

Fit of g^E Model Parameters to Activity Coefficients Obtained from Group Contribution Methods

UNIQUAC, NRTL, Wilson, Margules
original UNIFAC, modified UNIFAC (Dortmund), ASOG, NIST–modified
UNIFAC, COSMO models

DDBSP – Dortmund Data Bank Software Package



DDBST

Dortmund Data Bank
Software & Separation
Technology

DDBST - Dortmund Data Bank Software & Separation Technology GmbH

Marie-Curie-Straße 10

D-26129 Oldenburg

Tel.: +49 441 361819 0

support@ddbst.com

www.ddbst.com

Table of Contents

1	Introduction.....	3
2	Component Selection.....	4
3	Calculating Activity Coefficients.....	6
4	Fitting.....	6
4.1	Starting Values.....	7
4.2	Objective Function.....	8
4.3	Loops.....	8
4.4	Fitting.....	8
4.5	Batch Fitting.....	9
5	Results.....	12
5.1	Recalculated Values.....	12
5.2	Parameters.....	14
5.3	Save in ParameterDDB.....	15
5.4	Save as Aspen INP Files.....	16
6	Import Aspen Components.....	16
7	The UNIFAC Consortium.....	16
8	Appendix.....	17
8.1	Conversion between DDB and Aspen Simulator Equations.....	17
8.1.1	Introduction.....	17
8.1.2	Wilson.....	17
8.1.3	NRTL.....	18
8.1.4	UNIQUAC.....	20

1 Introduction

This program fits g^E model (Wilson, NRTL, Margules, and UNIQUAC) parameters to activity coefficients obtained from group contribution models (original UNIFAC, modified UNIFAC (Dortmund), modified UNIFAC (Lyngby), NIST–modified UNIFAC, and ASOG) and two COSMO models for binary systems.

This software can obtain parameters for a high number of systems in a short time. These parameters have the same quality as the underlying activity coefficient model, the program neither adds or loses quality. There's still a slight degradation of precision because of a remaining fit deviation.

T [K]	x1	x2	Act.Coeff.1	Act.Coeff.2	Act.Coeff.1 gE Model	Act.Coeff.2 gE Model	Deviation1 [%]	Deviation2 [%]
300.00	0.0000	1.0000	2.0005	1.0000	2.0040	1.0000	-0.1727	0.0000
300.00	0.0005	0.9995	1.9994	1.0000	2.0023	1.0000	-0.1470	0.0000
300.00	0.0010	0.9990	1.9983	1.0000	2.0007	1.0000	-0.1214	-0.0001
300.00	0.0015	0.9985	1.9972	1.0000	1.9991	1.0000	-0.0959	-0.0002
300.00	0.0020	0.9980	1.9961	1.0000	1.9975	1.0000	-0.0705	-0.0003
300.00	0.0025	0.9975	1.9949	1.0000	1.9959	1.0000	-0.0507	-0.0005

Figure 1: Main dialog

This deviation normally doesn't exceed 1 percent (absolute mean error in the activity coefficients).

The software is integrated in the standard DDB software package using the DDB's component list, the parameter files, and the standard software components.

2 Component Selection

This part of the main dialog allows to specify the two components of the binary systems either the standard component selection dialog or by typing a known DDB component number.

Component 1
123 Naphthalene
UNIQUAC-q/r: 3.4400/4.9808 Wilson-vmol: 125.0110
Component 2
12 Diethyl ether
UNIQUAC-q/r: 3.0160/3.3949 Wilson-vmol: 104.7520

11 → Press Return → 11 Ethanol

The program displays some data relevant data for the g^E models

- UNIQUAC: Molecular surface and volume
- Wilson: Liquid volume at approx. 298 K

If any of these values are 0 the corresponding model cannot be used.

The values are stored in the basic component file and can be modified by the DDB Component Editor.

Component 1
12 Diethyl ether
UNIQUAC-q/r: 3.0160/3.3949 Wilson-vmol: 104.7520
Component 2
15987 3-Chloro-2-ethynyl-tetrahydrofuran
UNIQUAC-q/r: 3.5880/4.5696 Wilson-vmol: 0.0000

Density: 778.00 kg/m³ at
UNIFAC Groups and UNIQUAC Parameters
UNIFAC Groups: 2 1001 1020
mod. UNIFAC Groups: 2 1001 1020 0 0 0 0 0
UNIQUAC r-value: 1.89910
UNIQUAC q-value: 1.79600

Wilson volume is calculated from this density

UNIQUAC surface and volume

The two buttons with glasses at the right side of the component names open a dialog with additional information about necessary component specific data.

Component 1
12 Diethyl ether
UNIQUAC-q/r: 3.0160/3.3949 Wilson-vmol: 104.7520
Component 2
15987 3-Chloro-2-ethynyl-tetrahydrofuran
UNIQUAC-q/r: 3.5880/4.5696 Wilson-vmol: 0.0000

The UNIQUAC volumes and surfaces and the Wilson volume is repeated here but the lines about the group assignment (group numbers and names) for the four supported group contribution models add new important information. These groups must be available for all the components. Otherwise a group contribution model can't be used.

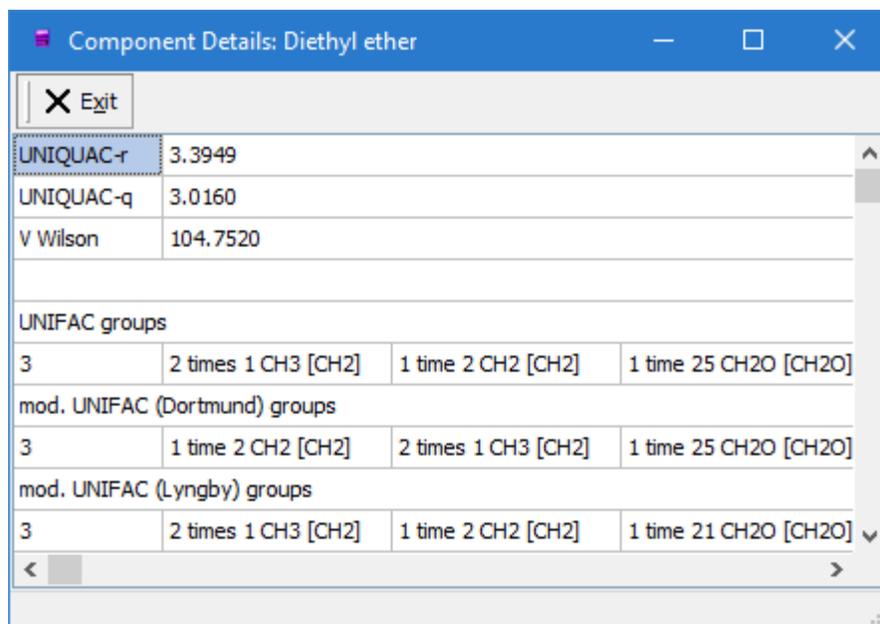


Figure 2: Component details

The button  opens a dialog with information of the availability of group assignments and interaction parameters for the different group contribution methods.

Green lines indicate that this model can be used to calculate activity coefficients – group assignment and interaction parameters are available. Red lines indicate either a missing group assignment or missing interaction parameters. The “Details” page allows to identify the problem.

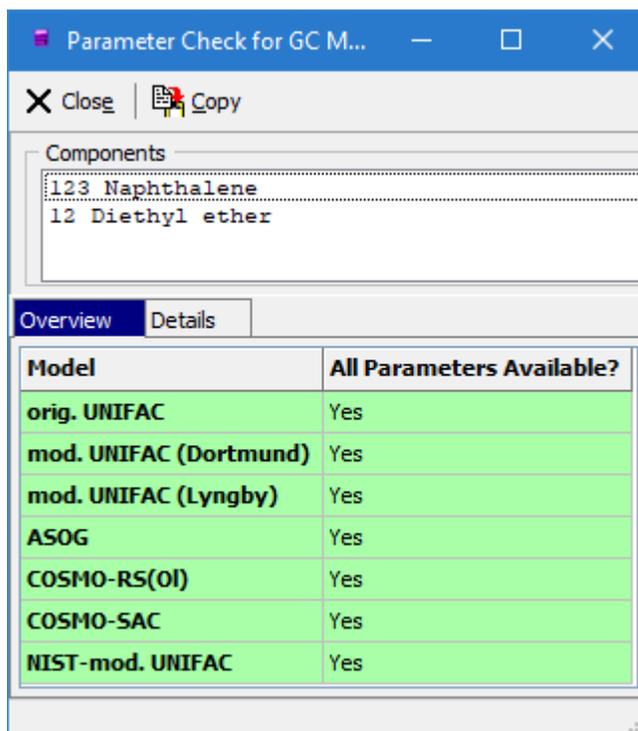


Figure 3: Parameter check

3 Calculating Activity Coefficients

Four different models are currently available for the calculation of activity coefficients:

- original UNIFAC
- modified UNIFAC (Dortmund)
- modified UNIFAC (Lyngby)
- ASOG
- COSMO-RS (Ol)
- COSMO-SAC
- NIST-modified UNIFAC

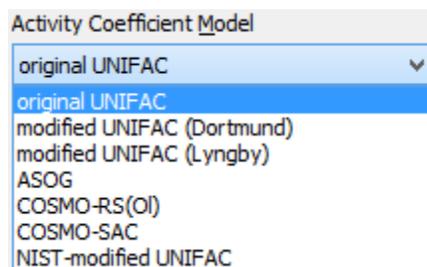


Figure 4: Activity coefficient models

The first two models are maintained by the UNIFAC consortium (www.unifac.org). The program recalculates the activity coefficient table on any model change and displays an error message if a calculation fails because of missing parameters or group assignments.

The last settings are the temperature range and the step width of the composition.



Figure 5: Temperature range

If the “Enhanced Resolution for Small Concentrations” option is set, the step width above 90 % and below 10 % is divided by a factor of 10 (e. g. from 5% to 0.5%) and the step width above 99% and below 1% is divided by a factor of 100 (e. g. from 5% to 0.05%).

If any of these values is changed the **Y** button has to be used to force the program to recalculate the activity coefficient table.

4 Fitting

The program allows to fit parameters for the g^E models UNIQUAC, NRTL, and Wilson.

For all three methods the temperature dependence of the parameters can be set to

$$a_{ij} + b_{ij} * T + c_{ij} * T^2 + d_{ij} * T * \ln(T) + e_{ij} * T^3 + \frac{f_{ij}}{T}$$

All parameters can be switched on and off.

The “Quick Select” combo box

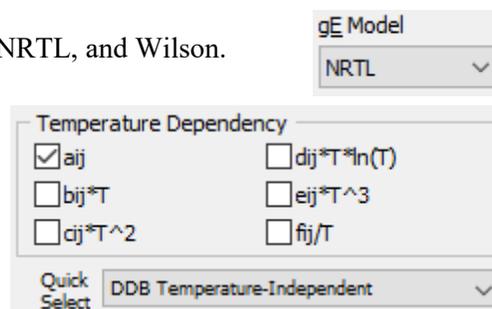


Figure 6: Temperature dependency

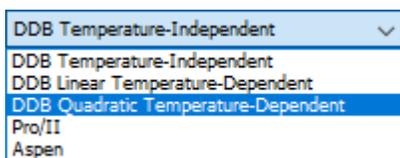


Figure 7: Quick select

defines some short cuts to often used parameter sets.

4.1 Starting Values

Starting values for the parameters can be modified by Start Parameters :

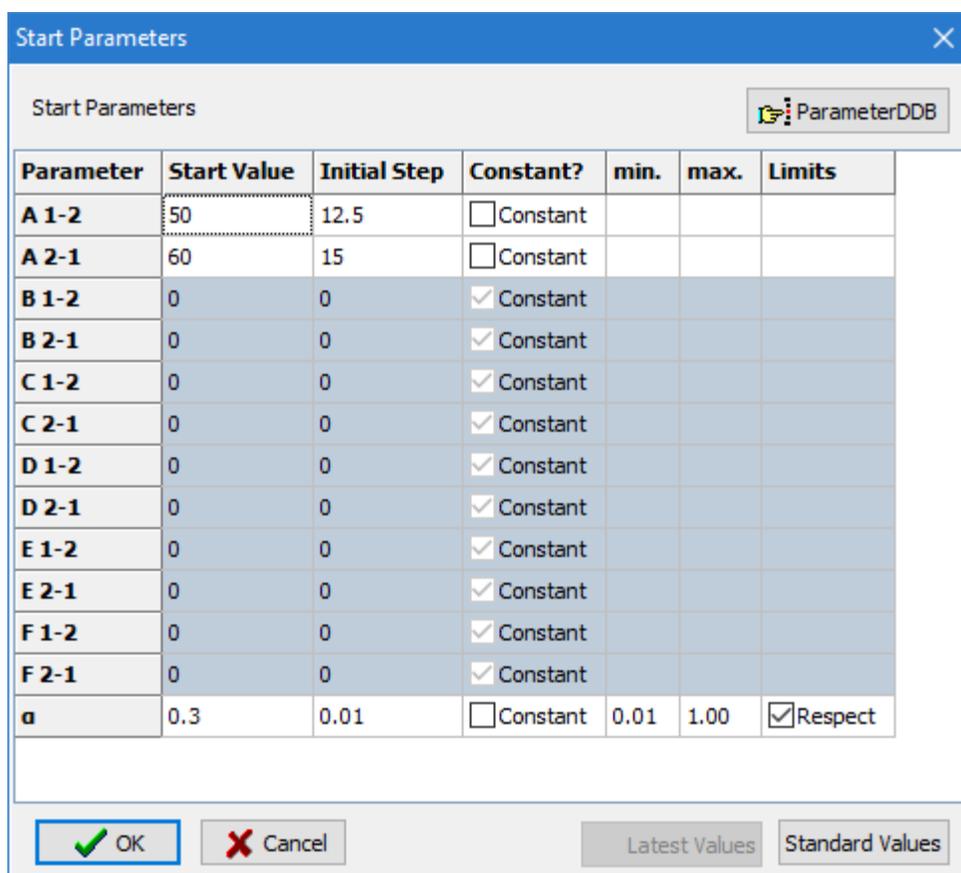
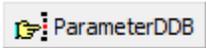


Figure 8: Start parameters

The grid is editable and allows to enter starting values as well as initial step widths. A parameter can be fixed by setting the initial step to 0. or, better, by checking the “Constant” check box. Unchecking this control will bring back the previously entered step width.

The Latest Values button allows using the parameters from the last fit and the Standard Values button resets

the parameters to some useful preset values. With the  button it is possible to search the DDB parameter data bank for existing parameters and to select these parameters as starting values.

Fixing the NRTL α parameter can either be done by simply setting the step width to zero or by checking “Constant” - this works exactly the same as for the normal parameters. Setting a variable α is done by entering an initial step different to zero or by unchecking “Constant”.

The last possible setting is to limit α to a maximum and minimum value.

4.2 Objective Function

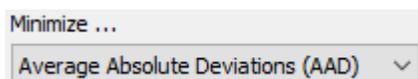


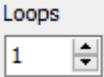
Figure 9: Objective function

The objective function can be set to

- Average Absolute Deviations
- Root Mean Square Deviations
- Mean Relative Deviation

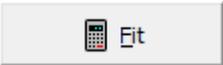
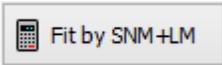
All three objective function have different advantages and downsides.

4.3 Loops

 The “Loops” setting is used to restart the fit procedure up to four times. This is a useful and often necessary feature because the changes in the objective function are rather small near the minimum and the optimization process is starving before reaching a *global* minimum. The re-start uses the obtained parameters from the previous fit but starts with the step widths of the initial fit.

This procedure guarantees in many cases a better minimum (in deviations) than a single fit.

4.4 Fitting

The  and  buttons starts the fit. The used algorithm is Simplex-Nelder-Mead (SNM) or a combined Simplex-Nelder-Mead and Levenberg-Marquardt (LM). This SNM optimization algorithm is known to be robust and will not fail. The drawback of this stability is that the SNM method needs a lot of function calls and can be slow. LM is far quicker but tends to fail in more cases.

The currently obtained objective function value and the parameters are displayed in a separate window.

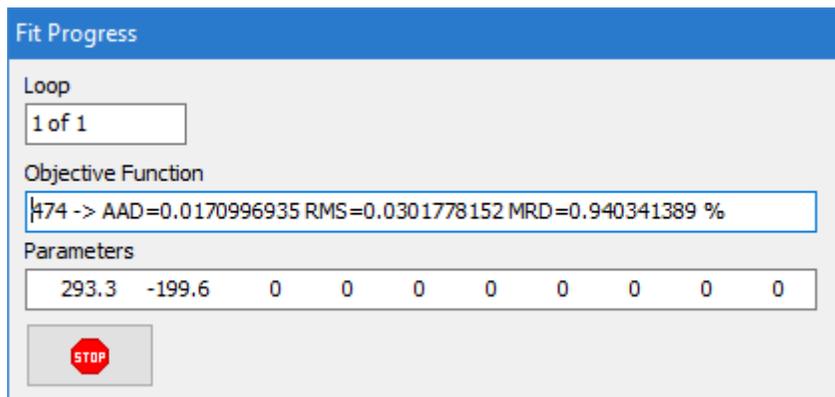


Figure 10: Fit progress

The  button allows stopping the current fit process.

The sequence of the parameters are always

- NRTL α first (if NRTL has been selected)
- A_{12} and A_{21} (temperature-independent parameter)
- B_{12} and B_{21} (linear temperature-dependent parameter)
- C_{12} and C_{21} (quadratically temperature-dependent parameter)
- and so on

4.5 Batch Fitting

The program allows to fit multiple systems in a row. These systems will be build from component lists which can be created by other DDB software like the component management program, the Artist structure editor, or the DDB main retrieval program.

The dialogs

#	#DDB	Name
1	31	Benzene
2	59	Divinylbenzene
3	121	alpha-Methyl styrene
4	123	Naphthalene
5	150	Styrene
6	415	1-Phenylnaphthalene
7	471	Biphenyl
8	521	m-Terphenyl
9	763	tert-Butylbenzene
10	797	2-Phenyl-2,4,6-trimethylheptane
11	925	Anthracene

Figure 12: List of components

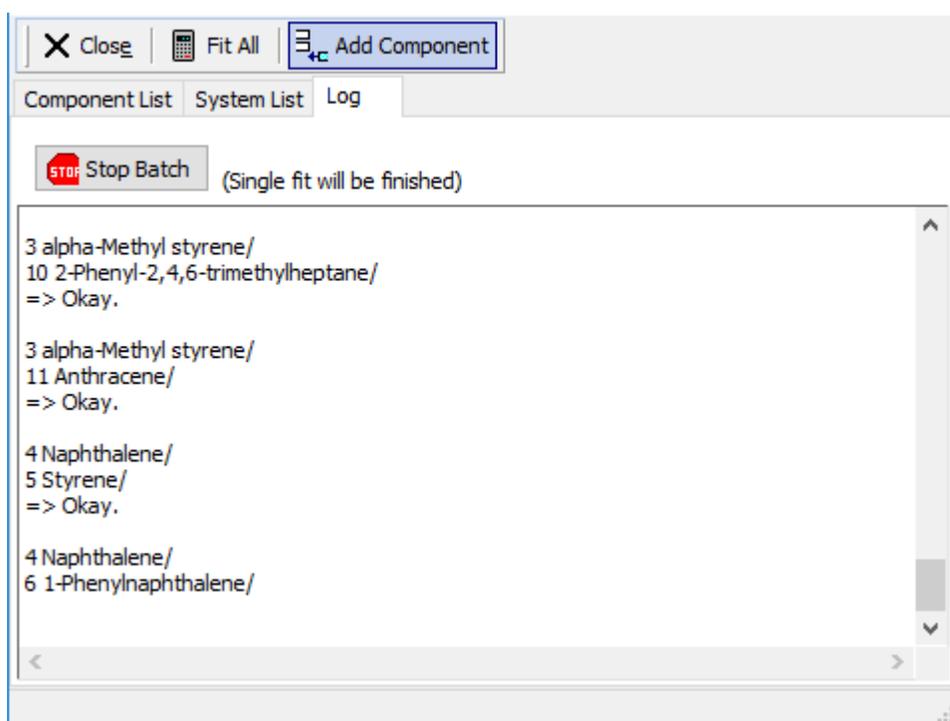
#	Component1	Component2	Fit?
1	1 Benzene/	2 Divinylbenzene/	Fit
2	1 Benzene/	3 alpha-Methyl styrene/	Fit
3	1 Benzene/	4 Naphthalene/	Fit
4	1 Benzene/	5 Styrene/	Fit
5	1 Benzene/	6 1-Phenylnaphthalene/	Fit
6	1 Benzene/	7 Biphenyl/	Fit
7	1 Benzene/	8 m-Terphenyl/	Fit
8	1 Benzene/	9 tert-Butylbenzene/	Fit
9	1 Benzene/	10 2-Phenyl-2,4,6-trimethylheptane/	Fit

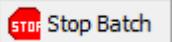
31 Benzene
59 Divinylbenzene

Fit Method
 Simplex Method by Nelder/Mead (SNM)
 SNM+Levenberg/Marquardt Method

Figure 11: List of systems

display both the list of components and the list of binary systems built from these components. The  button in the single grid rows will start a fit only for the single binary system displayed in the single row whereas the  button in the tool bar will start a batch fit for all the systems. This batch fit produces a protocol which is displayed in the third page of this dialog. The method used for the fitting



The batch fit can be stopped by the  button. A

single fit will be finished before the batch stops. Any results from the already executed fits are saved in the “Result” windows (see next chapter) and are not lost.

The log window itself only stores information on the success of a fit or an error message if a fit can't be performed due to, for example, missing group/group interaction parameters of a UNIFAC model.

5 Results

5.1 Recalculated Values

The recalculated values are shown alongside the source values in the data grid at the bottom of the dialog.

T [K]	x1	x2	Act.Coeff.1	Act.Coeff.2	Act.Coeff.1 gE Model	Act.Coeff.2 gE Model	Deviation1 [%]	Deviation2 [%]
300.00	0.0000	1.0000	0.9728	1.0000	0.9736	1.0000	-0.0775	0.0000
300.00	0.0500	0.9500	0.9758	0.9999	0.9764	0.9999	-0.0637	-0.0025
300.00	0.1000	0.9000	0.9786	0.9997	0.9792	0.9997	-0.0571	0.0002
300.00	0.1500	0.8500	0.9812	0.9993	0.9818	0.9993	-0.0569	-0.0019

Figure 13: Result table

This table also shows the relative deviations.

This table can be saved as Excel file, copied to the Windows clipboard, and a diagram with the results can be shown.

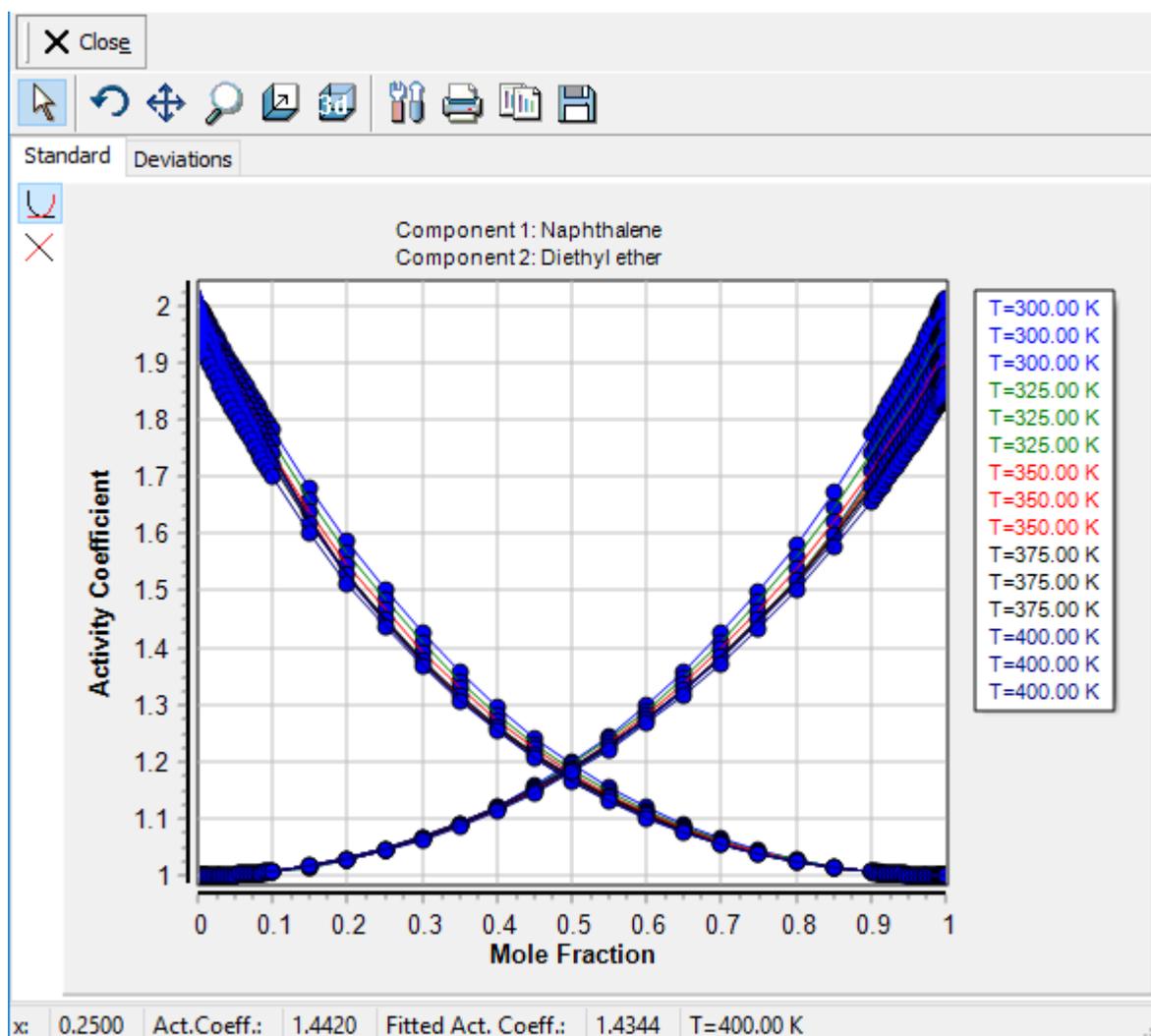


Figure 14: Result diagram

The diagram allows switching between a normal scale and a logarithmic scale, it can also be copied and saved, and a 3D display is also possible.



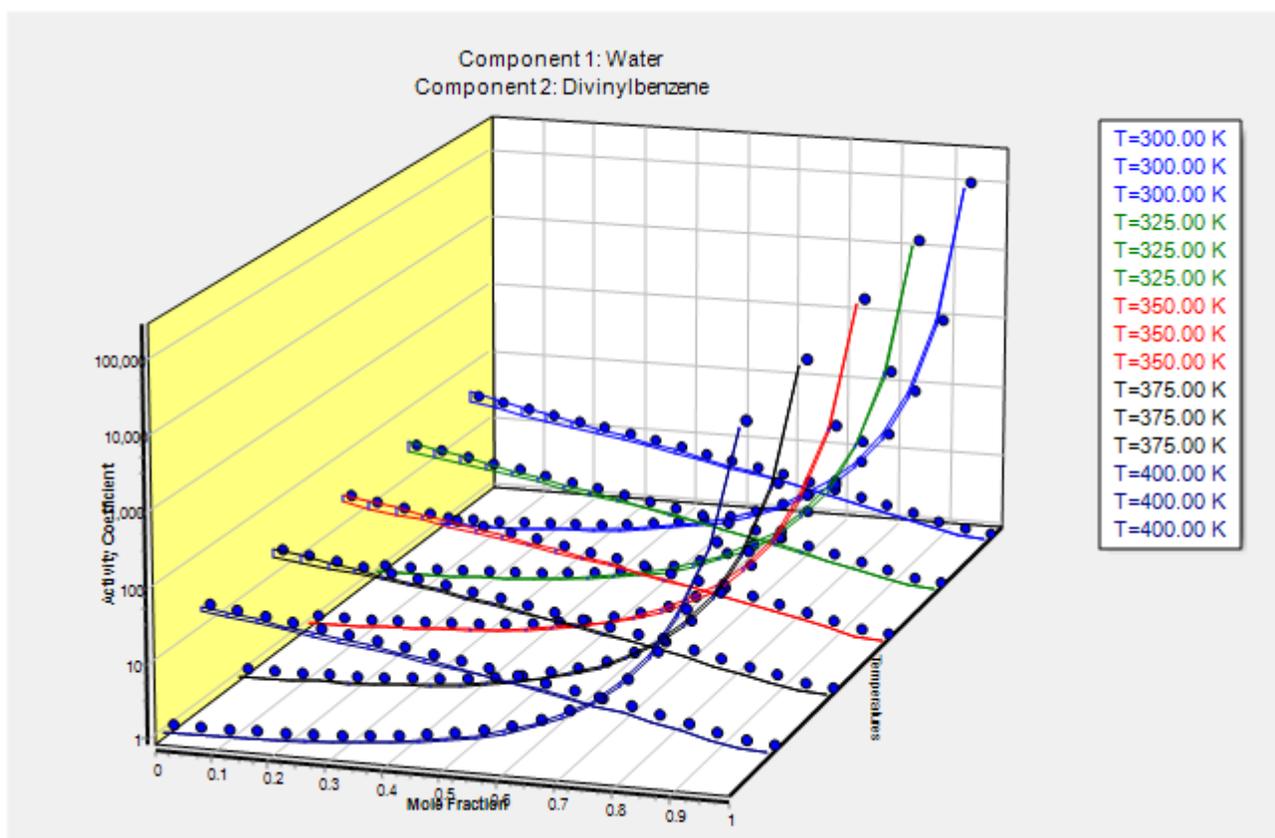


Figure 15: Result diagram 3D

5.2 Parameters

The obtained parameters are also displayed in a table. This table is used to collect all fit results.

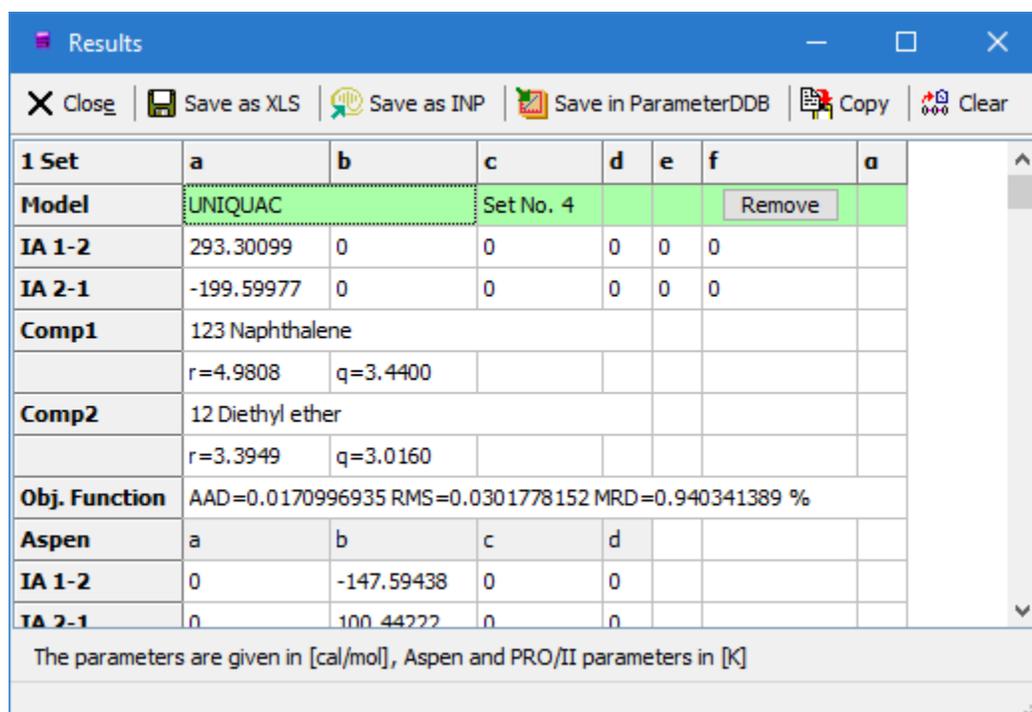
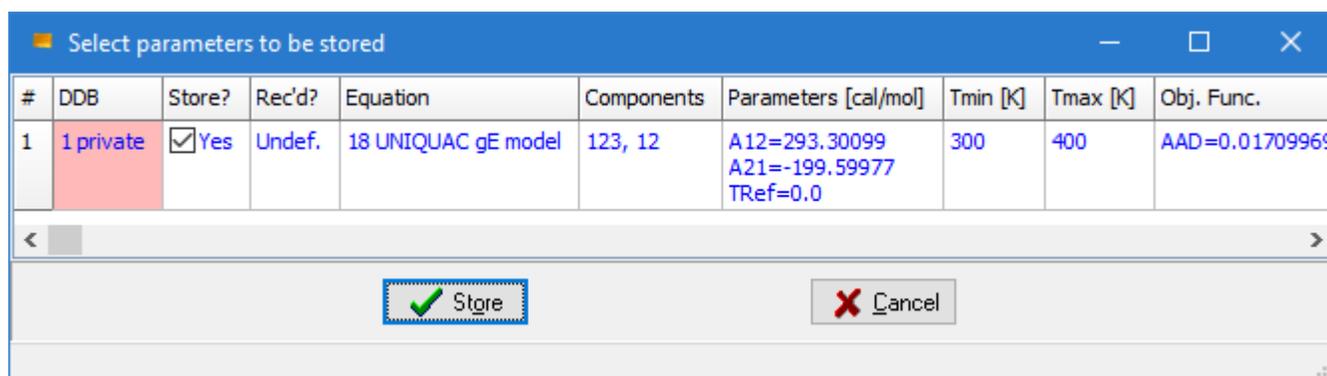


Figure 16: Fit Results Dialog

The parameters sets can be saved as Microsoft Excel file, as Aspen INP file, to the ParameterDDB, and copied to the Windows clipboard.

The parameters saved in the Aspen INP file are fully Aspen compatible whereas all other locations store the parameters in a DDB software specific format. The ParameterDDB management software supports to copy g^E model parameters in the DDB format as well as in the Aspen format and the Pro/II format.

5.3 Save in ParameterDDB



This export is displaying an additional selection dialog where all fitted parameter sets are listed. Single datasets can be included and excluded here before storing the entire list.

This ParameterDDB is also used to export parameters in various other formats.

5.4 Save as Aspen INP Files

This export only needs the specification of a file name to create an INP file. This fit program is able to perform a match of the internal (DDB) components with the Aspen IDs and will write appropriate Aspen names in the INP file. If a match can't be found a unique identifier is used which can be easily modified in Aspen.

6 Import Aspen Components

This feature allows to import components from Aspen simulator files. Supported files are the INP, the BKP, and the APT files.

This fit program will not only load the component identifiers but it also loads group assignments which will be matched automatically to DDB group assignments.

After a fit has been performed the program will write an INP file with the appropriate Aspen component identifiers allowing to seamlessly import the fitted g^E model parameters.

7 The UNIFAC Consortium

This program fully supports the parameter files developed by the UNIFAC consortium (www.unifac.org).

8 Appendix

8.1 Conversion between DDB and Aspen Simulator Equations

8.1.1 Introduction

This section describes the different forms of the g^E model equations used in the Aspen simulator (short Aspen) and the DDB software package (short DDBSP) for the three models Wilson, NRTL, and UNIQUAC. Wherever it is possible a conversion of parameters is described.

8.1.2 Wilson

DDBSP Form

$$\ln y_i = -\ln\left(\sum_j x_j \Lambda_{ij}\right) + 1 - \sum_k \frac{x_k \Lambda_{ki}}{\sum_j x_j \Lambda_{kj}}$$

with

$$\Lambda_{ij} = \frac{v_j}{v_i} e^{-\Delta \lambda_{ij}/T}$$

v_i Molar volume of liquid pure component i at 25°C

$\Delta \lambda_{ij}$ are interaction parameters between component i and j [cal/mol] and can be made temperature-dependent:

$$\Delta \lambda_{ij} = a_{ij} + b_{ij}T + c_{ij}T^2 + d_{ij}T \cdot \ln T + e_{ij}T^3 + \frac{f_{ij}}{T}$$

Aspen Form

$$\ln y_i = 1 - \ln\left(\sum_j x_j \Lambda_{i,j}\right) - \frac{\sum_k x_k \Lambda_{k,j}}{\sum_j x_j \Lambda_{k,j}}$$

$$\Lambda_{i,j} = \exp\left(a_{i,j} + \frac{b_{i,j}}{T} + c_{i,j} \ln T + d_{i,j}T\right)$$

Conversion between DDBSP and Aspen Parameters

The Aspen form does not support the terms for e_{ij} and f_{ij} in the temperature-dependency of the interaction parameters. Therefore parameter sets where e_{ij} or f_{ij} are non-zero can not be converted to the Aspen equation.

On the other hand, Aspen incorporated the ratio of molar volumes at 25°C $\left(\frac{v_j}{v_i}\right)$ within the parameter $a_{i,j}$.

This makes it impossible to convert Aspen-style Wilson parameters into DDBSP parameters if any of the pure component Wilson volumes are missing.

The last difference is that DDBSP and Aspen are using different units for the interaction parameters – [cal/mol] in DDBSP and [K] in Aspen which leads to the conversion factor R (gas constant) in $\frac{\text{cal}}{\text{K}\cdot\text{mol}}$.

Aspen parameters	DDBSP parameters
$a_{i,j}$	$\ln\left(\frac{v_j}{v_i}\right) - \frac{b_{ij}}{R}$
$b_{i,j}$	$-\frac{a_{ij}}{R}$
$c_{i,j}$	$-\frac{d_{ij}}{R}$
$d_{i,j}$	$-\frac{c_{ij}}{R}$

$$R = 1.9858775 \frac{\text{cal}}{\text{K}\cdot\text{mol}}$$

8.1.3 NRTL

DDBSP Form

$$\ln y_i = \frac{\sum_j \tau_{ji} G_{ji} x_j}{\sum_k G_{ki} x_k} + \sum_j \frac{x_j G_{ij}}{\sum_k G_{kj} x_k} \left(\tau_{ij} - \frac{\sum_n x_n \tau_{nj} G_{nj}}{\sum_k G_{kj} x_k} \right)$$

with

$$G_{i,j} = \exp(-\alpha_{ij} \tau_{ij})$$

and

$$\tau_{ij} = \frac{\Delta g_{ij}}{T}$$

Δg_{ij} are interaction parameters for component i and j [cal/mol]

$$\Delta g_{ij} = a_{ij} + b_{ij} T + c_{ij} T^2 + d_{ij} T \cdot \ln T + e_{ij} T^3 + \frac{f_{ij}}{T}$$

α_{ij} is the non-randomness parameter: $\alpha_{ij} = \alpha_{ji}$.

Aspen Form

$$\ln y_i = \frac{\sum_j \tau_{ji} G_{ji} x_j}{\sum_k G_{ki} x_k} + \sum_j \frac{x_j G_{ij}}{\sum_k G_{kj} x_k} \left(\tau_{ij} - \frac{\sum_l x_l \tau_{lj} G_{lj}}{\sum_k G_{kj} x_k} \right)$$

with

$$G_{ji} = \exp(-S_{ji} \tau_{ji})$$

and

$$\tau_{ji} = a_{j,i} + \frac{b_{j,i}}{T} + e_{j,i} \ln T + f_{j,i} T$$

and

$$S_{ji} = c_{ji} + d_{ji}(T - 273.15) \quad (\text{non-randomness parameter})$$

Conversion between DDBSP and Aspen Forms

The conversion between DDBSP and Aspen parameters is mainly a conversion of units. DDBSP uses [cal/mol] and Aspen uses [K] which leads to the conversion factor R (gas constant) in $\frac{\text{cal}}{\text{K} \cdot \text{mol}}$.

DDBSP does not support a temperature-dependent non-randomness parameter which means that the Aspen parameter $d_{i,j}$ is always 0 and an Aspen parameter set including a temperature-dependency of the non-randomness parameter cannot be converted into DDBSP parameters.

The sequence of parameters is also slightly different.

Aspen parameters	DDBSP parameters
$a_{i,j}$	$\frac{b_{ij}}{R}$
$b_{i,j}$	$\frac{a_{ij}}{R}$
$c_{i,j}$	α_{ij}
$d_{i,j}$	0
$e_{i,j}$	$\frac{d_{ij}}{R}$
$f_{i,j}$	$\frac{c_{ij}}{R}$

$$R = 1.9858775 \frac{\text{cal}}{\text{K} \cdot \text{mol}}$$

8.1.4 UNIQUAC

DDBSP Form

$$\ln y_i = \ln y_i^C + \ln y_i^R$$

$$\ln y_i^C = 1 - V_i + \ln V_i - 5 q_i \left(1 - \frac{V_i}{F_i} + \ln \frac{V_i}{F_i} \right)$$

$$\ln y_i^R = q_i \left(1 - \ln \frac{\sum_j q_j x_j \tau_{ji}}{\sum_j q_j x_j} - \sum_j \frac{q_j x_j \tau_{ij}}{\sum_k q_k x_k \tau_{kj}} \right)$$

with

$\ln y_i^C$ Combinatorial part of the activity coefficient of component i

$\ln y_i^R$ Residual part of the activity coefficient of component i

$\tau_{ij} = e^{-\Delta u_{ij}/T}$ Interaction parameters between component i and j [cal/mol]

$$\Delta u_{ij} = a_{ij} + b_{ij} T + c_{ij} T^2 + d_{ij} T \cdot \ln T + e_{ij} T^3 + \frac{f_{ij}}{T}$$

$$V_i = \frac{r_i}{\sum_j r_j x_j} \quad \text{Volume fraction}$$

$$F_i = \frac{q_i}{\sum_j q_j x_j} \quad \text{Surface fraction}$$

r_i Relative van der Waals volume of component i

q_i Relative van der Waals surface of component i

Aspen Form

The activity coefficient is calculated from a residual and a combinatorial part.

$$\ln y_i = \ln y_i^C + \ln y_i^R$$

The combinatorial part is determined by

$$\ln y_i^C = \ln \frac{V_i}{x_i} + \frac{z}{2} q_i \ln \frac{F_i}{V_i} + l_i - \frac{V_i}{x_i} \sum_j x_j l_j$$

The coordination number is $z = 10$

The residual part is determined by

$$\ln y_i^R = q_i \left(1 - \ln \sum_j F_j \tau_{j,i} - \sum_j \frac{F_j \tau_{i,j}}{\sum_k F_k \tau_{k,j}} \right) \quad \text{or} \quad \ln y_i^R = q'_i \left(1 - \ln \sum_j F'_j \tau_{j,i} - \sum_j \frac{F'_j \tau_{i,j}}{\sum_k F'_k \tau_{k,j}} \right)$$

with

$$l_i = 5(r_i - q_i) - (r_i - 1)$$

$$V_i = \frac{r_i x_i}{\sum_j r_j x_j} \quad \text{Volume fraction}$$

$$F_i = \frac{q_i x_i}{\sum_j q_j x_j} \quad \text{or} \quad F'_i = \frac{q'_i x_i}{\sum_j q'_j x_j} \quad \text{Surface fraction, the F' version is used for alcohols and Water.}$$

$$\tau_{j,i} = \exp \left(a_{j,i} + \frac{b_{j,i}}{T} + c_{j,i} \ln T + d_{j,i} T \right)$$

Conversion between DDBSP and Aspen Form

The conversion between DDBSP and Aspen parameters is mainly a conversion of units. DDBSP uses [cal/mol] and Aspen uses [K] which leads to the conversion factor R (gas constant) in $\frac{\text{cal}}{\text{K} \cdot \text{mol}}$.

There's one bigger difference regarding alcohols and Water containing mixtures. Aspen uses a slightly different equation form for alcohols and Water mainly caused by a different relative van der Waals surface constant (q') making it impossible to convert DDBSP parameters into Aspen parameters if these components are encountered.

Aspen parameters	DDBSP parameters
$a_{i,j}$	$-\frac{b_{ij}}{R}$
$b_{i,j}$	$-\frac{a_{ij}}{R}$
$c_{i,j}$	$-\frac{d_{ij}}{R}$
$d_{i,j}$	$-\frac{c_{ij}}{R}$

$$R = 1.9858775 \frac{\text{cal}}{\text{K} \cdot \text{mol}}$$