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
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.

![Figure 1: Main dialog](image)

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.

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.

The two buttons with glasses at the right side of the component names open a dialog with additional information about necessary component specific data.
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](image1)

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](image2)
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 (OI)
- COSMO-SAC
- NIST-modified UNIFAC

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.

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 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 gE 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
defines some short cuts to often used parameter sets.

4.1 Starting Values

Starting values for the parameters can be modified by:

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 button allows using the parameters from the last fit and the 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 $\alpha$ 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 $\alpha$ is done by entering an initial step different to zero or by unchecking “Constant”.

The last possible setting is to limit $\alpha$ to a maximum and minimum value.

### 4.2 Objective Function

![Figure 9: Objective function](image)

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.

![Fit Progress](image)

Figure 10: Fit progress

The **STOP** button allows stopping the current fit process.

The sequence of the parameters are always:

- NRTL $\alpha$ 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
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.

![Figure 13: Result table](image)

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.
5.2 Parameters

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

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

with

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

$v_i$ 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:

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

**Aspen Form**

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

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

**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 cannot 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 \( \frac{\text{cal}}{\text{K} \cdot \text{mol}} \).

<table>
<thead>
<tr>
<th>Aspen parameters</th>
<th>DDBSP parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{i,j} )</td>
<td>( \ln \left( \frac{v_j}{v_i} \right) - \frac{b_{ij}}{R} )</td>
</tr>
<tr>
<td>( b_{i,j} )</td>
<td>( -\frac{a_{ij}}{R} )</td>
</tr>
<tr>
<td>( c_{i,j} )</td>
<td>( -\frac{d_{ij}}{R} )</td>
</tr>
<tr>
<td>( d_{i,j} )</td>
<td>( -\frac{c_{ij}}{R} )</td>
</tr>
</tbody>
</table>

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

### 8.1.3 NRTL

**DDBSP Form**

\[
\ln \gamma_i = \sum_j \tau_{ij} G_{ji} x_j + \sum_k \sum_j G_{kj} x_k \left( \tau_{ij} \frac{\sum_n x_n \tau_{nj} G_{nj}}{\sum_k G_{kj} x_k} \right)
\]

with

\[
G_{i,j} = \exp \left( -\alpha_{ij} \tau_{ij} \right)
\]

and

\[
\tau_{ij} = \frac{\Delta g_{ij}}{T}
\]

\( \Delta g_{ij} \) are interaction parameters for component \( i \) and \( j \) [cal/mol]

\[
\Delta g_{ij} = a_{ij} + b_{ij} T + c_{ij} T^2 + d_{ij} T \ln T + e_{ij} T^3 + f_{ij} T
\]

\( \alpha_{ij} \) is the non-randomness parameter: \( \alpha_{ij} = \alpha_{ji} \).
Aspen Form

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

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

\[S_{ji} = c_{ji} + d_{ji} (T - 273.15)\] (non-randomness parameter)

Conversion between DDBSP and Aspen Forms

The conversion between DDBSP and Aspen parameters is mainly a conversion of units. DDBSP uses \([\text{cal/mol}]\) and Aspen uses \([\text{K}]\) 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 \(d_{j,i}\) 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.

<table>
<thead>
<tr>
<th>Aspen parameters</th>
<th>DDBSP parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_{i,j})</td>
<td>(b_{ij}) / (R)</td>
</tr>
<tr>
<td>(b_{i,j})</td>
<td>(a_{j,i}) / (R)</td>
</tr>
<tr>
<td>(c_{i,j})</td>
<td>(\alpha_{ij})</td>
</tr>
<tr>
<td>(d_{i,j})</td>
<td>0</td>
</tr>
<tr>
<td>(e_{i,j})</td>
<td>(d_{ij}) / (R)</td>
</tr>
<tr>
<td>(f_{i,j})</td>
<td>(c_{ij}) / (R)</td>
</tr>
</tbody>
</table>

\[R = 1.985 \, 8775 \frac{\text{cal}}{\text{K} \cdot \text{mol}}\]
8.1.4 UNIQUAC

**DDBSP Form**

\[ \ln y_i = \ln y_i^C + \ln y_i^R \]

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

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

**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

\[ \ln y_i^C = \ln \frac{V_i}{x_i} + \frac{z}{2} q_i \ln \frac{F_i}{V_i} + l_i - \frac{V_i}{x_i} \sum_j x_j l_{ij} \]

The coordination number is \( z = 10 \)

The residual part is determined by
\[
\ln \gamma_i^R = q_i \left( 1 - \ln \sum_j F_{j,i} \tau_{j,i} - \sum_k \frac{F_j \tau_{i,j}}{k} \right) \quad \text{or} \quad \ln \gamma_i^R = q_i' \left( 1 - \ln \sum_j F_{j,i} \tau_{j,i} - \sum_k \frac{F_j' \tau_{i,j}}{k} \right)
\]

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.}\]

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

**Conversion between DDBSP and Aspen Form**

The conversion between DDBSP and Aspen parameters is mainly a conversion of units. DDBSP uses \([\text{cal/mol}]\) and Aspen uses \([\text{K}]\) 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.

<table>
<thead>
<tr>
<th>Aspen parameters</th>
<th>DDBSP parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_{i,j})</td>
<td>(- \frac{b_{j,i}}{R})</td>
</tr>
<tr>
<td>(b_{i,j})</td>
<td>(- \frac{a_{j,i}}{R})</td>
</tr>
<tr>
<td>(c_{i,j})</td>
<td>(- \frac{d_{j,i}}{R})</td>
</tr>
<tr>
<td>(d_{i,j})</td>
<td>(- \frac{c_{j,i}}{R})</td>
</tr>
</tbody>
</table>

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