. However, that is still insufficient for many applications.

Table 2 Databases available in Aspen Plus version 2006 [1] prior to implementation of the technology described in this work.

Data in the database	Source of data
1972 organic components	DIPPR® [19] and in-house development
2477 inorganic components	Barin [20]
3462 solids	Wagman et al. [21]
800 ions	Wagman et al. [21]
Binary interaction parameters for ~4000 unique mixtures	Various literature sources [22,23] and in-house development

Property databases are traditionally created by physical property experts [19–23]. A comprehensive review of the developments of factual databases was given by Rarey and Gmehling [24]. These databases are developed over an extended period of time (years) by collecting experimental data from thermodynamic journals, reference books, and open and/or private on-line databases. For many compounds considered, there are only limited experimental data available. As a result, the missing data must be estimated from molecular structure using group-contribution or other methods [2]. The experimental and estimated “data” are evaluated for consistency to produce recommended values (see “Pure-component properties” section for more information on consistency checks). For temperature-dependent properties, the data are then fitted to the appropriate equations using data regression. This exercise is time-consuming and requires significant know-how. The expertise required is limited to a small number of groups (universities and research institutions, large industrial companies, government laboratories). In addition, experienced property evaluators continue to become rarer, particularly as the baby-boomer generation has reached retirement age.

There are no comprehensive libraries of binary interaction parameters that are publicly available. These parameters are generally developed in-house or collected from various literature sources (e.g., [22,23,25,26]). The efforts to develop binary interaction parameters in-house are very significant. Our experience with such development consumed over 10 person-years. The resultant library covers about 4000 unique binary mixtures and includes binary interaction parameters for the NRTL [16], Wilson [27], and UNIQUAC [17] models with three combinations of vapor-phase equations of state (ideal gas, Redlich–Kwong [28] and Hayden–O’Connell [26]). This results in about 36 000 sets of binary parameter values [1]. There is no doubt as to the necessity and benefits of such efforts and investments. However, this example underscores the magnitude of the work required for such an undertaking and the very large gaps that currently exist—considering that there are vapor–liquid equilibrium (VLE) data for about 30 000 unique binary mixtures in the NIST SOURCE Data Archive (as of 30 June 2010; see also “Experimental data section”). Any efforts to systematically analyze such data with the aim to develop a library of binary interaction parameters will be very extensive and costly. In addition, other data such as excess enthalpy and infinite-dilution activity coefficients should also be included in the parameter optimization process, thus increasing the scope of the work further. The database of binary interaction parameters is not limited to VLE systems for activity coefficient models. Binary interaction parameters for liquid–liquid equilibrium (LLE) applications, equation-of-state models, and Henry’s law constants are also available [1]. These databases should also be extended to make use of the currently available experimental data. Such systematic and inclusive efforts, if executed manually, currently appear imprac-

tical and will inevitably result in a certain number of systems having been poorly determined or overfitted, as pointed out by Rarey and Gmehling [24].

Property databases discussed above are delivered with the process simulator as static databases. Some of the databases may be updated with each new release of the software, but many are not due to the limitations discussed above. There is also a significant lag between the time the static database is developed and when it is available in the simulator, thus delaying its availability for industrial use. The static data evaluation process shown in Fig. 2 depicts the crucial property evaluation process to reside outside the domain of the process simulator resulting in inefficiencies.

Static data evaluation is also used to supply property data directly to the property calculation engine as shown in Fig. 2. This occurs when components present in the simulation are not available in the library so the user has to supply the data needed. This task has traditionally been performed by thermodynamics experts within the company and to an increasing extent by other users of the program. The same methodology employed by database developers is also used here. However, an additional impediment in this case lies in the accessibility to experimental data due to the high costs involved as well as to the lack of awareness of the available resources. In addition, most users are not thermodynamics experts; they lack the knowledge and confidence to perform the evaluation. As a result, the property data supplied to the simulator may contain errors or can be inconsistent. Given the increasing importance of new materials and processes, this challenge must be addressed.

From the preceding discussion, it is clear that the number and types of components available in the simulator limit the scope and accuracy of process models that can be developed. Therefore, it is important to increase the scope of the components and mixtures available as well as to improve the ability to handle new components that are not present in the database. However, it is also evident that the static evaluation approach alone is not sufficient and practical. Therefore, we proposed to address the problems in two ways: (1) by implementing the dynamic data evaluation concept and associated software, and (2) by greatly expanding the static database for pure components using the experimental data and evaluation tools developed for dynamic data evaluation.

DYNAMIC DATA EVALUATION

An alternative to the static evaluation approach lies in the concept called dynamic data evaluation [4] developed at NIST and in bringing both the experimental data and dynamic data evaluation inside the process simulation domain, as shown in Fig. 3. The new methodology solves the primary deficiencies of the static approach by (1) removing the impediment to accessing experimental data by providing them directly within the process simulator, (2) offsetting the lack of data evaluation expertise by providing software that captured such expertise and using it to perform data evaluation, and (3) enabling the user to initiate data evaluation as needed, dynamically—on demand when developing the property data for a process model. The data evaluation results can be used directly in property calculations as shown in Fig. 3. Over time, as results are validated in the field, the data and parameters can also be permanently stored in the database.

The three key elements of dynamic data evaluation are described in the following sections.

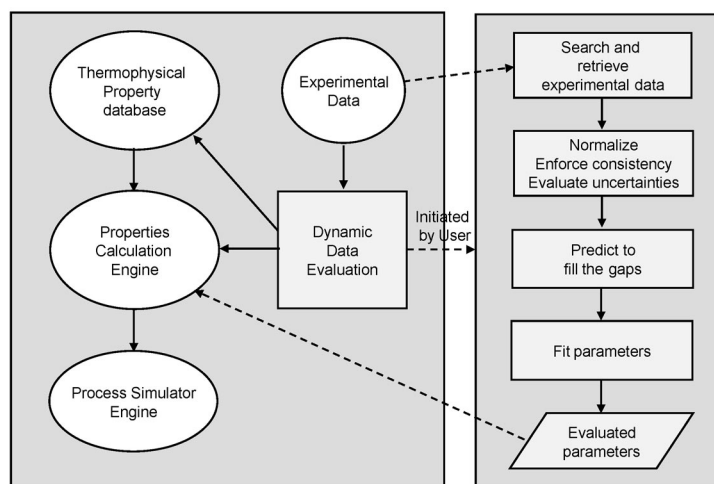


Fig. 3 Schematic diagram showing the key elements of dynamic data evaluation. The dynamic data evaluation process resides within the process simulation domain (shaded area on the left of the figure); it uses an extensive database of experimental data together with expert software (shaded area on the right) to provide evaluated property parameters required in a particular process simulation project. The data evaluation is initiated by the user when needed—on demand.

Experimental data

Implementation of the dynamic data evaluation concept has required the development of a large electronic database storing essentially all experimental thermodynamic data known to date with detailed descriptions of relevant metadata and uncertainties. The assessment and storage of the uncertainties are essential to the success of the *dynamic* approach. This database is referred to as the *NIST SOURCE Data Archive* (or SOURCE in the remainder of this paper). The current state of SOURCE is shown in Table 3. Experimental data are continually and systematically collected and evaluated at the NIST TRC's Data Entry Facility in Boulder, Colorado, USA. The staff of this facility, working in conjunction with the University of Colorado and the Colorado School of Mines, gathers property data from contemporary and historical scientific journals from around the world. The data are captured and uncertainties are assessed with the aid of an expert system called the *Guided Data Capture* program [29]. This expert system checks the data using various thermodynamic relationships to ensure the assessed uncertainties are reliable and that the data are accurate and self-consistent. This process eliminates many of the transcription errors associated with traditional property data collection, and ensures the integrity of SOURCE that underlies the dynamic data evaluation approach. Approximately 500 000 data points are added to the database annually.

The SOURCE database can be integrated into the process simulator domain as depicted in Fig. 3. However, due to the nature of the software release cycle (typically annually), SOURCE that is being used in the simulator is not always up to date. NIST has developed a web-based data deployment mechanism, which allows data to be updated daily [4]. In principle, the process simulator can take advantage of this web-based update mechanism to keep SOURCE up to date for the users. However, such an update mechanism from the NIST web site introduces security risks and can incur high administrative costs related to entitlement verification. As a result, it was not adopted in this work. An alternative approach was therefore developed. The most current SOURCE data are provided to the users via an update posted to the company customer support web site every three months. Interested users with current software support entitlement can download and install the SOURCE database quarterly. This practical compromise provides relatively up-to-date data to the users at minimum cost.

Table 3 Status of experimental data stored in the SOURCE data archive, as of 30 June 2010. A data point refers to a measured property of interest, e.g., mixture liquid viscosity at a given temperature, pressure, and composition.

Data description	No. of data points
VLE	817 568
LLE	142 394
Activity coefficients at infinite dilution in pure solvents	26 636
Activity coefficients at infinite dilution in mixtures	5 169
Critical data of mixtures	14 121
SLE	110 330
Azeotropic data	4 942
Excess enthalpies	263 541
Excess heat capacities	46 913
Excess volumes	693 379
Pure-component data	1 196 181
Reaction heats	5 662
Reaction equilibrium data	8 229
Mixture transport properties	258 284
Mixture surface tension, refraction, speed of sound	213 862
Total number of data points	4 025 274

Dynamic data evaluation software

The dynamic data evaluation concept was implemented in the software program *NIST ThermoData Engine* (TDE) and had been described in details in refs. [4,30,31]. As discussed earlier, the goal is to alleviate the lack of data evaluation expertise by providing software that captured such expertise and using it to perform data evaluation. TDE makes use of the experimental and metadata from SOURCE together with software heuristics to evaluate the data for consistency, estimate missing properties to fill gaps, and fit the data to produce critically evaluated properties and parameters that can be used directly in process calculations as shown schematically in Fig. 3. Dynamic data evaluations for pure components and binary mixtures are described in the following sections.

Pure-component properties

Pure-component properties of interest in process simulation are listed in Table 1. The algorithm used in pure-component property evaluation is shown in Fig. 4 and is described in details in ref. [4]. A brief description of the algorithm is given below.

Experimental pure-component property data are retrieved from the database and are divided into blocks or types: phase diagram (e.g., phase boundary pressure, critical temperature), volumetric (e.g., density, molar volume, critical density), energetic (e.g., enthalpy of vaporization or sublimation, heat capacity at constant pressure), and other (e.g., viscosity, surface tension). Phase diagram properties are used to delineate the phase regions and their boundaries. The phase diagram, volumetric and energetic properties are tied by thermodynamic consistency conditions—that is, they are related through mathematical thermodynamic identities. Properties in the “other” block have no influence on the first three blocks of properties and are processed last.

Normalization refers to the process of combining closely related properties into single properties. For example, specific density, molar volume, and compressibility factor are combined into one property. This eliminates the complexity and confusion that will be caused by applying separate equations for each of these related properties.

For each block, after normalization, missing properties are added from predictions based on group-contribution, corresponding-states, and other methods; appropriate methods are selected based

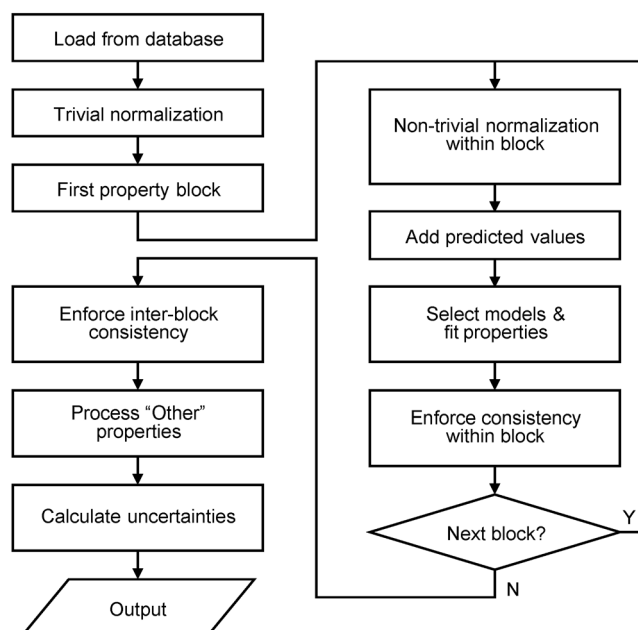


Fig. 4 Algorithm for dynamic data evaluation of pure-component properties [4].

on the nature of the chemical system and property; data are fitted to the selected equations; and consistency within block is enforced. When all three blocks are processed, then inter-block consistency is enforced. Finally, the “other” properties, which depend on the first three blocks, are processed.

Three levels of consistency enforcement are used: single-property enforcement, in-block enforcement, and inter-block enforcement [4]. Single-property enforcement constrains a single fitting equation to be consistent with other properties. This is used to ensure that saturated vapor pressures are consistent with heat capacity differences between the liquid and gas phases at low pressures and to ensure that condensed-state phase boundary lines converge at triple points. In-block procedures are used for all vapor and sublimation pressures as well as for saturated liquid and gas densities and single-phase and saturated gas densities. An inter-block procedure refines the gas density and enthalpy of vaporization through improved consistency with vapor pressures and liquid densities [4].

Table 4 lists the estimation methods used in TDE. The methods range from the familiar Joback [35] and Benson [45] group-contribution methods to the work by Nannoolal and co-workers [34] and the recent quantitative structure–property relationship (QSPR) work of Kazakov and co-workers [37]. Most of these methods are group-contribution-based, which require functional groups for the compounds to be known. TDE contains an algorithm that automatically maps the molecular structure obtained from SOURCE (or supplied by user) into the functional groups for the different estimation methods involved.

For compounds for which limited or no experimental data are available, the thermodynamic constraints and property evaluation procedure described above still apply. However, group-contribution and corresponding-state methods are relied upon to provide the required information, instead of the experimental data. As a result, the uncertainties of the evaluated properties are relatively larger.

For pure-component properties, dynamic data evaluation can be performed unattended, without user intervention. This is possible because (1) the rules involved in performing thermodynamic consistency test have been implemented and refined over many years of field use, (2) automatic generation of groups from molecular structure has been successfully implemented, (3) extensive estimation methods

exist to fill the data gaps, and (4) data regression of pure-component properties is relatively straightforward to carry out.

Table 4 Estimation methods available in ThermoData Engine [4].

Property	Method
Normal boiling point	Constantinou–Gani [32], Marrero–Pardillo [33], Nannoolal et al. [34]
Critical temperature	Modified Joback [35], Constantinou–Gani [32], Marrero–Pardillo [33], Wilson–Jasperson [36], NIST/TRC QSPR [37]
Critical pressure	Modified Joback [35], Constantinou–Gani [32], Marrero–Pardillo [33], Wilson–Jasperson [36], NIST/TRC QSPR [37]
Critical density	Modified Joback [35], Constantinou–Gani [32], Marrero–Pardillo [33]
Density (saturated liquid)	Modified Rackett [11], Riedel [38]
Heat capacity (ideal gas)	Modified Joback [35]
Heat capacity (saturated liquid)	Modified Bondi [39]
Vapor pressure	Ambrose–Walton [40]
Viscosity (gas)	Lucas [41]
Viscosity (saturated liquid)	Sastri-Rao [42]
Thermal conductivity (gas)	Chung [43]
Thermal conductivity (saturated liquid)	Chung [44]
Enthalpy of formation (Ideal gas)	Benson [45]

Binary mixture properties

Experimental property data of binary mixtures are needed in the development and validation of mixture property models (correlations) used in process simulation. The most important properties are phase-equilibrium properties—VLE, LLE, and solid–liquid equilibrium (SLE) data. These data are used to determine binary interaction parameters for either an activity coefficient or an equation-of-state model. Excess enthalpy and infinite-dilution activity coefficients data are also often used to improve the quality and extend the application range of these models. The need for experimental data is not limited to those related to phase equilibrium. To develop an accurate model for other properties such as liquid viscosity, binary liquid viscosity data are required. Mixture properties available in SOURCE are shown in Table 3.

Dynamic data evaluation of binary mixture data follows similar steps as those described in Fig. 4 for pure components—that is, retrieval of the data from SOURCE, assessment of data quality, and finally fitting the data to an appropriate mixture model. However, due to the large amount of data and the complexity involved in fitting nonlinear mixture model to data, especially when more than one type of data are involved (i.e., multi-property regression analysis), dynamic data evaluation cannot be performed automatically (i.e., unattended). User intervention is required. After retrieval of the binary data, quality of the VLE data can be assessed (see “Quality assessment of binary VLE data” section). For other types of data, their quality beyond the uncertainty cannot yet be assessed. Finally, appropriate data are used to determine binary interaction parameters of mixture model via regression.

Quality assessment of binary VLE data

The most important data for vapor–liquid processes are binary VLE data. As shown in Table 3, currently there are more than 800 000 VLE data points in SOURCE. These data comprise about 30 000 unique binary mixtures. The challenge facing dynamic data evaluation of binary VLE data is how to deal with this very large amount of data that are of varying quality and sometimes conflicting.

Kang and co-workers [46] developed a quality assessment algorithm for binary VLE data based on thermodynamic consistency tests. The algorithm combines four widely used tests of VLE consistency based on the requirements of the Gibbs–Duhem equation: Herrington or area test [47,48], point or differential test [49,50], infinite dilution test [49,50], and Van Ness modeling test [51,52], with a check of consistency between the binary VLE data and the pure-compound vapor pressures. Mathematical details of all tests are described fully in [46].

Kang and co-workers [46] also developed a numerical quality factor $F_{\text{test}i}$ for each test to replace the simple pass/fail results normally associated with these tests. Since the four tests evaluate different aspects of the data, conflicting results do often occur. Pass/fail results for these tests do not help users determine the overall quality of the datasets. Table 5 shows an example of the conflicting pass/fail results for four hypothetical VLE data sets of a binary mixture. For data set 1, tests 1 and 2 passed while tests 3 and 4 failed. For data set 3, the opposite is true. Given these results, it is not possible to compare the quality of the data sets.

Table 5 Hypothetical consistency test results for four sets of VLE data, including the overall quality factor, which can be used to provide quantitative results in the evaluation of binary VLE data.

Data set	Test 1*	Test 2*	Test 3*	Test 4*	Overall quality factor
1	Pass	Pass	Fail	Fail	0.4
2	Fail	Pass	Pass	Fail	0.3
3	Fail	Fail	Pass	Pass	0.8
4	Pass	Fail	Fail	Pass	0.5

*The four test methods are: 1. Herrington or area test [47,48], 2. Point or differential test [49,50], 3. Infinite dilution test [49,50], and 4. Van Ness modeling test [51,52].

$F_{\text{test}i}$ values for the four tests were combined with F_{pure} value of the consistency between the end-point pressures of the VLE data with pure-component vapor pressure assessed using the algorithm described in the “Pure-component properties” section to calculate an overall quality factor Q_{VLE} :

$$Q_{\text{VLE}} = F_{\text{pure}} (F_{\text{test}1} + F_{\text{test}2} + F_{\text{test}3} + F_{\text{test}4}); Q_{\text{VLE}} \leq 1$$

Q_{VLE} and the F factors were formulated such that if a particular test cannot be performed, the F factor for that test is set to $0.5F_{\text{test}i,\text{max}}$, where $F_{\text{test}i,\text{max}}$ is 0.25 for tests 1 to 4 and 1.0 for the pure-component test. If none of the tests can be applied, Q_{VLE} is 0.25.

$F_{\text{test}i}$ has lower and upper limits: $0.025 \leq F_{\text{test}i} \leq 0.25$. F_{pure} has lower and upper limits: $0.1 \leq F_{\text{pure}} \leq 1.0$. When all tests passed: $F_{\text{test}1,\text{max}} + F_{\text{test}2,\text{max}} + F_{\text{test}3,\text{max}} + F_{\text{test}4,\text{max}} = 1$. When all tests failed: $F_{\text{test}1,\text{min}} + F_{\text{test}2,\text{min}} + F_{\text{test}3,\text{min}} + F_{\text{test}4,\text{min}} = 0.1$. Application of the quality assessment to the same four hypothetical data sets theoretically results in a quantitative comparison of the data as shown in Table 5.

The algorithm is applicable to all non-reacting chemical systems at subcritical conditions. It can be applied to VLE data sets with at least three state variables reported: temperature, pressure plus liquid, and/or vapor composition. To apply the Gibbs–Duhem-based tests, p - T - x - y data are necessary. All four tests can be applied to isothermal data, while two (Herrington and Van Ness) can be applied to isobaric data. In the case of p - T - x or p - T - y data, only the pure-component consistency test can be applied.

The overall quality factor, Q_{VLE} provides very useful information to the user. It helps identify data sets that are thermodynamically consistent that should/can be used in the mixture model evaluation (i.e., testing accuracy of existing binary interaction parameters) and development (i.e., determining new binary interaction parameters). An additional benefit of Q_{VLE} is that it can be used to provide relative weight in data regression [46]. The four thermodynamic consistency tests and pure-component vapor pressure test as well as the overall quality factor, Q_{VLE} calculations have been implemented in TDE.

On-demand data evaluation

Dynamic data evaluation is initiated as needed by the user as shown in Fig. 3, usually at the start of a project. The most current experimental data are available through SOURCE for evaluation of built-in property models available in the simulator as well as for developing the pure-component and mixture models. TDE is the enabling technology that makes it possible to develop critically evaluated properties and data quickly by encapsulating the expert data-evaluation knowledge in the software. The dynamic nature of the data evaluation also means that the required thermophysical properties are developed as needed and only for the systems of interest, thus making the problem more manageable. This is in contrast to the daunting task of developing the static database in anticipation of all possible needs and broad requirements as described in the preceding section.

The combined pure-component and binary interaction parameters constitute the critically evaluated properties and data that can be used directly to calculate the physical properties required by the simulator (see Fig. 3).

IMPLEMENTATION IN THE PROCESS SIMULATOR

The static database for pure-component properties has been greatly expanded by using TDE and a script that enables automatic evaluations of all pure components available in the SOURCE database. The number of components increased from approximately 9000 to more than 30000 [1] over the last several years.

The dynamic data evaluation concept has been implemented in a process simulator, Aspen Plus [1], by integrating the essential elements of the TDE software using dynamic link library (DLL) and by delivering a snapshot of the SOURCE database with the product. The SOURCE database is kept up to date by an automatic update that can be obtained from the company customer support web site on a quarterly basis. TDE is integrated seamlessly into the user interface and the workflow of process model development. It can be accessed using either a menu option or a toolbar button as shown in Fig. 5. Availability of comprehensive experimental data in process simulator is a new development and, we believe, is currently unique among the major process simulation programs.

This section describes the new workflow resulting from the implementation of this new technology. The workflow follows the steps shown in Fig. 6 to develop critically evaluated property data required for a process model—normally at the beginning of a project. These are recommended steps or best practices that serve to illustrate how the new technology can be applied and is not meant to be a recipe. The user can deviate from the sequence as needed. It is also important to remember that dynamic data evaluation can be initiated by the user at any time. Finally, this section aims to provide an overview of the new technology, therefore, it does not describe many of the features of the physical property system of the simulator. Interested readers should refer to the online help feature of the simulator.

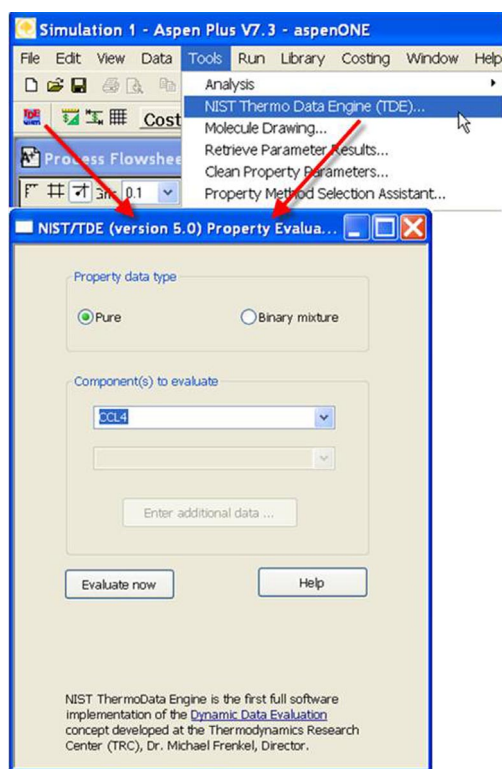


Fig. 5 NIST ThermoData Engine is integrated with the simulator and can be launched directly to perform property evaluation for pure-component and binary mixtures. The figure is shown for pure-component property evaluation. For binary mixture, two components must be specified.

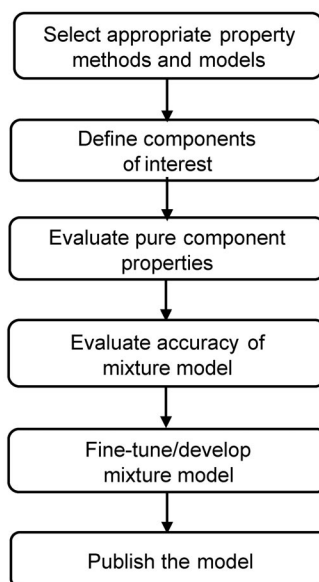


Fig. 6 Schematic diagram showing the key steps involved in developing accurate property models in process simulation.

Select property methods and models

Property methods and models most appropriate for the process of interest must be selected. Property method in this simulator refers to a collection of calculation procedures that is used to calculate all the required properties (see Table 1). For a given property, one or more specific models may be used. The Peng–Robinson equation of state is an example of a model and it is used to calculate fugacity coefficients, molar volume, enthalpy, entropy, and Gibbs free energy. One or more methods may be required to be able to model the entire process. For example, in the low-pressure section of the process, a property method based on an activity coefficient model may be used while in the high-pressure section, a property method based on an equation of state is more appropriate.

Property method selection is an important task that can be daunting for new users or for users who are not familiar with thermodynamics, especially when there are a large number of methods and models to choose from. For the simulator described in this work, there are more than 90 built-in methods and several hundreds models. To assist the users, the simulator offers a Property Method Section Assistant, which is rule-based software that provides suggestions of possible property methods to be used based on the answers that the user provided to a series of questions. For example, for general chemical processes, the assistant will recommend that an activity coefficient-based method, such as NRTL [16], Wilson [27], or UNIQUAC [17] can be used. However, if the process is at high pressures (e.g., >10 bar), an equation of state with advanced mixing rules, such as predictive Soave–Redlich–Kwong (PSRK) [15] should be used. If two liquid phases exist, a model capable of representing phase splitting must be used, such as an activity coefficient model (NRTL [16] or UNIQUAC [17]), or an equation-of-state model with advanced mixing rules (PSRK [15] or many others). Additional help is available for specific processes, such as azeotropic separations or those involving carboxylic acids or hydrogen fluoride. For carboxylic acids, the assistant will recommend that an activity coefficient model be used together with an equation of state that accounts of the effect of vapor-phase association such as Nothnagel [53] or Hayden–O’Connell [26]. For electrolytic systems, phase and chemical equilibrium must be considered, and an electrolyte model, such as the electrolyte NRTL model [54], is recommended.

Define components of interest

The user has access to a library of over 30 000 components as a result of the expansion of the static database described earlier. Flexible search helps locate the components quickly.

Evaluate pure-component properties

TDE can be used to evaluate properties of pure components. TDE is integrated with the simulator and can be launched directly using either a menu option or a toolbar button as shown in Fig. 5. The component of interest is then selected for evaluation.

Results of the evaluation include the recommended property parameters, the underlying experimental data, and the estimated properties that were used. Results of dynamic data evaluation for trimethylamine are shown in Fig. 7. The evaluated properties are shown on the left-hand side of the form. For each property, the critically evaluated parameter values are shown on the right-hand side. Figure 7 shows the parameter values of the Wagner equation for vapor pressure.

Evaluation results also include all the available experimental data, predicted values, and evaluated results. An example of the experimental data is shown in Fig. 8 for vapor pressure. Experimental values are shown with uncertainty and literature citation. Data uncertainty and citation provide a measure of quality and traceability to the original source. The form also indicates whether or not a data point is accepted or rejected by the evaluation algorithm. The experimental data and parameters can be saved with the simulation case file using the *Save* button.

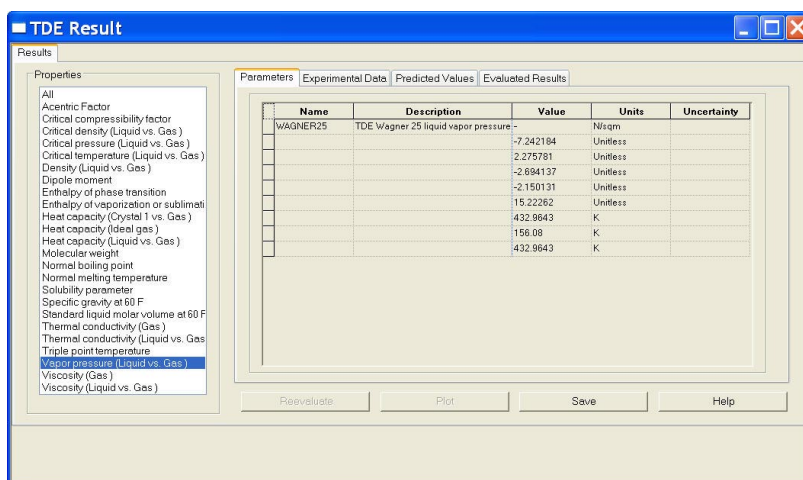


Fig. 7 Results of pure-component properties evaluation performed by TDE. Critically evaluated parameters for the vapor pressure model (Wagner [9]) for trimethylamine are shown. Results also include all available experimental data, predicted values used to fill gaps of experimental data, and evaluated results.

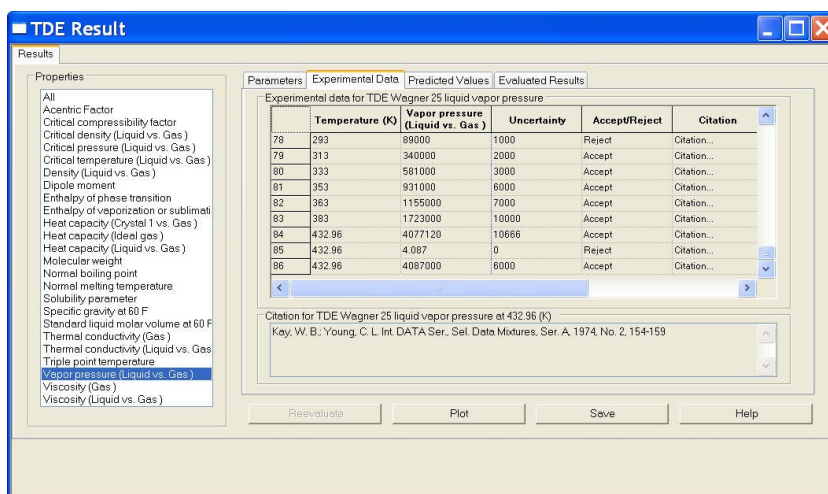


Fig. 8 Results of pure-component properties evaluation performed by TDE. Experimental vapor pressure data of trimethylamine are shown.

Model prediction can be compared against experimental data. Figure 9 shows a typical plot of the property vs. temperature. In this case, the natural logarithm of vapor pressure of trimethylamine is plotted vs. inverse temperature. Accepted experimental data, those that were rejected, predicted data, and the evaluation results (best fit to the Wagner equation) are shown.

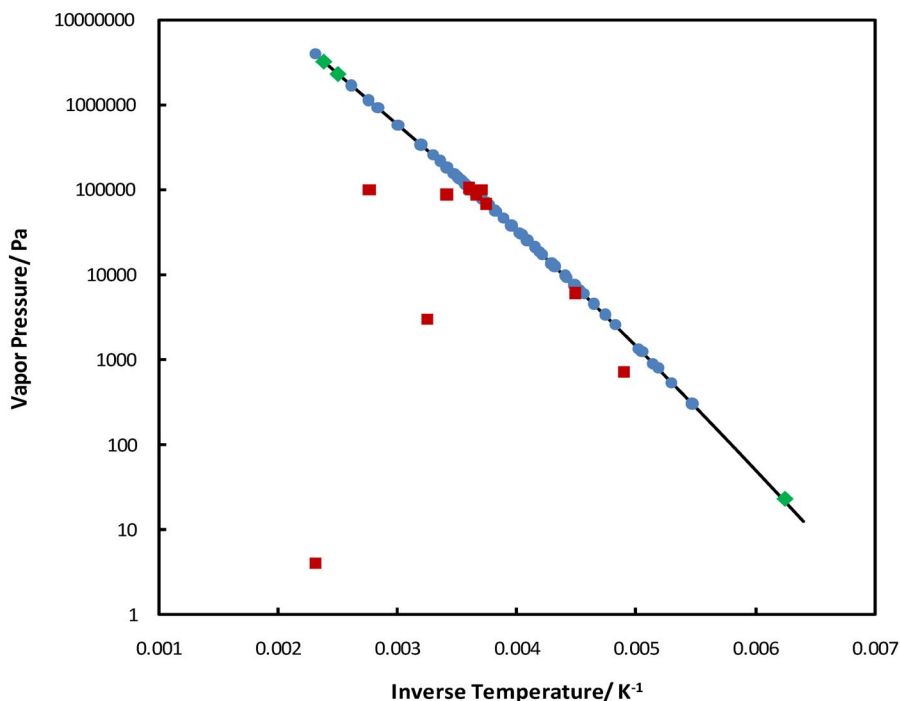


Fig. 9 Vapor pressure of trimethylamine as a function of temperature. ●: accepted experimental data; ■: rejected data; ◆: predicted data using Ambrose–Walton method [40]; —: evaluation results (best fit to the Wagner equation [9]).

As discussed earlier, it is important to improve the ability to handle new compounds that are not in the database. TDE can use molecular structure information imported from an MDL MOL file or drawn using a molecular drawing tool. The drawing tool was developed and included with the Aspen Plus simulator and can be invoked as shown in Fig. 10. As a result, in principle, TDE can be used to estimate the properties for *any* organic compound (including those which contain sulfur, fluorine, chlorine, bromine, and iodine), thus providing direct access to property values for literally millions of compounds, all with reliable and conservative estimates of uncertainty.

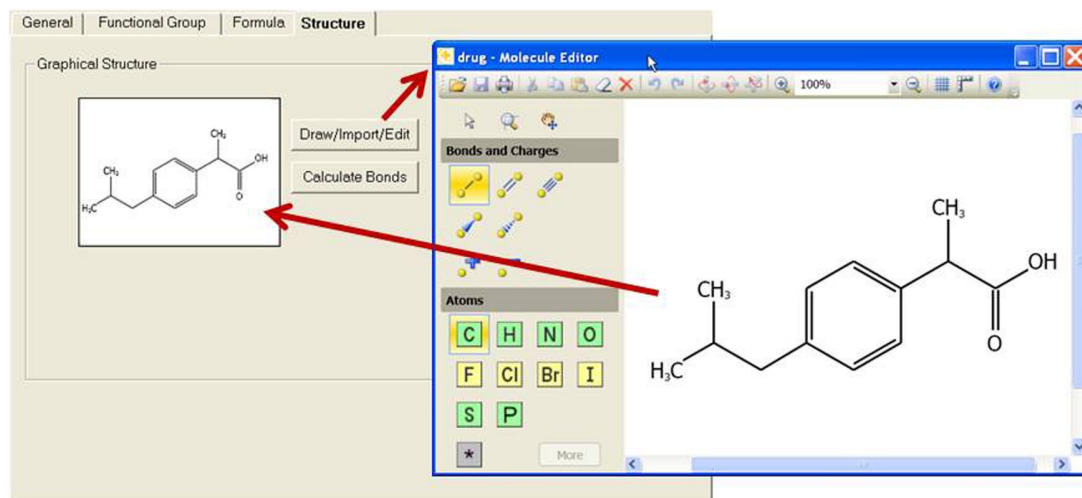


Fig. 10 Molecular structure can be drawn and provided to TDE for use in pure-component property evaluation. Atomic connectivity information is used by TDE to automatically generate functional groups required by the appropriate estimation methods.

Evaluate accuracy of mixture model

After evaluation of pure components has been completed, the user can start the evaluation of mixture properties, which is a more complex task. As discussed earlier and shown in Table 2, the simulator contains a large number of built-in binary interaction parameters. It is important to ascertain whether or not the binary interaction parameters are available for the components of interest. And if they are available, whether or not they are appropriate for the process conditions (temperature, pressure, composition) and types of phase equilibrium (VLE, LLE, SLE) under development. If the binary interaction parameters are not available, then they must be determined, subject to the degree of importance of the binary pair to the modeling problem at hand (e.g., whether they constitute key components, trace components that affect product specification, or trace components with no important consequence).

To evaluate the accuracy of built-in binary interaction parameters or to develop new ones, experimental mixture data are required. For a component pair for which no experimental data are available, methods such as UNIFAC [55] can be used to fill the gaps. However, it must be emphasized that for final design and for key components, binary interaction parameters based on experimental data, even limited ones, are essential.

The TDE dialog box shown in Fig. 5 can be used to select the two components of interest to search for any available experimental mixture data. As an example, the binary mixture water and NMP (*N*-methyl-2-pyrrolidone) will be used. All available experimental data for the selected binary mixture from SOURCE are located and displayed on the results form as shown in Fig. 11. The data are categorized and displayed in a tree view on the left-hand side for convenience and easy access. For this system, binary diffusion coefficient, VLE, density, excess enthalpy, heat capacity at constant pressure, surface tension, and viscosity data are available. On the right-hand side, an overview of each data set is shown, including the number of data points, year of publication, temperature range, and pressure range. The temperature and pressure ranges are particularly useful as they can be used to identify the data sets most appropriate for the process conditions of interest. Fitting data that are well beyond the temperature and pressure ranges of the process can lead to a sub-optimum model, resulting in unnecessary sacrifice of accuracy. However, these “extra” data can be used to test the extrapolation capability of the

	Name	Points	Year	Temp. ranges (K)	Press. ranges (N/sqm)
1	Binary diffusion coefficient 001	5	1995	298 – 348	101325
2	Binary diffusion coefficient 002	5	1995	298 – 348	101000
3	Binary diffusion coefficient 003	12	1985	278	101000
4	Binary diffusion coefficient 004	12	1995	283 – 313	101325
5	Binary VLE 001	13	1967	339 – 427	26660
6	Binary VLE 002	13	1967	373 – 475	101320
7	Binary VLE 007	9	1975	357 – 414	53320
8	Binary VLE 008	11	1975	374 – 433	101320
9	Binary VLE 009	8	1975	339 – 386	26660
10	Binary VLE 011	24	1986	374 – 471	101000
11	Binary VLE 013	12	1992	373 – 475	101300
12	Binary VLE 015	24	1994	374 – 471	101325
13	Binary VLE 017	46	1996	351	1210 – 43370
14	Binary VLE 018	40	1996	380	4640 – 130030
15	Binary VLE 004	13	1972		101000
16	Binary VLE 010	11	1980	273 – 313	20 – 4702
17	Density 001	30	2008	298	101000
18	Density 002	16	1985	278	101000
19	Density 003	65	2004	298 – 343	101150

Fig. 11 The main TDE results form for binary mixture data. Data shown are for NMP–water. All available experimental data are shown together with the number of data points, year of publication, and temperature and pressure ranges. These information can help quickly identify data sets that should be further evaluated for suitability in the application of interest.

model, which is always a good practice. In many instances, overfitting the data (i.e., using more parameters than is necessary) can result in poor extrapolation behavior.

Figure 12 displays isobaric data for NMP–water at 0.5332 bar from Golubkov et al. [56]. The data can also be shown graphically by clicking on the *Plot* button. See Fig. 13 for T - x - y plot of the selected NMP–water data, labeled binary VLE 007 in Fig. 12.

As discussed in the “Quality assessment of binary VLE data” section, it is useful to perform thermodynamic consistency tests for the VLE data. The overall quality assessment results can help the user make informed decision regarding the data sets to use as well as the weighting factors to assign to them when performing data regression, especially when there are many data sets available. The thermodynamic consistency test is initiated by clicking on the *Consistency* button (see Fig. 12). Consistency test results are displayed in the summary view (see Fig. 14) with the overall data quality (Q_{VLE}) and results for each test. For this system, the overall data quality for the available data sets is relatively poor. Most of the isobaric data do not cover the entire composition range, similar to the data set shown in Fig. 13. Nevertheless, the quality of one data set relative to another can still be inferred using Q_{VLE} .

Given experimental data (e.g., VLE, LLE, excess enthalpy), it is possible to verify the accuracy of the mixture model—provided that built-in binary interaction parameters exist. For the NMP–water system, their NRTL binary interaction parameters are available in the simulator as shown in Table 6. Only VLE data were used to determine these binary interaction parameters. The data covered a wide temperature range (293–451.49 K) and pressures up to atmospheric. As an illustrative example, the two sets of VLE data with the highest overall data quality (sets labeled binary VLE 007 and binary VLE 008, in Fig. 12) will be used to verify the accuracy of the built-in binary interaction parameters. Binary VLE 007 is at 0.5332 bar, while binary VLE 008 is at 1.0132 bar with a combined temperature

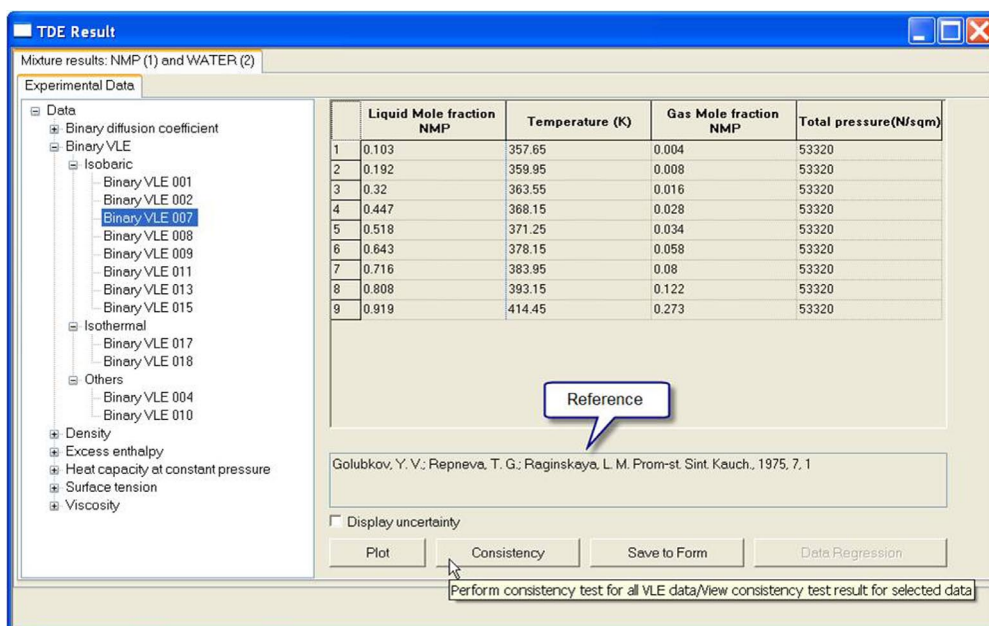


Fig. 12 Experimental VLE data of NMP–water at 0.5332 bar from Golubkov et. al. [56]. The data can be plotted using the *Plot* button. VLE data can also be evaluated for thermodynamic consistency using the *Consistency* button. Data sets of interest can be saved to the simulation case file using the *Save to Form* button.

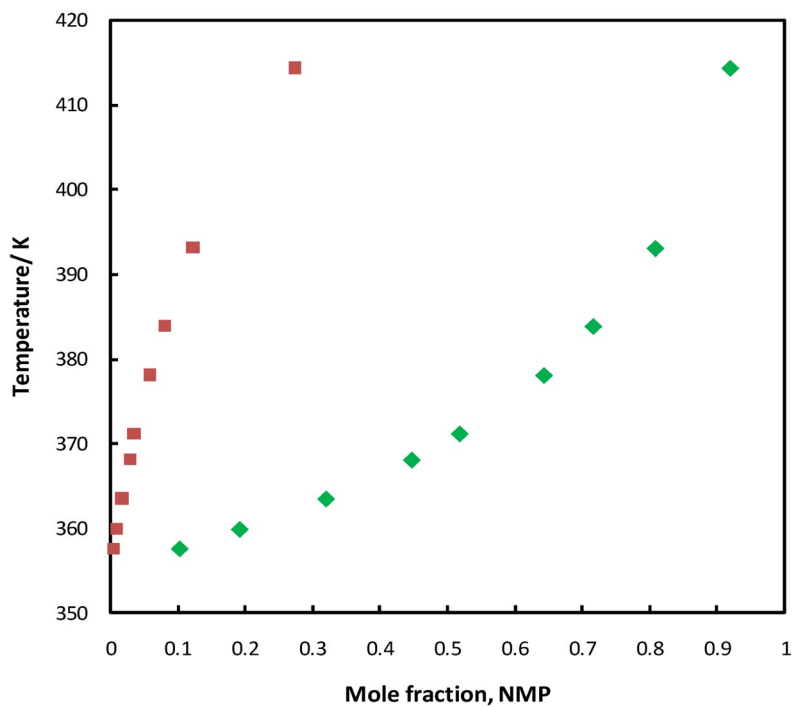


Fig. 13 T - x - y diagram of NMP–water data at 0.5332 bar [56].

	Name	Points	Year	Temp. ranges (K)	Press. ranges (N/sqm)	Overall data quality	Herington test	Van Ness test	Point test	Inf diluti
1	Binary VLE 001	13	1967	339 -- 427	26660	0.24129	Passed	Failed	Not Available	Failed
2	Binary VLE 002	13	1967	373 -- 475	101320	0.17848	Failed	Passed	Not Available	Failed
3	Binary VLE 007	9	1975	357 -- 414	53320	0.5007	Failed	Passed	Not Available	Failed
4	Binary VLE 008	11	1975	374 -- 433	101320	0.33437	Failed	Failed	Not Available	Failed
5	Binary VLE 009	8	1975	339 -- 386	26660	0.11204	Failed	Failed	Not Available	Failed
6	Binary VLE 011	24	1986	374 -- 471	101000	0.10118	Failed	Failed	Not Available	Failed
7	Binary VLE 013	12	1992	373 -- 475	101300	0.20179	Failed	Passed	Not Available	Failed
8	Binary VLE 015	24	1994	374 -- 471	101325	0.1018	Failed	Failed	Not Available	Failed
9	Binary VLE 017	46	1996	351	1210 -- 43370	0.28823	Not Performed	Not Performed	Not Performed	Not Pe
10	Binary VLE 018	40	1996	380	4640 -- 130030	0.30171	Not Performed	Not Performed	Not Performed	Not Pe
11	Binary VLE 004	13	1972		101000	0.012812	Not Performed	Not Performed	Not Performed	Not Pe
12	Binary VLE 010	11	1980	273 -- 313	20 -- 4702	0.026803	Not Performed	Not Performed	Not Performed	Not Pe

Fig. 14 Summary view of VLE data for NMP–water system, showing overall data quality and results of the four thermodynamic consistency tests.

range of 357–433 K. In addition, excess enthalpy data of Kuramaki et al. [57] and Noll et al. [58] will also be used. These data cover a temperature range of 298.15–416.15 K. Both types of data are well within the valid temperature and pressure ranges of the built-in binary interaction parameters.

Table 6 Binary interaction parameters for the NRTL model [16]* for the water (1)–NMP (2) mixture.

	A_{12}	A_{21}	B_{12}	B_{21}	α_{12}	E_{12}	E_{21}
Original	7.4081	-2.2438	-2279.96	533.85	0.3	0.0	0.0
Regressed	11.6476	20.1942	-1234.28	-1154.55	0.3	-1.204	-3.005

* $\tau_{12} = A_{12} + B_{12}/T + E_{12} \ln(T)$ [1]. $\tau_{21} = A_{21} + B_{21}/T + E_{21} \ln(T)$. α_{12} is the non-randomness factor.

The NRTL model with built-in binary parameters represents the VLE data quite well, as shown in Fig 15. Results for excess enthalpy are shown in Fig. 16. Significant deviations exist, especially at 298.15 K for water mole fraction less than 0.6. Deviation on a percentage basis is very large (up to 100 %) for both temperatures. This is not too surprising since excess enthalpy data were not used in the original data regression work. However, considering the very large deviation, it is useful to explore the possibility of improving the representation of excess enthalpy while maintaining the accuracy of VLE prediction.

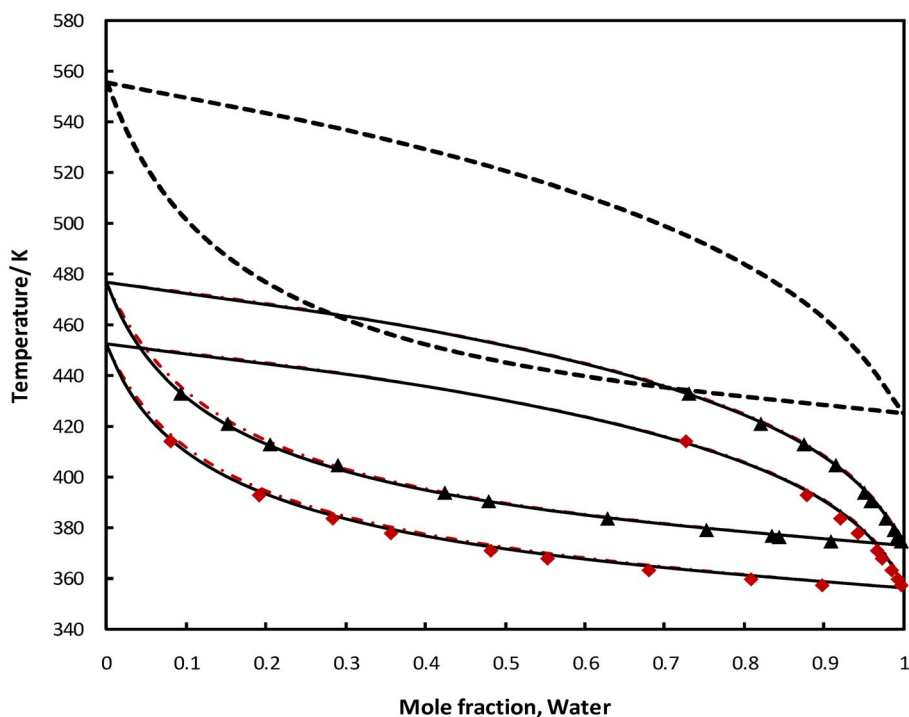


Fig. 15 Comparison between NRTL model prediction using binary interaction parameters from Table 6 and experimental VLE data of NMP–water binary mixture. Data from Golubkov et. al. [56]; 0.5332 bar (◆); 1.0132 bar (■). NRTL results are shown as lines. - - -: built-in binary interaction parameters; —: regressed binary interaction parameters determined in this work; - · -: extrapolation to 5.0 bar using the regressed parameters.

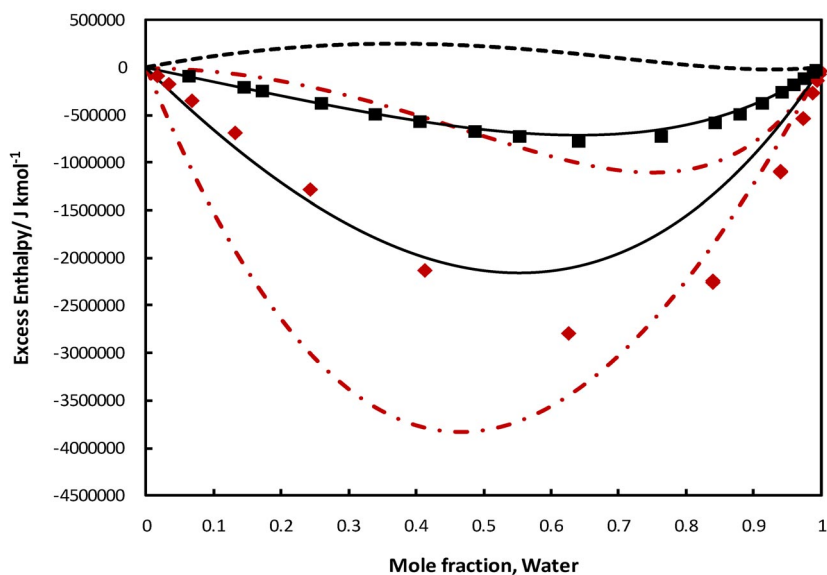


Fig. 16 Comparison between NRTL model prediction using binary interaction parameters from Table 6 and experimental excess enthalpy data of NMP–water binary mixture. Data: 298.15 K (◆) [57]; 416.15 K (■) [58]. NRTL results are shown as lines. - - -: built-in binary interaction parameters; —: regressed parameters determined in this work; - · -: extrapolation to 500 K using the regressed parameters.

Fine-tune/develop mixture model

The NRTL model can be further improved by performing data regression—in this example, using both VLE and excess enthalpy data. The same four sets of data used in the previous step are used here. Significantly higher weight was given to the two sets of VLE data. This corresponds to the higher importance of phase equilibrium data.

Data regression can be initiated using the *Data Regression* button shown in Fig. 12. The new binary interaction parameters are reported in Table 6. Results of the data regression for VLE and excess enthalpy are shown in Figs. 15 and 16, respectively. As can be seen (solid lines in the figures), VLE and excess enthalpy data are adequately correlated. The excess enthalpy representation was significantly improved while maintaining the accuracy of VLE.

To further ascertain that the revised binary interaction parameters are reasonable, they are used to predict VLE and excess enthalpy at other conditions. VLE calculations are made at 5 bar and results are added to Fig. 15. For excess enthalpy, the extrapolation is performed at 500 K. Results are shown in Fig. 16. These results do not show any unusual behavior.

Since transport properties also have significant impact on process design and the associated capital costs as discussed earlier, it is important to develop and improve transport property models. Unfortunately, this is an area where insufficient attention had been paid thus far. As an illustrative example, liquid viscosity of the NMP–water system is studied. As shown in Fig. 11, liquid viscosity data are available for this binary mixture. The experimental data of Henni et al. [59] are compared against the model developed by Song and co-workers [60]. This model uses two binary interaction parameters (one symmetric and another antisymmetric) to capture the nonlinear mixing behavior of liquid viscosity. The binary interaction parameters are also temperature-dependent, which provides better fit of experimental data over a wide temperature range. First, calculations are made using the model at 298, 323, and 343 K without the use of binary interaction parameters. Results are very poor when compared with experimental data as shown in Fig. 17. Temperature-dependent binary interaction parameters are then

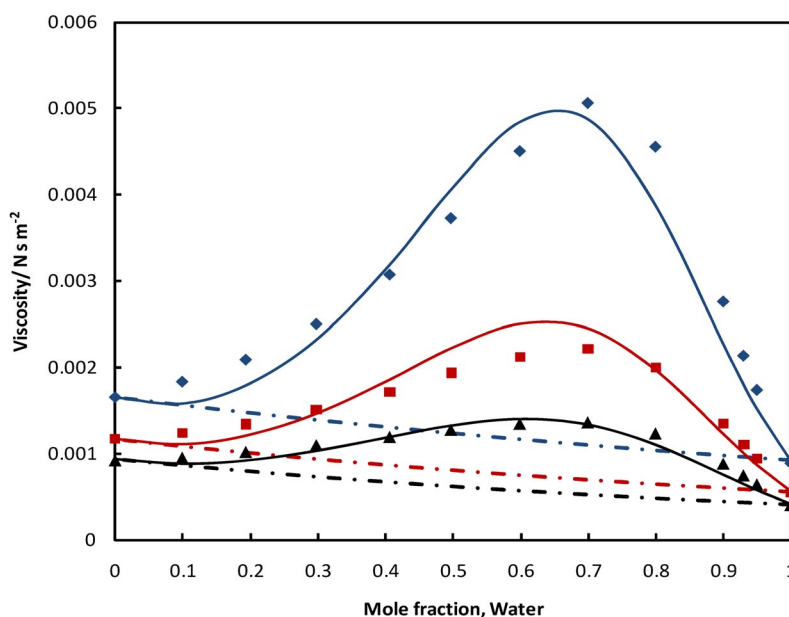


Fig. 17 Comparison between prediction using the Song model [60] and experimental liquid viscosity data for the NMP–water binary mixture. Data of Henni et. al. [59] at 298 K (◆), 323 K (■), and 343 K (▲) are used. Model results are shown as lines. - - -: zero binary interaction parameters; —: regressed parameters shown in Table 7.

obtained by regression using the available experimental data; results are reported in Table 7. The predicted liquid viscosity values are also compared against experimental data in Fig. 17. As shown, the model can adequately correlate the very nonlinear, unsymmetric liquid viscosity data of this system.

Table 7 Binary interaction parameters for the liquid viscosity model [60]* for the water (1)–NMP (2) mixture.

a_{12}	b_{12}	c_{12}	d_{12}
-51.329	67.58	65.479	-85.08

*Symmetric binary interaction parameter $k_{12} = a_{12} + b_{12}/T_r$, $k_{12} = k_{21}$. Antisymmetric binary interaction parameters $l_{12} = c_{12} + d_{12}/T_r$, $l_{12} = -l_{21}$. $T_r = T/T_{ref}$; $T_{ref} = 298.15$ K.

Publish model

Finally, the critically evaluated property data and parameters can be used with confidence in the simulation. After further validation in the field, the evaluated data can also be permanently stored in the database (see Fig. 3). In some organizations, property data developed for a project or process are packaged and deployed using a secured database [1] when working with a 3rd party vendor or joint venture. The secured database possesses unique characteristics in that the property data that are stored in the database can be used in the process calculations but cannot be viewed or altered. This helps facilitate exchange of process simulation files and enforce physical property standards while protecting the companies' intellectual property. As a result of increasing globalization, the need to cooperate with 3rd party vendors and joint ventures in other countries also increases. At the same time, companies recognize the need to be more vigilant about their intellectual property, such as proprietary thermophysical property data and parameters.

LIMITATIONS AND FUTURE PLANS

Dynamic data evaluation is currently applicable to non-reacting chemical systems. It cannot be used for electrolyte and polymer systems. Experimental data available in the SOURCE database are primarily for organic compounds. The estimation methods available for filling the gaps of experimental data of pure components are also for those containing C, H, N, O, S, F, Cl, Br, I, and P atoms only.

Although TDE makes it possible to develop critically evaluated properties quickly and confidently, it is ultimately the responsibility of the engineer to ascertain that the resultant properties meet the needs of the process model under development. TDE is an enabling technology, not a panacea, and should be used in conjunction with scientific and engineering judgment.

The software technology related to the evaluation of pure-component properties is relatively well developed. Much work remains on the evaluation of binary mixture properties as significant user interactions and expertise are still required. Future work will continue to be carried out in cooperation with our colleagues at TRC. The plan will be focused on improving data analysis, quality assessment, and regression of nonlinear models using multi-property data regression analysis with the aim to increase the level of automation for critical evaluation of binary mixture data. The quality assessment algorithm will be enhanced to handle binary VLE mixtures containing supercritical components, binary LLE, and other types of data. In addition, the extensive pure-component experimental data can be used to further refine and extend the group-contribution methods to improve the accuracy as well as the type of atoms and functional groups that can be handled. Results of this research can also be used to further fine-tune the heuristics used in the selection of the best method for estimating a given property for a given class

of compound. The ability to easily compare properties of compounds within a homologous series will also be a very useful addition. TDE is expected to be extended to cover ternary data.

Finally, for completeness, we would like to note that other valuable tools exist that provide structure-based estimation, experimental data access and analysis, and process synthesis tools, for example, DDB Software Package DDBSP [24] and DECHEMA DPP [61]. However, they are not available within the domain of the process simulator and therefore lack close interoperability and, to some extent, compatibility. More importantly, these software are separately licensed at significant cost relative to the cost of the simulator, thus they are not accessible to the majority of the users unlike the technology described in this work, which is provided by default as an integral part of the simulator.

CONCLUSIONS

In this paper, we described the important role thermophysical properties play in the development of process simulation models, reviewed past practice in obtaining the necessary property data, and introduced a new technology that can be used to overcome the shortcomings of the current method. The new technology is software that implements the dynamic data evaluation concept, which makes use of experimental thermophysical property data, structure-based property estimation system, data analysis, and regression algorithms to generate critically evaluated property data and parameters. The implementation of the new technology in a process simulator, Aspen Plus, has been described and illustrated via examples.

Direct access to the “raw” experimental data is a fundamental shift in the process simulation market. The enabling technology, NIST ThermoData Engine, allows the required thermophysical properties to be obtained “on-demand” with confidence. Engineers can more easily develop accurate mixture models based on experimental data that cover conditions relevant to the process of interest—with full traceability of data and parameters to the original source. The technology helps accelerate process model development and ensures higher-quality models. The resulting models also help companies optimize operating conditions and run their processes with greater agility, helping them stay profitable in a volatile market.

ACKNOWLEDGMENTS

The author wishes to acknowledge the technical cooperation and supports of the members of the TRC of NIST. Special thanks go to Dr. Michael Frenkel, director of TRC whose leadership, vision, and steadfast support make this work possible. The author also wishes to thank his colleagues at Aspen Technology: Dr. Maohua Cao and Dr. Shu Wang for developing the software integration between Aspen Plus and TDE; Dr. Vikas Dhole, Dr. Chau-Chyun Chen, Mr. David Tremblay, and Mr. Willie Chan whose management support and encouragement are critical to the successful introduction of the technology into our software.

REFERENCES

1. Aspen Plus[®] Versions 2006, 2006.5, V7, V7.1, and V7.2. Aspen Technology, Inc. Burlington, MA, USA.
2. B. E. Poling, J. M. Prausnitz, J. P. O’Connell. *The Properties of Gases and Liquids*, 5th ed., McGraw-Hill, New York (2001).
3. H. I. Britt, R. Luecke. *Technometrics* **15**, 233 (1973).
4. M. Frenkel, R. D. Chirico, V. Diky, X. Yan, Q. Dong, C. Muzny. *J. Chem. Inf. Model* **45**, 816 (2005).
5. J. R. Fair. *Advanced Process Engineering*, AIChE Monogram Series **13**, 76 (1980).

6. R. Dohrn, O. Pfohl. Presented at the 17th European Symposium on Applied Thermodynamics, Vilamoura, Portugal (1999).
7. D. Y. Peng, D. B. Robinson. *Ind. Eng. Chem. Fundam.* **15**, 59 (1976).
8. (a) C. Antoine. *C. R. Hebd. Séances Acad. Sci.* **107**, 681 (1888); (b) C. Antoine. *C. R. Hebd. Séances Acad. Sci.* **107**, 778 (1888); (c) C. Antoine. *C. R. Hebd. Séances Acad. Sci.* **107**, 836 (1888).
9. W. Wagner. *Cryogenics* **13**, 470 (1973).
10. Aspen HYSYS[®] Version V7.1. Aspen Technology, Inc. Burlington, MA, USA.
11. H. G. Rackett. *J. Chem. Eng. Data* **15**, 514 (1970).
12. R. E. Thek, L. I. Stiel. *AIChE J.* **13**, 626 (1967).
13. F. A. Aly, L. L. Lee. *Fluid Phase Equilib.* **6**, 169 (1981).
14. E. N. da C. Andrade. *Nature* **125**, 309 (1930).
15. T. Holderbaum, J. Gmehling. *Fluid Phase Equilib.* **70**, 251 (1991).
16. H. Renon, J. M. Prausnitz. *AIChE J.* **14**, 135 (1968).
17. D. S. Abrams, J. M. Prausnitz. *AIChE J.* **21**, 116 (1975).
18. P. Chang, C. R. Wilke. *J. Phys. Chem.* **59**, 592 (1955).
19. W. V. Wilding, R. L. Rowley, J. L. Oscarson. *Fluid Phase Equilib.* **150–151**, 413 (1998).
20. I. Barin. *Thermochemical Data of Pure Substances*, VCH, Weinheim (1989).
21. D. D. Wagman, W. H. Evans, V. B. Parker, R. H. Schumm, I. Halow, S. M. Bailey, K. L. Churney, R. L. Nuttall. *J. Phys. Chem. Ref. Data* **11**, Suppl. 2 (1982).
22. J. Gmehling, U. Onken, W. Arlt, P. Grenzheuser, B. Kolbe, U. Weidlich, J. Rarey. *Vapor-Liquid Equilibrium Data Collection*, 31 parts, DECHEMA Chemistry Data Series, Frankfurt (1977–1978).
23. H. Knapp, R. Doring, L. Oellrich, U. Plocker, J. M. Prausnitz, R. Langhorst, S. Zeck. *Vol. VI: Vapor-Liquid Equilibrium for Mixtures of Low Boiling Substances*, DECHEMA Chemistry Data Series, Frankfurt (1982).
24. J. Rarey, J. Gmehling. *Pure Appl. Chem.* **81**, 1745 (2009).
25. S. Watanasiri, M. R. Brule, K. E. Starling. *AIChE J.* **28**, 626 (1982).
26. J. G. Hayden, J. P. O'Connell. *Ind. Eng. Chem. Process Des. Dev.* **14**, 209 (1975).
27. G. M. Wilson. *J. Am. Chem. Soc.* **86**, 127 (1964).
28. O. Redlich, J. N. S. Kwong. *Chem. Rev.* **44**, 233 (1949).
29. V. Diky, R. D. Chirico, R. C. Wilhoit, Q. Dong, M. Frenkel. *J. Chem. Inf. Comput. Sci.* **43**, 15 (2003).
30. V. Diky, C. Muzny, E. Lemmon, R. D. Chirico, M. Frenkel. *J. Chem. Inf. Model.* **47**, 1713 (2007).
31. V. Diky, R. D. Chirico, A. F. Kazakov, C. Muzny, M. Frenkel. *J. Chem. Inf. Model.* **49**, 503 (2009).
32. L. Constantinou, R. Gani. *AIChE J.* **40**, 1697 (1994).
33. J. Marrero-Morejon, E. Pardillo-Fontdevila. *AIChE J.* **45**, 615 (1999).
34. Y. Nannoolal, J. Rarey, D. Ramjugernath, W. Cordes. *Fluid Phase Equilib.* **226**, 45 (2004).
35. K. G. Joback, R. C. Reid. *Chem. Eng. Commun.* **57**, 233 (1987).
36. G. M. Wilson, L. V. Jasperson. AIChE Meeting, New Orleans, LA (1996).
37. A. Kazakov, C. D. Muzny, V. Diky, R. D. Chirico, M. Frenkel. *Fluid Phase Equilib.* **298**, 131 (2010).
38. (a) L. Riedel. *Chem.-Ing.-Tech.* **26**, 259 (1954); as modified in (b) J. L. Hales, R. Townsend. *J. Chem. Thermodyn.* **4**, 763 (1972).
39. B. E. Poling, J. M. Prausnitz, J. P. O'Connell. *The Properties of Gases and Liquids*, 5th ed., p. 6.22, McGraw-Hill, New York (2001).
40. D. Ambrose, J. Walton. *Pure Appl. Chem.* **61**, 1395 (1989).

41. B. E. Poling, J. M. Prausnitz, J. P. O'Connell. *The Properties of Gases and Liquids*, 5th ed., p. 9.9, McGraw-Hill, New York (2001).
42. S. R. S. Sastri, K. K. Rao. *Chem. Eng. J.* **50**, 9 (1992).
43. T. H. Chung, L. L. Lee, K. E. Starling. *Ind. Eng. Chem. Fundam.* **23**, 8 (1984).
44. T. H. Chung, M. Ajlan, L. L. Lee, K. E. Starling. *Ind. Eng. Chem. Res.* **27**, 671 (1988).
45. B. E. Poling, J. M. Prausnitz, J. P. O'Connell. *The Properties of Gases and Liquids*, 5th ed., Section 3-5, McGraw-Hill, New York (2001).
46. J. W. Kang, V. Diky, R. D. Chirico, J. W. Magee, C. D. Muzny, I. Abdulagatov, A. F. Kazakov, M. Frenkel. *J. Chem. Eng. Data* **55**, 3631 (2010).
47. E. F. G. Herington. *J. Inst. Pet.* **37**, 457 (1951).
48. J. Wisniak. *Ind. Eng. Chem. Res.* **32**, 1531 (1993).
49. K. Kojima, H. M. Moon, K. Ochi. *Fluid Phase Equilib.* **56**, 269 (1990).
50. K. Kurihara, Y. Egawa, K. Ochi, K. Kojima. *Fluid Phase Equilib.* **219**, 75 (2004).
51. H. C. Van Ness, S. M. Byer, R. E. Gibbs. *AIChE J.* **19**, 238 (1973).
52. P. L. Jackson, R. A. Wilsak. *Fluid Phase Equilib.* **103**, 155 (1995).
53. K.-H. Nothnagel, D. S. Abrams, J. M. Prausnitz. *Ind. Eng. Chem. Process Des. Dev.* **12**, 25 (1973).
54. C. C. Chen, H. I. Britt, J. F. Boston, L. B. Evans. *AIChE J.* **28**, 588 (1982).
55. A. Fredenslund, R. L. Jones, J. M. Prausnitz. *AIChE J.* **21**, 1086 (1975).
56. Y. V. Golubkov, T. G. Repneva, L. M. Raginskaya. *Prom-st. Sint. Kauch.* **7**, 1 (1975).
57. S. Murakami, R. Tanaka, R. J. Fujishiro. *Solution Chem.* **3**, 71 (1974).
58. O. Noll, K. Fischer, J. Gmehling. *J. Chem. Eng. Data* **41**, 1434 (1996).
59. A. Henni, J. J. Hromek, P. Tontiwachwuthikul. *J. Chem. Eng. Data* **49**, 231 (2004).
60. Y. Song, P. M. Mathias, D. Tremblay, C. C. Chen. *Ind. Eng. Chem. Res.* **42**, 2415 (2003).
61. U. Westhaus, R. Sass. *Fluid Phase Equilib.* **49**, 222 (2004).