Benefits of the NIST ThermoData Engine in aspenONE Engineering V7.3

Thermophysical Properties On Demand

David Tremblay – Product Management Director, Aspen Plus Family
Suphat Watanasiri – Senior Director of Technology, Physical Properties

4/25/2011

This document describes the technical and business benefits of the NIST ThermoData Engine (TDE) and NIST Source Database in aspenONE V7.3.
Synopsis

The NIST ThermoData Engine in Aspen Plus® V7.3 provides on-demand thermophysical property data evaluation, helping engineers to rapidly develop accurate, high-fidelity chemical process models.

The ThermoData Engine feature in Aspen Plus V7.3 has been significantly extended to include access to over three million points of experimental phase equilibrium and mixture property data.

On July 20, 2009, Aspen Technology Inc. and the US National Institute of Standards and Technology (NIST) jointly received an R&D 100 award for their collaborative development culminating in aspenONE V2006.5. Winning an R&D 100 Award demonstrates a mark of excellence known to industry, government and academia that the product is one of the most innovative ideas of the year. Since 1963, the R&D 100 Awards have honored such revolutionary technologies as the ATM (1973), the printer (1986), HDTV (1998), and others.

What is the Significance of the NIST Technology to our Customers?

The twenty-first century is proving difficult for companies in the Continuous Processing Industries (CPI) 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 high-value specialty chemicals and intermediates. Pharmaceutical companies are developing synthesis routes for a wide slate of products while reducing production costs for their blockbuster drugs. New processes are being developed to manufacture biofuel, and to synthesize liquid fuel from natural gas and synthetic gas from coal, coke, oil sands, and biomass. Oil companies are working to improve yields through better characterization and 'molecule management' in their facilities. High-fidelity process models are required to design, operate, and optimize these processes to maximize yields, minimize energy and capital costs, and to ensure environmental compliance. Aspen Plus® with NIST ThermoData Engine (TDE) is a general purpose process simulator that provides new tools to significantly increase the number and variety of chemical components that may be considered by the process-design engineer, while dramatically reducing the time required for compilation/critical-evaluation of the component properties, which underlie high-fidelity process-model development.

This document contains some forward looking statements; it does not imply any contractual obligations.
Process simulation models use 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 and to size and rate equipment require reliable chemical component thermophysical property data. For example, phase equilibrium, which controls the separation of species in distillation columns, depends on the component vapor pressures, while heat transfer depends on their thermal conductivity, heat capacity, heat of formation, and heat of vaporization. The accuracy (i.e., fidelity) of the process simulation models is directly related to the accuracy of the underlying thermophysical property information. Poor data or poorly estimated property parameters can lead to very large errors in the calculated capital and operating costs of the plant, which in turn can lead to poor investment decisions. Worse, a poorly based model could lead to infeasible process designs, poor plant operability, or unsafe plants. Thus, it is absolutely essential that process simulation models are based on precise property data.

Chemical industry trends towards new chemicals and new chemical families present an enormous challenge to providing the necessary property information in a timely and cost effective manner.

When developing process models, one of the most difficult challenges has been to collect, evaluate, regress, and estimate thermophysical property parameters for the various chemical species involved. With traditional methods, property evaluators spend weeks or even a year or more collecting data from thermodynamic journals, reference books, in-house data systems, and open and/or private on-line databases. Some property data are not available and must be estimated from molecular structure or limited experimental data. Traditionally, experienced experts in thermodynamics have performed this task using data regression and estimation tools available in the simulator or with third-party data fitting and spreadsheet tools. Highly qualified and experienced professionals who know the tricks of the trade have successfully applied their knowledge of thermodynamics to determine the various property parameters for the key species in the process. This approach is extremely slow and costly, and often fails completely because the evaluator cannot keep pace with industry requirements for new chemicals and properties.

Less experienced engineers may use the tools with highly mixed results, often yielding a set of thermophysical property parameters which appear to match the data well on the surface only. Unfortunately, if these parameters are not validated properly, they may not be thermodynamically consistent, and can lead to infeasible results. For example, the resulting model might predict a
product separation where none exists, resulting in a complete failure of the process design. Furthermore, the difficulty in obtaining missing property parameters leads to the use of relatively crude and correspondingly less accurate models. Experienced property evaluators continue to become rarer, particularly as the baby-boomer generation has reached retirement age and a new wave of less experienced engineers replacing them in the workforce.

Over the last 30 years, process simulation software developed by different companies has been equipped with databases containing thermophysical property data for a very limited number of chemical species (typically 1500 to 2000 pure components). Obviously, the scope of these databases limits severely the range of components which can be considered in any given process model. Given the increasing importance of new materials and processes, one might wonder why simulation companies have not developed larger databases more quickly. Efforts to expand the scope of these databases have been hampered by the cost and effort required to collect data from many published sources, to verify the integrity of these data, and to fill any gaps in the data using estimation methods. These traditional 'static' workflows for gathering and checking data are highly labor intensive, error prone, and inefficient, with very long turnaround times of many weeks to even years.

As an alternative to the static approach, the new concept of dynamic data evaluation was developed at the National Institute of Standards and Technology (NIST). Implementation of this concept has required the successful development of a large electronic database storing essentially all experimental thermodynamic data known to date with detailed descriptions of relevant metadata and uncertainties; the NIST SOURCE Data Archive. The assessment and storage of the uncertainties are essential to the success of the dynamic approach. (The inclusion of uncertainties for the experimental data is key and is absent in all competing products.) The combination of this electronic database with expert-system software, designed to automatically produce recommended property models based on available experimental data and a set of prediction methods, enables on-demand generation of critically evaluated data. Property evaluation times are measured in seconds or minutes rather than weeks, months, or years. This approach also contrasts sharply with the static approach, which must be initiated far in advance of need. The dynamic data evaluation process dramatically reduces the effort and costs associated with developing, validating, and maintaining a thermophysical property database.

This document contains some forward looking statements; it does not imply any contractual obligations.
NIST Thermodynamics Research Center (TRC) maintains a Data Entry Facility in Boulder Colorado. 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 developed by NIST (Guided Data Capture software). 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 the NIST SOURCE Data Archive that underlies the dynamic data evaluation approach.

The dynamic data evaluation concept was fully implemented in the software program NIST ThermoData Engine or TDE. TDE makes use of the experimental and metadata from the SOURCE Data Archive together with expert system software – structure-based property estimation, a rule-based engine to enforce thermodynamic consistency, and data regression – to produce critically evaluated data and property parameters. Aspen Plus® implements the NIST TDE technology, thus putting the process designer at the focal point of property information processing, rather than subjecting him or her to the slow pace and whims of external data evaluators. Property evaluation is done on demand in real time for exactly the components that are needed.

A snapshot of the NIST SOURCE Data Archive is included with Aspen Plus®. This addition has greatly increased the scope of pure component datasets available in the software – from those for approximately 9,000 species to more than 23,000 unique species, which is more than an order of magnitude greater than that for all competitors. Further, for the first time ever, process engineers have direct access to the experimental data itself; whereas all other data collections provided with process simulation software are limited to fitted parameters derived from experimental data. The current version of the SOURCE Data Archive includes over one million experimental data points for thermophysical properties of pure compounds (nearly 4 million for all types of chemical systems). With NIST TRC processing approximately 500,000 data points per year, the scope of this database is expected to grow rapidly with each new version of the software.

AspenTech provides quarterly updates to the NIST source database through the customer support knowledgebase. This puts the newest available data into the hands of Aspen Plus users on a regular basis.
Direct access to the “raw” experimental data is a fundamental shift in the process simulation market. In this new product, a thermodynamic expert can extract the experimental data electronically and use it to optimize property parameters over particular temperatures and pressures of interest. They can also use the raw data to fit existing property models or their own custom property models. These features enable the development of high-quality customized databases for particular processes, which in turn lead to more accurate process designs, better cost estimates, and optimization of process operating conditions. The ready access of this data can save weeks of data gathering efforts. Further, the experimental data are pre-assessed, validated, and free of transcription errors. The sources for all data (literature citations) are readily available, providing, for the first time, unequivocal traceability for all property evaluations.

The NIST ThermoData Engine in Aspen Plus provides easy access to millions of sets of pure component physical property data on demand. The uncertainty and source of every data point is fully documented. The data can be plotted against model predictions with a single mouse click.

Experimental data are easily displayed visually in charts, and can be quickly compared against model predictions. This allows experts and novices to verify the agreement between their models and the available experimental data in minutes.
Benefits of the NIST ThermoData Engine in aspenONE Engineering

Engineers can quickly and visually validate model predictions against experimental data. The circles and diamonds show experimental data and estimated data used by NIST TDE to fit the property model parameters. The square points are data points rejected by the expert system in NIST TDE. The curve shows the property values predicted by the model.

The **NIST SOURCE Data Archive** within **Aspen Plus®** is highly complementary to the legacy databases available in the software. Incorporation of **NIST ThermoData Engine (TDE)** is accomplished in a variety of ways, including the IUPAC standard for thermophysical property data communications, *ThermoML*, which was developed at NIST and dynamic link library (DLL) technology. The use of this advanced technology has made the product integration seamless to the user, and it puts an enormous amount of data into the hands of engineers and scientists. The NIST archive is especially rich in data for intermediate-weight organic compounds. This is especially important for modeling processes to produce specialty chemicals and chemical intermediates. There are also considerable amounts of data for compounds associated with biodiesel, coal liquefaction and gasification, and a wide range of common heat transfer fluids.

The **NIST SOURCE Data Archive** includes molecular structure information in addition to experimental data. The structural data is used together with available experimental data to evaluate missing component properties on demand through the **NIST ThermoData Engine (TDE)** expert system. **TDE** is the first expert system which brings together automated experimental data regression, structure-based property estimation, and a rule-based engine which applies thermodynamic methods and relationships to produce critically evaluated property parameters.

This document contains some forward looking statements; it does not imply any contractual obligations.
With NIST TDE, novices and experts alike can rapidly obtain the key property parameters for a wide range of compounds. This expert system leverages the experimental data and uncertainties included in the NIST SOURCE Data Archive, as well as proprietary experimental data which might be entered by the user. Further, this system can use molecular structure information imported from a MOL file or drawn using a molecular structure drawing tool included in Aspen Plus® V7.2. This means that the NIST TDE expert system can be used to estimate the properties for any organic compound (including those which contain sulfur, fluorine, chlorine, iodine, and bromine), thus providing direct access to property values for literally millions of compounds, all with reliable and conservative estimates of uncertainty.

The NIST ThermoData Engine can estimate the properties of new components drawn using a built-in tool or imported as chemistry MOL files from other software packages. The expert system will use the structural information and any additional user-specified data to estimate a wide range of physical property parameters. The system uses thermodynamic relationships to ensure the parameter set is self-consistent and reliable.

Enforcement of thermodynamic consistency is built into TDE, thus ensuring that resulting properties and parameters are self-consistent. For example, closely related properties, such as vapor pressure, boiling point, critical properties, vapor density, and heat of vaporization, exhibit the required thermodynamic relationships. The uncertainties for some property evaluations can be improved considerably through addition of key experimental data, such as even a single boiling temperature. Such additions can be done easily by the user, thus allowing ready addition of company proprietary data for essential species. This allows even the novice user to take full advantage of the TDE expert system for his or her limited data, which is critical for design of processes for new products.

This document contains some forward looking statements; it does not imply any contractual obligations.
In summary, *Aspen Plus*® with *NIST ThermoData Engine (TDE)* enables a new paradigm in Chemical Engineering. For the first time, engineers can use an expert system and extensive data archive of pure component structures and experimental thermophysical data to evaluate the key property parameters for a wide range of organic compounds on demand. This saves weeks or even months of effort, brings expert knowledge into the hands of less experienced engineers, and enables the development of more detailed and more accurate simulation models. Ultimately, it provides a unique opportunity for development of hundreds of new products and processes on an industrial scale.

**Improvements with V7**

AspenTech and NIST have continued to collaborate to improve the implementation of the NIST technology. With the release of aspenONE V7 in January of 2009, the NIST TRC database was expanded to include data for 18,840 pure components in V7.0, 19,724 species in V7.2, and over 24,000 species in V7.3. The property estimation engine in NIST TDE was improved to make better use of user-specified experimental data. These improvements have made it even easier to find or estimate the pure component parameters required to build accurate process models.

AspenONE V7 also introduced the NIST REFPROP (REFerence fluid PROPerties) models, developed under agreement with the National Institute of Standards and Technology's Standard Reference Data Program (SRDP). REFPROP is based on the most accurate pure fluid and mixture models currently available. The REFPROP models are designed specifically for utility fluids, including common refrigerants, steam, ammonia, natural gas, carbon dioxide, hydrogen, air, and noble gases. Engineers can apply these methods within Aspen Plus and HYSYS process simulations or within AspenTech’s heat exchanger design programs. These methods improve the reliability of the process and equipment designs, allowing engineers to have more confidence in their designs.

AspenONE V7.2 introduced the NIST implementation of the GERG\(^1\) 2008 equation of state model. GERG 2008 provides extremely accurate thermophysical property predictions for 21 pure components and their mixtures, representative of natural gas systems. The GERG model is appropriate for modeling rich natural gases, liquefied natural gases, liquefied petroleum gases, and hydrogen-light hydrocarbon mixtures in the gas, liquid, or

---

1 Groupe Européen de Recherches Gazières.
supercritical regions. The GERG model is the ISO-20765 reference equation for natural gas custody transfer.

**Update of NIST SOURCE Data Archive**

With NIST TRC processing approximately 500,000 data points per year, the scope of this database is expanding rapidly. However, access to these additional data within aspenONE had been limited by the frequency of the release. In order to bring the latest data to the users more quickly, NIST and AspenTech have collaborated to provide quarterly updates of the NIST TRC database on the AspenTech Customer Support website. Please contact AspenTech support engineers or search the online knowledge base to learn how to obtain the latest database file.

**Binary Property Data in aspenONE V7.3**

The NIST TDE feature in Aspen Plus V7.3 is significantly enhanced through the addition of over three million points of binary mixture and phase equilibrium data, including VLE data for over 30,000 unique pairs of components.

With V7.3, engineers and scientists can search the database and extract phase equilibrium, infinite dilution activity coefficient, and heat-of-mixing data for thousands of component pairs, saving additional weeks or months of effort. The database includes data sets for vapor-liquid equilibrium, liquid-liquid equilibrium, and solid-liquid equilibrium (solubility). These data are invaluable for validating or fitting binary coefficients for equations of state and/or activity coefficient models. Well used, these data can help significantly improve the predictive accuracy of a process simulation model, which in turn leads to more reliable equipment designs and more accurate cost estimates.

Binary mixture data is also available for a wide range of thermophysical and transport properties, including binary diffusion coefficients, mixture viscosities, heat capacities, etc. These can be especially valuable for improving equipment design. For example, heat exchanger sizing calculations are especially sensitive to the thermal conductivity and viscosity of the mixture. Column sizing and rating depends on the accuracy of mixture viscosity, density, and binary diffusion coefficients, especially when using predictive rate-based column models.

Each set of data includes full citations and uncertainty estimates. The data can be easily saved as property data objects. The existing Aspen data regression system can be used to evaluate the agreement between the model and the data, or you can regress the
binary parameters in your model to fine-tune the model predictions within a desired range of temperature, pressure, or composition.

The quality of binary VLE data can be assessed using built-in 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, Point or differential test, Infinite Dilution test, and Van Ness modeling test, with a check of consistency between the binary VLE data and the pure compound vapor pressures. Each set of data is assigned a numerical quality score, making it easy to compare the quality of data from different sources.

With aspenONE V7.3, you can use the NIST ThermoData Engine to extract binary mixture and phase equilibrium data for over 30,000 unique pairs of components. Within minutes you can get a list of available data sets for any pair of compounds, access a set of data, and validate the model predictions against the data or refit the data using existing data regression features in Aspen Properties or Aspen Plus.
Summary

Aspen Plus® with NIST ThermoData Engine (TDE) enables a new paradigm in Chemical Engineering. For the first time, engineers can use an expert system and an extensive data archive of pure component structures and experimental thermophysical property data to evaluate the key property parameters for a wide range of organic compounds on demand. This saves weeks or even months of effort, brings expert knowledge into the hands of less experienced engineers, and enables the development of more detailed and more accurate simulation models. Ultimately, it provides a unique opportunity for development of hundreds of new products and processes on an industrial scale in a cost-effective manner. The resulting models also help companies optimize operating conditions and run their processes with greater agility, helping them stay profitable in a volatile market.

Aspen Plus V7.3 builds on this foundation to deliver over three million points of binary mixture and phase equilibrium data. These data can be used together with Aspen’s Data Regression feature to evaluate the agreement between model predictions and data, or to fine-tune the model to better fit the data. On-demand access to binary data saves additional weeks or months of effort of collecting data. Ultimately, these data make it easier than ever before to improve model accuracy and rigor, resulting in more reliable equipment and process designs.