# PEAXACT
# User Manual

Version 3.7.2
2015-06-23

S•PACT GmbH
Burtscheider Str. 1
52064 Aachen
Germany

phone: +49 241 9569 9812
fax: +49 241 4354 4308
e-mail: support@s-pact.de

# CONTENTS

# 1 QUICK START

## 1.1 What is PEAXACT?

PEAXACT is a comprehensive software product for *quantitative analysis of spectroscopic and chromatographic data* using several state-of-the-art and innovative analysis methods. PEAXACT contains ready-to-go calibration-free methods as well as calibration methods for the prediction of features from unknown samples. It offers quick-and-easy access to established methods like Peak Integration or Partial-Least-Squares and is the only software product that provides the innovative ***Indirect Hard Modeling*** method.

PEAXACT can be used for laboratory analytics as well as for process analytics and reaction monitoring either interactively using the user interface or remotely in combination with third-party software.

### Separation to Related Software Products

PEAXACT strongly focuses on quantitative data evaluation of spectra, i.e. determination of the quantity of pure component and mixture properties from measured spectra. PEAXACT is optimized for IR and Raman spectra, but can also be used for other kinds of peak-shaped data like NMR spectra and chromatograms.

PEAXACT is not designed for the identification of chemical structures of measured compounds.

PEAXACT does not replace measurement software for spectrometers such as OPUS (Bruker Optics), HoloGrams (Kaiser Optical Systems Inc.), iC-Software-Suite (Mettler Toledo), and others. In fact, PEAXACT operates on data measured with such software products.

PEAXACT is equipped with several data analysis methods. Some of these methods may also be distributed with other software tools. However, unique hard modeling methods like Indirect Hard Modeling (IHM) and Hard Modeling Factor Analysis (HMFA) are available in PEAXACT only.

## 1.2 Getting Help

### User Manual

This user manual documents a certain version of PEAXACT. You can find the version number and publication date on the title page.

We are continuously working on improving the manual. The latest document version is distributed as PDF file with each PEAXACT software update. The file is located in subdirectory `help` of the PEAXACT installation directory.

### Technical Support

The Technical Support can be contacted in different ways:

- E-mail to support@s-pact.de
- Web form at http://www.s-pact.de/support
- From the PEAXACT desktop (see Request Support)

> Note: A subscription of S•PACT Software Maintenance Service (SMS) is required to be eligible for technical support. The first year of SMS is included with new PEAXACT product licenses.

### Blog

The PEAXACT Blog was launched as a free source of information complementary to the user manual. It contains tutorials, how-tos, and tips & tricks.
See: http://www.s-pact.de/blog

# 1.3 Installation

## 1.3.1 Requirements

- Microsoft Windows XP SP2 or later (32 bit or 64 bit)
- Intel or AMD x86 or x64 CPU with SSE2 support (2 GHz recommended)
- 1 GB of disk space (2 GB recommended)
- 1 GB RAM (2 GB recommended)

## 1.3.2 Licensing

PEAXACT software is furnished under a license agreement. The software may be used only under the terms of the license agreement.

License conditions vary with respect to license type, license option, and license modules. This Section only gives a short overview of the different license options. For the full and legally valid conditions please refer to the license agreement document.

> Note: Depending on the modules of your particular license, you may not be able to access all PEAXACT features.

### License Options

**Individual**

PEAXACT can be installed and operated by a single designated named user on up to 2 designated computers.

**Group**

PEAXACT can be installed on a specified number of computers (no remote access computers / no terminal servers) and can be operated by an unlimited group of users (one user per computer at a time).

**Network**

PEAXACT can be installed on any computer in a network (including remote access computers / terminal servers) and can be operated by a specified number of users simultaneously.

## 1.3.3 Installation

### Step 1: Before You Install

- Make sure your computer fulfills the system requirements.
- When upgrading an existing installation, visit http://www.s-pact.de/peaxact/whatsnew and read the upgrade notes and compatibility considerations.
- Make sure you have administrator privileges to perform the installation.
- Make sure your license is valid for the major version you want to install. If you do not have a license yet, you can get a free trial license or purchase a license after installation.

Note: The PEAXACT version is a concatenation of three numbers `<major version>.<minor version>.<maintenance version>`, e.g. `3.7.0`.

### Step 2: Install PEAXACT

- Download the latest PEAXACT Installer from http://www.s-pact.de/downloads

Note: The installer's filename is `peaxactInstaller_<version>_<arch>.exe` `<version>` is the version number; `<arch>` is the software architecture (`win32` or `win64`). PEAXACT is available for 32 bit and 64 bit Windows platforms. The 32 bit version also runs on 64 bit platforms, but not vice versa.

Note: Only one installation of PEAXACT can exist at a time. Installing a newer version will update the existing installation automatically. Installing a different architecture or older version requires uninstallation of the existing version first.

**Online Installation (Web Installation)**

- If you are going to install PEAXACT on a computer which is connected to the internet, you do not need to download any additional files.

- Run the PEAXACT Installer and follow the setup instructions. Additional runtime packages are downloaded and installed automatically if detected missing.

**Offline Installation**

- If you are planning to install PEAXACT on a computer without internet access, you have to download additional runtime packages in advance from http://www.s-pact.de/peaxact/runtime
- Make sure to download runtime packages for the same architecture as the PEAXACT installer (32 bit or 64 bit)
- Save all installer files without renaming them to one folder on your hard drive / flash drive.
- Run the PEAXACT Installer file from this folder and follow the setup instructions. Runtime packages are installed automatically if detected missing.

### Step 3: After Installation

- After a new product installation, continue with License Activation.
- After upgrading an existing installation, check the upgrade notes at http://www.s-pact.de/peaxact/whatsnew for further upgrade steps.

## 1.4 License Activation

License activation involves loading a valid license file. If you already have a license file, go ahead to step 3.

### Step 1: Find out the computer's Host ID

> Note: This step is required for purchased licenses only! For free licenses, proceed with step 2.

For purchased licenses, activation associates the use of PEAXACT with designated computers by means of a Host ID. The Host ID is a MAC address (format xx-xx-xx-xx-xx-xx) or the serial number of volume c (format xxxx-xxxx) of the computer on which PEAXACT is installed.

- Click the Windows start menu and select **Programs > PEAXACT > Activate PEAXACT**
- Wait until the License Activation Dialog is displayed
- Take the Host ID from the dialog window, then click Cancel.
- If you purchased a license for multiple computers, get one Host ID for each computer.

> Note: You can also type `getmac` at the command prompt and use the first MAC address as Host ID.

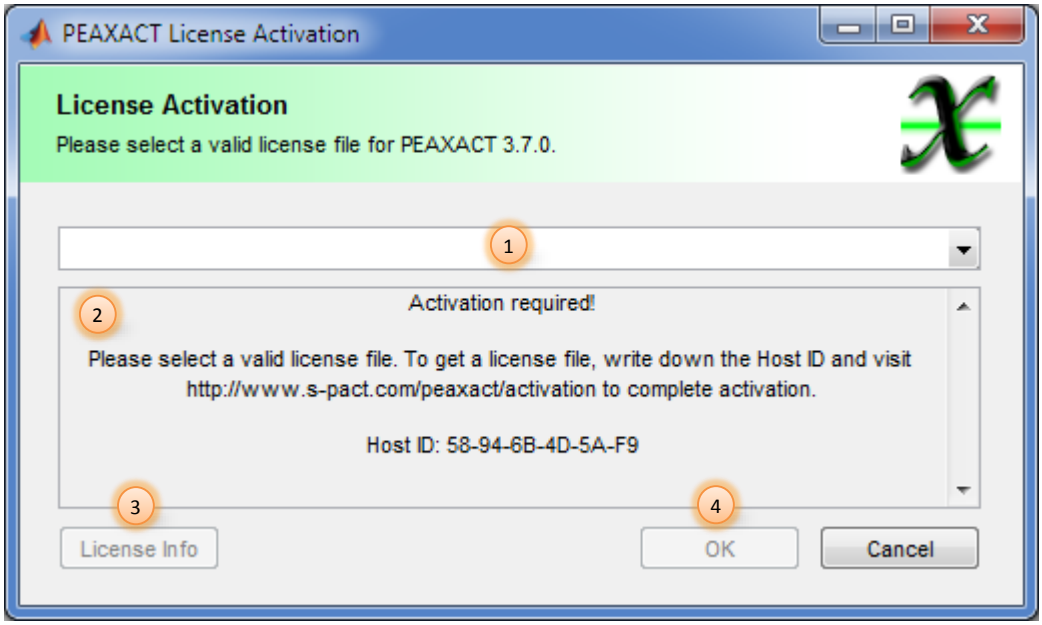
### Step 2: Request license file

- Visit http://www.s-pact.de/peaxact/activation and use the web form to request a license file.

### Step 3: Activate license

- Click the Windows start menu and select **Programs > PEAXACT > Activate PEAXACT**
- Wait until the License Activation Dialog is displayed

Note: The License Activation Dialog will also be shown if PEAXACT is started without a valid license.




**License Activation Dialog**

(1) License selection
(2) Status of activation
(3) Additional license information
(4) Apply and close

- Choose **Import License...** from the list (1) to browse for a valid license file. If the license is valid the license file is copied to the license directory. Alternatively, you can select **Free Viewer License** from the list (1), but this license would limit functionalities to reading and viewing model and data files only.
- Once a valid license is selected, you can click on the **License Info** button (3) to learn more about the license or on the **OK** button (4) to accept the selection.

**Per-machine license vs. per-user license**
If you perform the activation with administrator privileges, licenses will be activated per-machine, i.e. for all Windows users. Otherwise, licenses will be activated per-user, i.e. for the logged on user. Per-machine licenses take precedence over per-user licenses. Once a per-machine license is activated the License Activation Dialog gets locked for regular users.

# 1.5 Starting the PEAXACT Program

There are several ways to start the PEAXACT program on a Microsoft Windows platform:

- Click the Windows start menu and select **Programs** > **PEAXACT** > **PEAXACT 3**
- Double-click the **PEAXACT 3 icon** on the Windows desktop
- Open a command window, `cd` to the PEAXACT installation folder and type `peaxact` at the command prompt.

After starting PEAXACT, first the Status Window opens and then the Start Dialog opens. You can specify startup options to change the default startup behavior.

> Note: Starting PEAXACT may need several seconds (up to a minute) for loading the MATLAB Runtime Library.

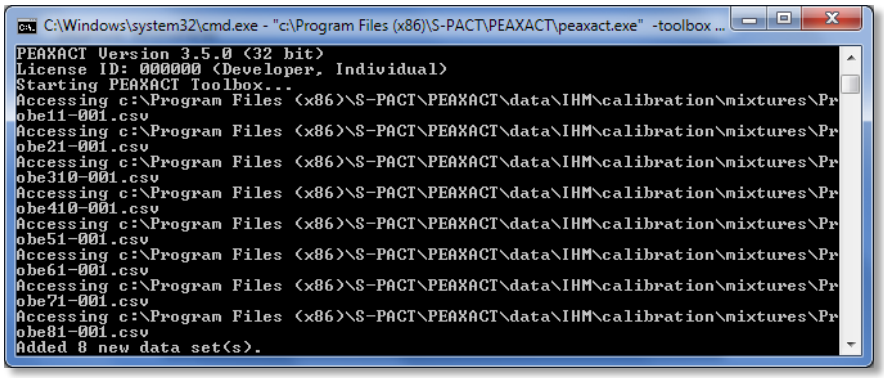
If you start PEAXACT without a valid license you will be prompted to activate a license.

If your computer is connected to the internet, a check for updates is performed at startup. You will be informed in case a new version is available.

If you have trouble starting PEAXACT, see Troubleshooting information.

## Status Window

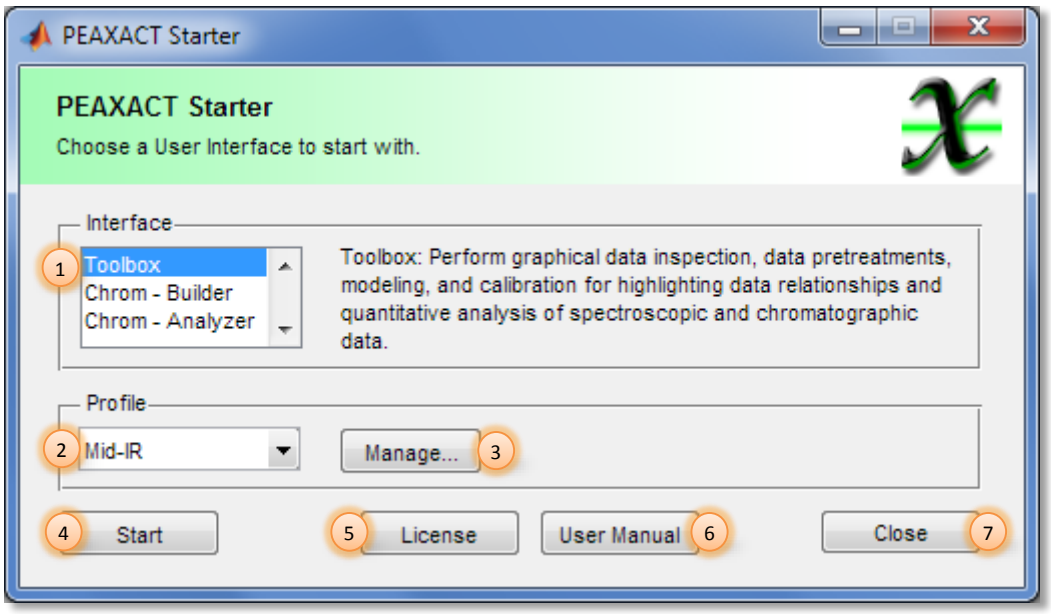
The Status Window is command window used for text output only, e.g. status and progress messages, warnings, and errors.




**Status Window**

## Start Dialog

The Start Dialog lets you choose a graphical user interface and a profile to start with.



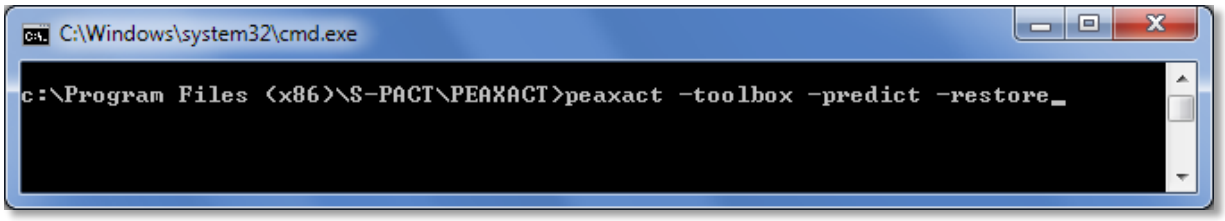

**Start Dialog**

(1) Select graphical user interface to start with (depends on license).

- Learn more about the Toolbox User Interface.
- Learn more about the Chrom User Interface.

(2) Choose user preferences profile
(3) Open User Preferences Editor
(4) Start with selected graphical user interface and user preferences profile
(5) Show License Activation Dialog
(6) Open this user manual (PDF file)
(7) Exit PEAXACT

## Startup Options




**Command Line Interface**

PEAXACT can be started from the command line with additional parameters. Please note that some command line options depend on licensed modules. Type `PEAXACT -help` to see all available options for your license.

Parameters in square brackets are optional. Angle brackets represent placeholders which must be replaced by specific values. The vertical bar `(|)` separates alternative arguments.

```
PEAXACT
PEAXACT -help | -manual
```

```
PEAXACT [-logLevel debug|exception|info|warning|severe|off]
        [-logFile <file> | -noFileLog] [-noConsoleLog]
        [-openglfix]
        [-<interface> [<more...>]]

-help         display help for command line usage
-manual       open user manual
-noConsoleLog disable text output to console
-noFileLog    disable text output to log file
-logLevel     set level of text output (default = info)
-logFile <file> set path and name of log file (default = APPDATADIR\peaxact.log)
-openglfix    fix graphics problems with some video cards
-<interface>  start specific user interface, e.g.:
  -toolbox    start PEAXACT Toolbox
  -chrom      start PEAXACT Chrom
-<interface> [<more...>]
```

There are more options available for each user interface. See Toolbox Command Line Options and Chrom Command Line Options.

# 2 INPUT AND OUTPUT FILES

PEAXACT reads and writes several files for different purposes. This Section explains these file types and their usage.

> Note: This manual uses the placeholder `APPDATADIR` in order to refer to the Windows user's directory for application data. In Windows XP it typically is
> `%UserProfile%\Local Settings\Application Data\S-PACT\PEAXACT`
> Since Windows Vista, it is
> `%LocalAppData%\S-PACT\PEAXACT`
> However, the actual directory can be seen in the User Preferences Editor or when typing `peaxact -help` at the command line.

## 2.1 Model File

The central element of PEAXACT is the ***model file***. Almost everything you do with PEAXACT is related to models. You use PEAXACT to create and edit models, load and save models, and apply models for data analysis.

The PEAXACT model is called ***master model***. Each model file stores one master model with all its sub-models and settings.

PEAXACT is backward compatible. It can read (but not write!) model files of previous versions. Please note that PEAXACT refuses reading model files saved with newer versions.

**File Extensions**

| | |
|---|---|
| `*.fpm` | PEAXACT Model file prior to version 3 (read-only) |
| `*.pxm` | PEAXACT Model file since version 3 |

## 2.2 Data File

A ***data file*** contains one or more samples, each of which consists of x-data (wavenumbers, frequencies ...) and y-data (intensities). PEAXACT supports various file formats and reads even 3D data (i.e. files containing multiple samples), and 4D data (i.e. files containing multiple frames of multiple samples). Many data files may also contain complementary information. However, this kind of data is ignored (but for a very few exceptions) because it is optional and varies too much between different file formats.

**Supported File Formats for Reading**

| | |
|---|---|
| `*.0` | OPUS files |
| `*.cdf` | Common Data Format files |
| `*.csv` | Text files (comma separated values) |
| `*.dpt` | Text files (blank separated values) |
| `*.txt` | Text files (delimiter separated values) |
| `*.dx *.jdx *.jcm` | JCAMP-DX files |

`*.mat` Matlab MAT files
`*.sp` Perkin Elmer Spectrum files
`*.spa *.srs` Thermo Nicolet OMNIC files
`*.spc` Galactic SPC file
`*.spe` WinSPEC files

**Text format**

Text formats like `CSV`, `DPT`, and `TXT` must contain tabularly structured data. The first table column is interpreted as x-data while following columns are interpreted as y-data. Table headers will be ignored.

**MATLAB format**

Data in MATLAB `MAT` files are recognized when the file contains one numeric matrix variable. The variable's name can be arbitrary. The first matrix column is interpreted as x-data while following columns are interpreted as y-data.

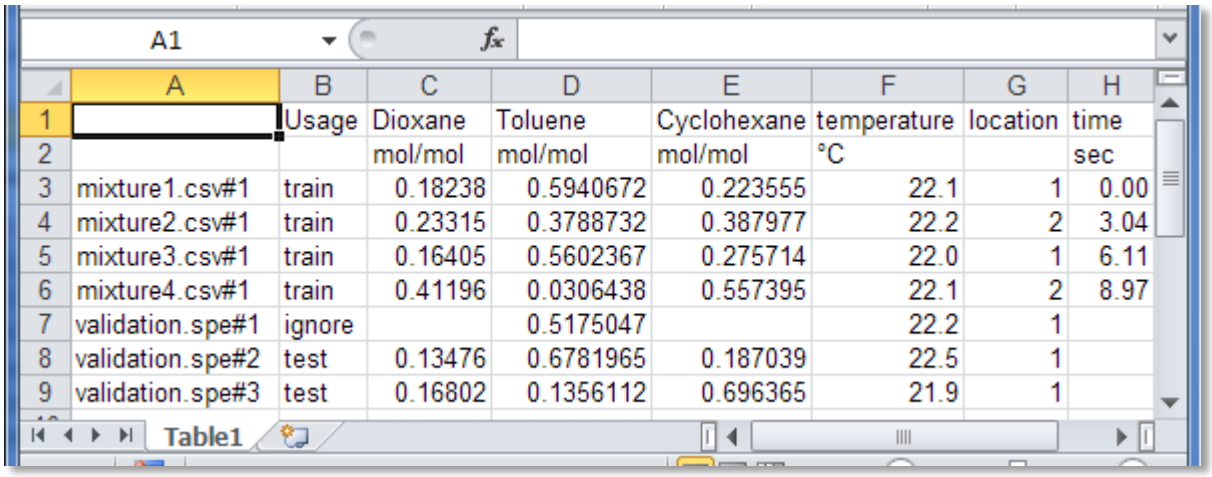
**Supported File Formats for Writing**

PEAXACT tries to avoid writing data files. However, there are a few operations that save data to file. You can choose between the following formats:

`*.mat` Matlab MAT files
`*.csv` Text files (comma separated values)

# 2.3 Data Table File

A ***data table*** file is a spreadsheet file (e.g. a Microsoft Excel file) that contains ***data sets***. Each row of the table represents a data set. Each column represents a data set ***feature***.



| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | Usage | Dioxane | Toluene | Cyclohexane | temperature | location | time |
| 2 | | | mol/mol | mol/mol | mol/mol | °C | | sec |
| 3 | mixture1.csv#1 | train | 0.18238 | 0.5940672 | 0.223555 | 22.1 | 1 | 0.00 |
| 4 | mixture2.csv#1 | train | 0.23315 | 0.3788732 | 0.387977 | 22.2 | 2 | 3.04 |
| 5 | mixture3.csv#1 | train | 0.16405 | 0.5602367 | 0.275714 | 22.0 | 1 | 6.11 |
| 6 | mixture4.csv#1 | train | 0.41196 | 0.0306438 | 0.557395 | 22.1 | 2 | 8.97 |
| 7 | validation.spe#1 | ignore | | 0.5175047 | | 22.2 | 1 | |
| 8 | validation.spe#2 | test | 0.13476 | 0.6781965 | 0.187039 | 22.5 | 1 | |
| 9 | validation.spe#3 | test | 0.16802 | 0.1356112 | 0.696365 | 21.9 | 1 | |

**Data Table**

The concept of data tables is an easy way to manage data that belongs together. See also Data Set Management for further information.

**File Extension**

`*.xls` Excel 97-2003-workbook
`*.csv` Text files with tabular data separated by a delimiter character

**Data Set**

A data set consists of at least a unique reference to a sample (see ***URI***), and optional *features*.

**URI**

A *uniform resource identifier (URI)* unambiguously identifies a specific sample in a data file. The URI is a name composed of the data file's name followed by the number sign (#) and an ***access ID*** (`URI = <dataFileName>#<accessID>`). A URI with missing access ID (and missing number sign) refers to all samples in the file.

**Access ID**

- For most 2D data files (single sample) and 3D data files (multiple samples), the access ID simply is the sample number within the file (e.g. `#3`).
- For 4D data files (multiple frames), the access ID is the frame number followed by the sample number within the frame (e.g. `#10-3`)
- For OPUS files, the access ID is the block name followed by the sample number within the block (e.g. `#AB-3`)

**Features**

Features are pieces of information that are related to samples. There are two kinds of features: ***predictive features***, e.g. concentrations, and ***non-predictive features***, e.g. timestamps. Non-predictive features are recognized by the following special names: `Usage`, `Timestamp`, `Time`, `Location`. Other feature names automatically indicate predictive features.

**Table Format**

PEAXACT is able to read data table files as long as the following formatting rules apply:

- First table column contains sample URIs (with absolute or relative paths)
- First table row contains feature names (except for first column)
- Rows without URI as well as columns without feature name are ignored
- URIs without access ID refer to all samples in a file
- Values of predictive features must be numeric
- Values of non-predictive features are treated in a special manner:
  - `Usage`: values can be `unspecified`, `train`, `test`, or `ignore`
  - `Timestamp`: date and/or time as text or serial number
  - `Time`: any numeric value, used to tag a temporal sequence
  - `Location`: any numeric value, used to tag a spatial sequence
- Empty cells are interpreted as missing values
- In case of duplicate URIs, feature values are merged (values from rows further down take precedence over values from rows further up)
- In case of duplicate feature names, only the first occurrence is considered
- For Excel files, only the first worksheet is considered

## 2.4 Session File

A ***session file*** is a snapshot of the PEAXACT workspace at the exact moment the session is saved. Loading a session file restores the workspace and allows you to continue working where you stopped before.

PEAXACT automatically saves session files to a specific directory on shutdown. There is a user preference you can set to change this behavior.

**File Extension**

`*.mat` PEAXACT Session file up to version 2.x (read-only)
`*.pxs` PEAXACT Session file since version 3

**Auto-Save Directory**

`APPDATADIR\sessions`

## 2.5 User Preferences File

A user preferences file contains a full set of PEAXACT options. It is also called a ***profile***. Profiles are located at a specific directory but can be imported from or exported to other directories. A profile is loaded on PEAXACT startup and saved each time individual options are changed by the User Preferences Editor.

PEAXACT defines some defaults profiles. These profiles can be loaded, edited and saved like any other profile. However, a defaults profile can always be reset to its initial state by simply deleting the file. The missing defaults profile is then restored during next PEAXACT startup.

**File Extension**

`*.profile` PEAXACT user preferences profile

**Directory**

`APPDATADIR\profiles`

**Defaults profiles**

```
Factory.profile
Chromatography.profile
Mid-IR.profile
NIR.profile
NMR.profile
Raman.profile
UV-Vis.profile
```

## 2.6 License File

The ***license file*** contains information about a license, e.g. the licensed release version. On startup, PEAXACT searches the Windows registry for a license filename, reads the file or asks for a valid license file if none could be found.

**File Extension**

`*.lic` PEAXACT License file

### License Filenames in the Windows Registry

`HKEY_LOCAL_MACHINE\S-PACT\PEAXACT\licenseSource`
takes precedence over:
`HKEY_CURRENT_USER\S-PACT\PEAXACT\licenseSource`

### License Directory

`%ProgrammData%\S-PACT\PEAXACT`

## 2.7 Report File

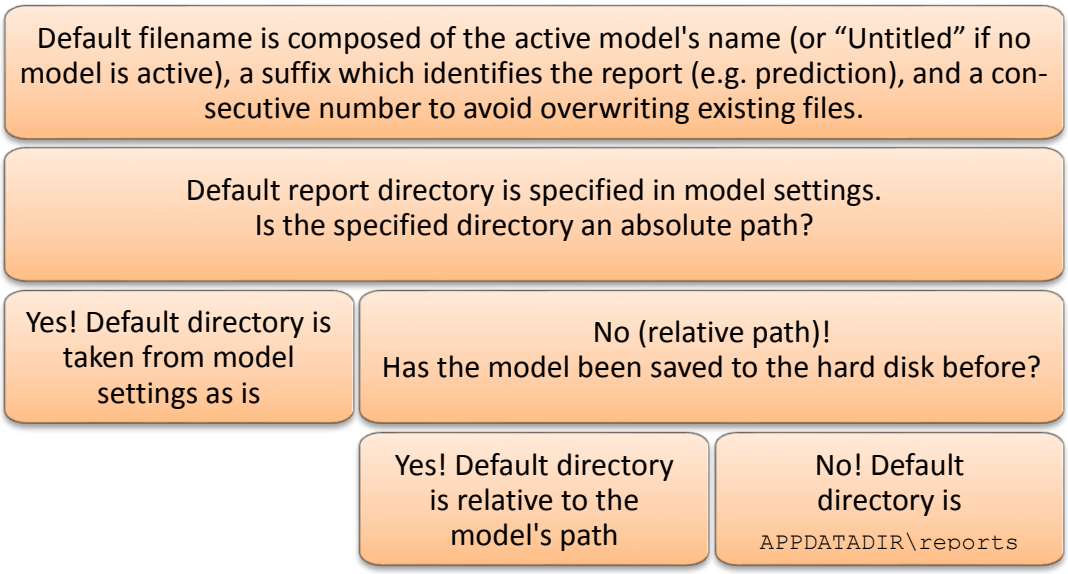
*Report files* are used to export graphical or tabular reports. Depending on the report, different file formats are available such as text, table, or image files.

### Supported Formats

| | |
|---|---|
| `.pdf` | Portable Document Format |
| `.xls` | Microsoft Excel Workbook |
| `.csv` | Comma separated values |
| `.dpt` | Tab separated values |
| `.txt` | Blank separated values |
| `.jpg` | JPEG Image File |
| `.png` | PNG Image File |
| `.tiff` | TIFF Image File |
| `.bmp` | Bitmap Image File |
| `.eps` | Encapsulated Post Script Image File |
| `.mat` | Matlab MAT File |
| `.fig` | Matlab Figure File |

### Default Directory and Filename

Writing reports typically involves a File Dialog which you can use to browse directories and choose a filename. A default filename is suggested according to the heuristic below. Default filenames are used automatically when PEAXACT is running in non-interactive mode.



# 2.8 Log File

By default, PEAXACT writes messages, warnings and errors to a ***log file***. The verbosity of the log file can be changed using the command line parameter `-logLevel`. File logging can be disabled completely using the command line parameter `-noFileLog`.

**Directory and Filename**

```
APPDATADIR\peaxact.log
```

# 3 TOOLBOX USER INTERFACE

## 3.1 Introduction

### 3.1.1 Product Overview

PEAXACT Toolbox is a graphical user interface for the quantitative analysis of spectroscopic data. Functionality covers all aspects of data-driven and model-based analysis, namely data visualization and inspection, modeling of data with the goal of highlighting useful information, and reporting of results. Using PEAXACT Toolbox, you get easy access to multiple analysis methods from a common platform.

Features include:

- 2D and 3D data visualization
- Data pretreatments such as resampling, baseline correction, standardization
- Interactive data modeling
- Data analysis with several analysis methods, e.g.
  - Peak search
  - Peak integration
  - Rank analysis based on principle components
  - Multivariate Curve Resolution (MCR-ALS)
  - Hard Modeling Factor Analysis (HMFA)
  - Peak fitting / Peak deconvolution
  - Component fitting
  - Indirect Hard Modeling (IHM)
  - Partial Least Squares (PLS)
- Customizable reporting
- Exporting to several output formats, e.g. PDF, XLS, CSV, JPG; PNG, EPS, FIG

### 3.1.2 Starting the Toolbox User Interface

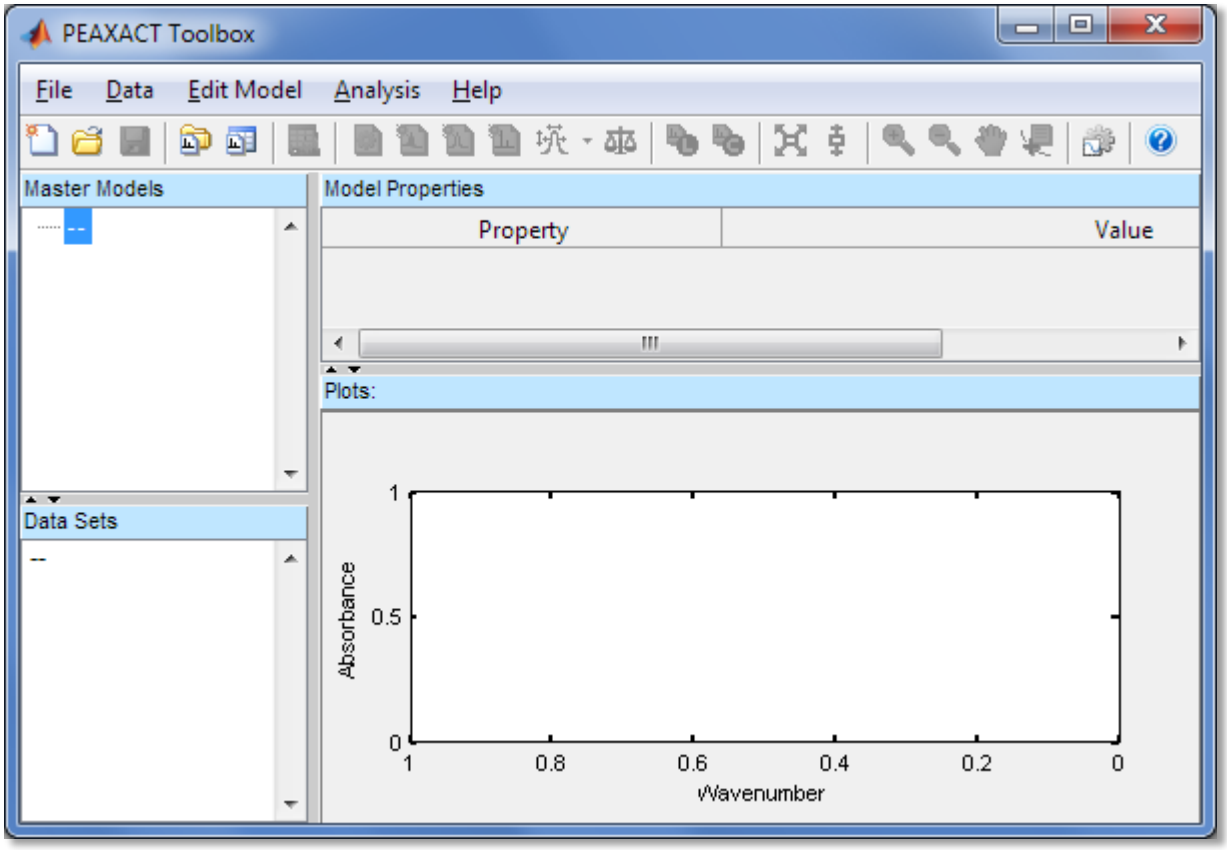
After starting the PEAXACT program choose the Toolbox user interface from the Start Dialog.

Alternatively, you can run PEAXACT with startup option `-toolbox` which directly starts the Toolbox user interface.

#### The Desktop

When you start the Toolbox user interface, the ***desktop*** appears. The desktop contains a set of control panels for managing models, data sets, and features associated with PEAXACT.

The following figure shows the default desktop. Learn more about the desktop in the next Section.




**PEAXACT Desktop**

## Command Line Options

In addition to PEAXACT startup options, the Toolbox user interface has individual command line options.

Parameters in square brackets are optional. Angle brackets represent placeholders which must be replaced by specific values. The vertical bar `(|)` separates alternative arguments.

```
PEAXACT -toolbox
PEAXACT -toolbox -help
PEAXACT -toolbox [-defaults <name> | -profile <name>]
                 [-predict] [-restore | -session <file>]
                 [<filename1> [<filename2> [...]]]

-help         display help for command line usage
-defaults <name> load defaults profile
-profile <name> load user profile
-predict      run prediction task without GUI
              requires calibrated model and data files to be loaded
-restore      reload most recently used session
-session <file> load session file
<filename*>   load additional files. These can be:
              model files, data files, or data table files
```

Additional notes:

- Option `-predict` must be used in combination with `-restore` or `-session <file>` or with `<filename*>` in order to load models and data sets for prediction analysis
- Starting prediction analysis from the command line has advantages over the graphical user interface. Learn more in Section Prediction Analysis.
- Option `-defaults <name>` overwrites any existing user preferences profile `name.profile` with default settings
- Filenames must be quoted if they contain spaces, e.g. `"c:\my data\file A.pxm"`

### 3.1.3 Quitting the PEAXACT Program

To quit the PEAXACT program, click the close box in the desktop window or choose "Quit" from the desktop File menu. You may need to cancel running operations first in order to do this.

PEAXACT closes after prompting you to save any unsaved files.

#### Enforced shutdown

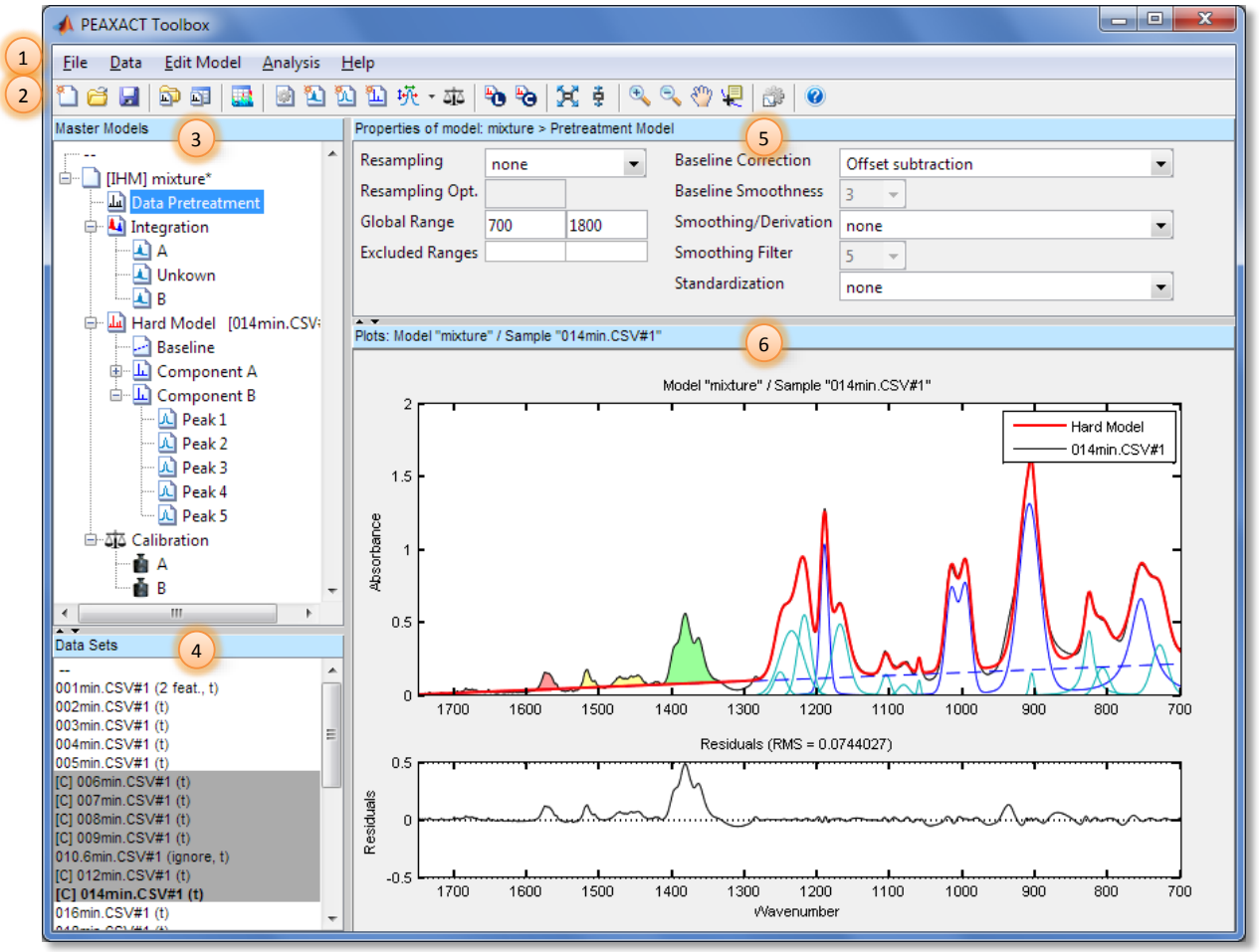
To force quitting the PEAXACT program at any time, click the close box in the Status Window.

> Caution: When closing the Status Window you will not be prompted to confirm quitting and all unsaved changes will be discarded.

# 3.2 Desktop

## About the Desktop

When you start the PEAXACT Toolbox user interface, it displays the ***desktop***. The desktop is a set of control panels for managing models, data sets, and features associated with PEAXACT.

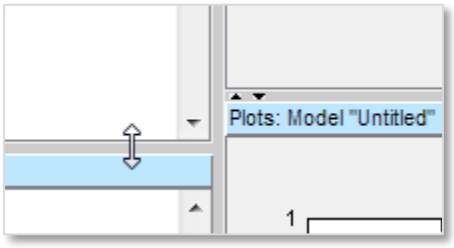



**PEAXACT Desktop with opened models and loaded data sets**

(1) Menu bar
(2) Toolbar
(3) Model Tree Panel
(4) Data Sets Panel
(5) Model Properties Panel
(6) Plot Panel

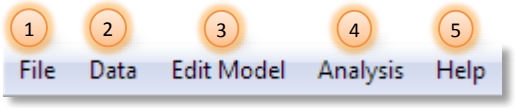
## Resizing Panels

The desktop is structured into four panels. You can change the panels' sizes to suit your needs. Drag the gray bar between two panels to resize them (see Figure below). Also, you can hide panels completely by clicking on the small black triangle on the resize bar.




## 3.2.1 Menu bar

The ***Menu bar*** groups available actions into 5 groups:




**Desktop Menu bar**

(1) File: Model File Management, User Preferences, Sessions
(2) Data: Data Set Management, Data Inspector
(3) Edit Model: Modeling Tools
(4) Analysis: Analysis & Reporting
(5) Help: Documentation, Support, Licensing, Web Resources...

## 3.2.2 Toolbar

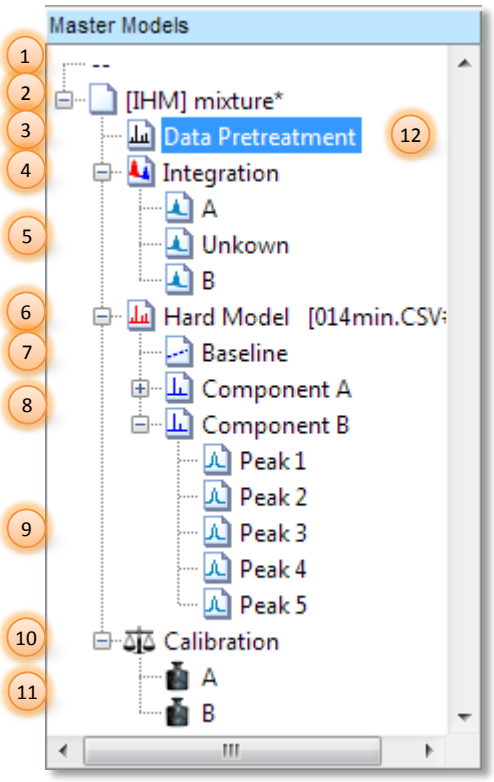
The ***Toolbar*** gives quick access to frequently used actions:


**Desktop Toolbar**

(1) New Model, Open Model, Save Model
(2) Load Data Files, Load Data Table Files
(3) Open Data Inspector
(4) Open Model Settings Editor, Add Integration Range, Auto-fit Peaks
Import Components, Adjust Hard Model Parameters, Calibration
(5) Linear Component Fit, Component Fit according to model settings
(6) Full zoom, Full y-zoom
(7) Zoom in, Zoom out, Move center, Data cursor
(8) Show User Preferences Editor
(9) Show User Manual

## 3.2.3 Model Tree Panel

The ***Model Tree Panel*** (located top left of the desktop) lists all opened master models, each one represented as tree structure. Use the tree to access certain model elements.




**Model Tree Panel**

(1) Empty item
(2) Master Model item, contains the following information

- Calibration type (if any): `[IHM]`, `[PLS]` or `[PI]`
- Model name (default name is `Untitled`)
- Asterisk symbol (`*`) when model has unsaved changes

(3) Pretreatment Model item
(4) Integration Model item
(5) Integration Range item, contains range name
(6) Hard Model item, optionally contains the name of the sample (in brackets) the model was last fitted to
(7) Baseline Model item
(8) Component Model item, contains component name
(9) Peak Model item
(10) Calibration Model item
(11) Calibrated Feature item
(12) Selected/active element

## Selecting and Activating Model Elements

Left-clicking an element of the tree selects it. The selected element becomes the active element. Only one element can be selected at a time.

Activating model elements affects other parts of the GUI:

- Display of the element's parameters (see Model Properties Panel)
- Display and/or highlighting of the element's plot (see Plot Panel)
- Enabling of element-specific actions in the Menu bar and Toolbar

By selecting a Master Model item (2) or any of its subentries (3)-(11), the model becomes the ***active model***. Many operations of PEAXACT implicitly refer to the active model.

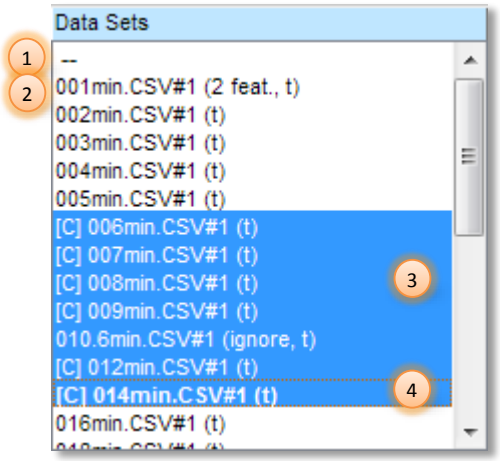
Sometimes it is favorable to not have any model active. Select the empty item (1) for this purpose.

### Context Menu

Right-clicking an item selects it and opens an element-specific context menu.

## 3.2.4 Data Sets Panel

The ***Data Sets Panel*** (located bottom left of the desktop) lists all data sets that are currently loaded. Use the panel to access certain data sets.




**Data Sets Panel**

(1) Empty item
(2) Data Set item

The name of each data set is composed of different kinds of information:

- `[C]` marks data sets for which a component fit is available
- Sample name and access ID

The name also informs about data set features:

- `train`, `test`, `ignore`: data set usage
- `# feat.`: number of available predictive features
- `loc.`: measurement location available
- `t`: measurement time or timestamp available

(3) Selected elements
(4) Active data set

### Selecting Data Sets

Many features of PEAXACT – including all analysis methods – are applied to selected data sets. Left-clicking and holding the mouse button selects one or more data sets. Doing this while the CTRL-key is pressed allows you to select multiple ranges. Press CTRL+A in order to select all items.

### Activating a Data Set

The last selected data set becomes the ***active data set***. Its name is displayed in bold characters. Some operations of PEAXACT implicitly refer to the active data set.

When activating a data set whose URI points to a file that doesn't exist, you will be prompted to update the data set. See Update Missing Files.

Activating a data set affects other parts of the GUI:

- Display of the sample's plot (see Plot Panel)
- Enabling of specific actions in the Menu bar and Toolbar

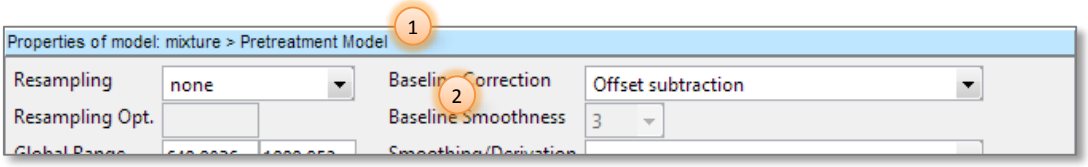
Sometimes it is favorable to not have any data set active. Select the empty item (1) for this purpose.

### Context Menu

Right-clicking into the data sets panel opens the context menu.

## 3.2.5 Model Properties Panel

The ***Model Properties Panel*** (located top right of the desktop) changes its appearance according to the selected model element in the Model Tree Panel. The panel displays parameters of the selected model element. Some parameters are displayed for information purposes only, others are editable. Use the controls (e.g. edit fields, pull-down menus, tables) to change parameter values.




**Model Properties Panel**

(1) Name of currently selected model element
(2) Controls for displaying and editing model parameters

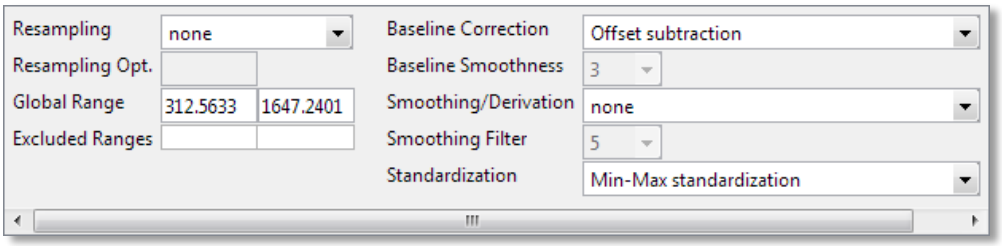
### Master Model Properties

When the Master Model item is selected in the Model Tree Panel, the Model Properties Panel displays master model properties as a tabular report. The table is read-only.

| Property | Value |
|---|---|
| GENERAL PROPERTIES | |
| Report time | 14-May-2013 22:39:50 |
| Report creator | PEAXACT 3.1.3 |
| Model file | Untitled.pxm |

## Pretreatment Model Properties

When the Data Pretreatment item is selected in the Model Tree Panel, the Model Properties Panel displays data pretreatment settings.




## Integration Model and Integration Range Properties

When the Integration Model item is selected in the Model Tree Panel, the Model Properties Panel displays a table containing the parameters of all integration ranges. When an Integration Range item is selected, the table contains a single row.

| Range Name | Lower bound | Upper bound | Baseline | Area |
|---|---|---|---|---|
| Toluene | 512.975 | 536.785 | Offset | 23547.2 |
| Dioxane | 478.452 | 504.94 | Straight Line | 4663.44 |
| Cyclohexane | 1247.11 | 1272.72 | Linear Fit | 97553.3 |

## Hard Model Properties

When the Hard Model item is selected in the Model Tree Panel, the Model Properties Panel displays a table containing the parameters of all component models.

| Component | Shift | Weight |
|---|---|---|
| empty | ☑ | 1 |
| Cyclohexane | ☑ | 1 |
| Dioxane | ☑ | 1 |
| Toluene | ☑ | 1 |

## Baseline Model Properties

When the Baseline Model item is selected in the Model Tree Panel, the Model Properties Panel displays a table containing the parameters of the hard model's baseline.

| | Value | Fixed? |
|---|---|---|
| Intercept | 0 | ☐ |
| Slope | 0 | ☐ |

### Component Model Properties

When the Component Model item is selected in the Model Tree Panel, the Model Properties Panel displays a table containing an overview of the component's peaks.

| Peak | Type | Position | Area | Group |
|---|---|---|---|---|
| Peak 1 | pseudovoigt | 798.665 | 8.91831 | 0 |
| Peak 2 | pseudovoigt | 1023.37 | 6.37353 | 0 |
| Peak 3 | pseudovoigt | 1259.57 | 5.46854 | 0 |

### Peak Model Properties

When the Peak Model item is selected in the Model Tree Panel, the Model Properties Panel displays a table containing the parameters of the selected peak.

| | Value | Fixed? |
|---|---|---|
| Position | 798.665 | ☐ |
| Max. | 0.774906 | ☐ |
| HWHM | 4.40046 | ☐ |
| Gaussian Part | 0.519634 | ☐ |

☐ Distinctive Peak (HMFA)

### Calibration Model Properties

When the Calibration Model item is selected in the Model Tree Panel, the Model Properties Panel displays a table containing all calibrated features and predicted values for the active data set.

| Feature | Prediction |
|---|---|
| Cyclohexane | 0.71142 |
| Dioxane | 0.155877 |
| Toluene | 0.132703 |

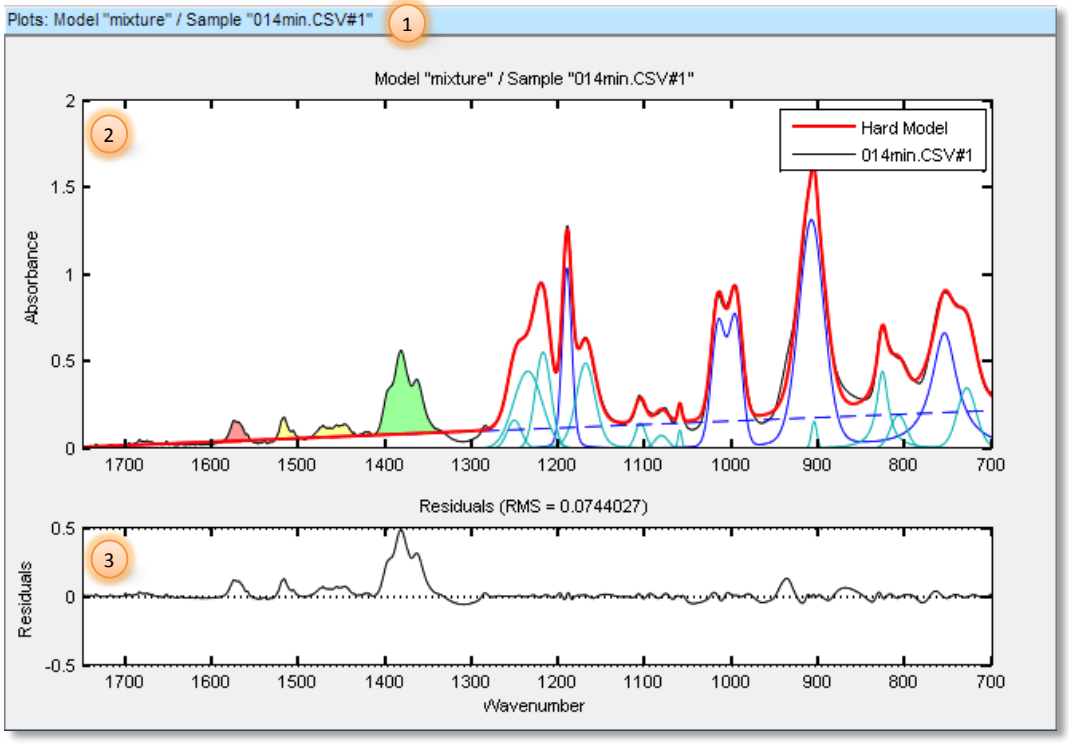
### Calibrated Feature Properties

When the Calibrated Feature item is selected in the Model Tree Panel, the Model Properties Panel displays calibration properties as a tabular report. The table is read-only.

| Property | Value |
|---|---|
| CALIBRATION: COMPONENT SUMMARY | |
| Feature name | Cyclohexane |
| Range | [0 1] |
| Regression function | simple |
| $R^2$ | 0.999319 |

## 3.2.6 Plot Panel

The ***Plot Panel*** (located bottom right of the desktop) changes its appearance according to the items selected in the Model Tree Panel and in the Data Sets Panel. The Plot Panel is used for graphical output. It also offers a way to interactively edit the model by using the mouse for modifying plots.



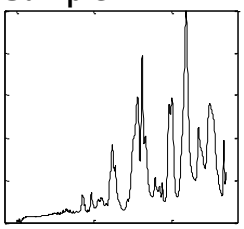

**Plot Panel**

(1) Name of active model and/or active data set
(2) Model and/or sample plot
(3) Residual plot

The residual plot is only visible if both, a master model and a data set are active, and in addition, if the master model contains a hard model.

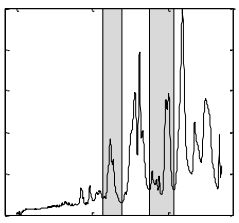
## Available Graphs

### Sample



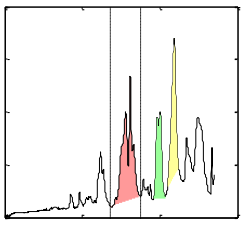
Graph of xy-data of the active data set's sample (black line).
If a model is active in the Model Tree Panel, the sample is modified according to the active model's pretreatment model.
The sample plot cannot be modified interactively.
(Hint: Use the Data Inspector to plot multiple samples)

### Excluded Ranges



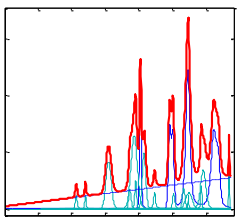
Graph of excluded ranges of the active model's data pretreatment model (transparent gray patches).
Excluded range patches can be modified with the mouse (in modeling mode if enabled in the user preferences).

**Integration Ranges**

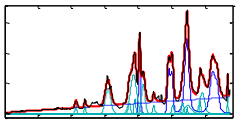


Graph of integration ranges of the active model's integration model (colored patches).
Integration range patches can be modified with the mouse (in modeling mode).

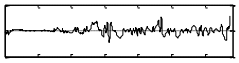
**Hard Model and Residuals**



Graph of baseline model (dashed blue line), component models (solid blue lines), and peak models (cyan lines) of the active model's hard model (red line).
Baseline, component, and peak plots can be modified with the mouse (in modeling mode). The red line cannot be modified; it is the sum of the other plots.



If a data set is active in the Data Sets Panel, the sample's xy-data is plotted, too, and the differences (residuals) between hard model (red line) and sample (black line) are shown in a second graph.



### Interactive Modeling

Most model-related lines and patches can be modified interactively with the mouse. Available actions differ from graph to graph but typically include:

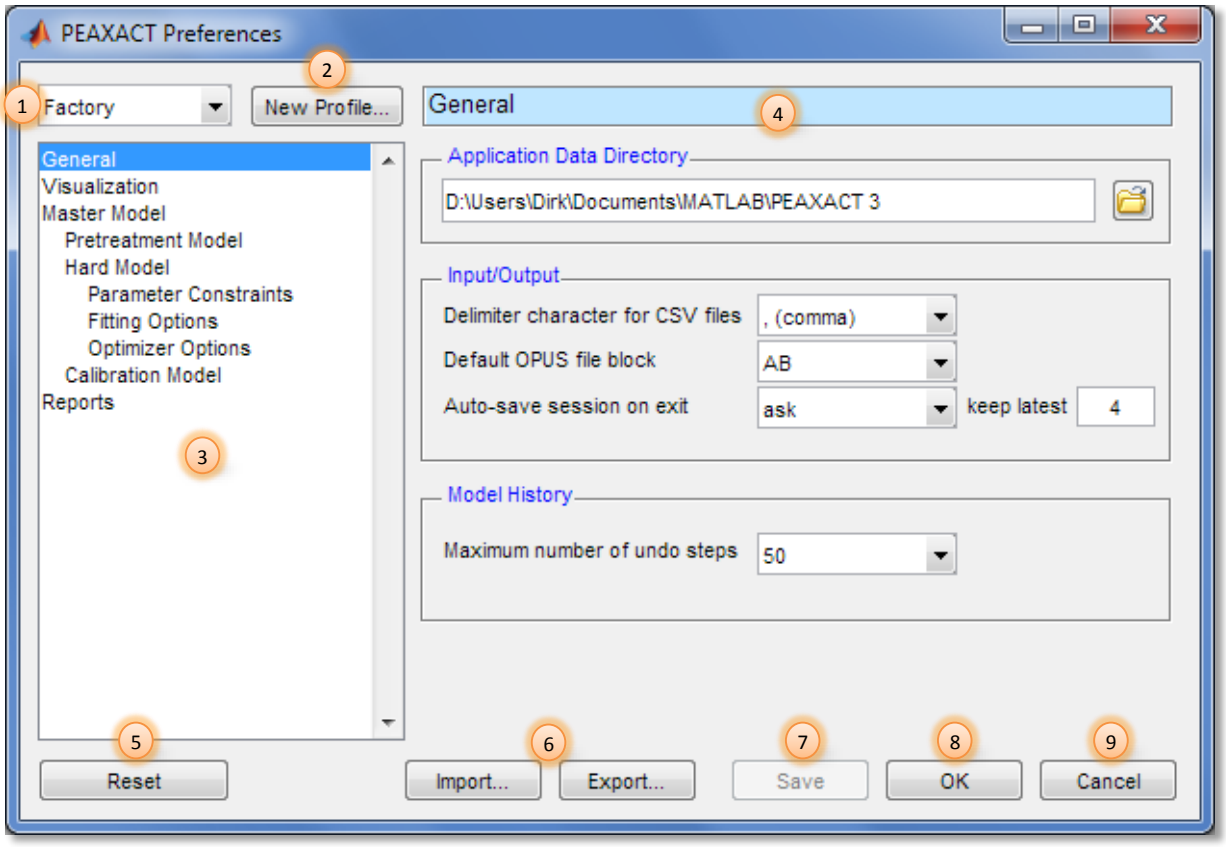
- left-clicking a graph in order to select the corresponding element in the Model Tree Panel
- dragging a graph while the left mouse-button is pressed in order to change the graphs' position or shape
- right-clicking a graph in order to open the corresponding context-menu
- hitting the DEL-key in order to delete the corresponding element from the model

Detailed descriptions can be found in Section Modeling.

## 3.2.7 User Preferences Editor

| Menu bar: | File > Preferences |
|---|---|

The ***User Preferences Editor*** can be used to adjust PEAXACT settings.




**User Preferences Editor**

(1) Select profile
(2) Create new profile
(3) Select category
(4) Preferences of selected category
(5) Reset changes of selected profile
(6) Import / Export profile file
(7) Save changes to file
(8) Activate profile and close dialog
(9) Discard changes and close dialog

## Profile Management

User preferences are preferred settings. A full set of preferences is called a user preferences profile (or simply profile). The user may have several profiles for different purposes but only one profile can be active at a time. The User Preferences Editor allows managing multiple profiles and choosing the active profile.

Please note that the User Preferences Editor looks similar to the Model Settings Editor. This is because model settings originate from user preferences, i.e., every time a new model is created, the preferred settings for new models are copied from the active profile to the individual model and may then later be changed independently using the Model Settings Editor.

## Selecting, Editing, and Activating a Profile

Use the pull-down menu (1) to select a profile and to load the preferences into the editor. Choose any category from the list (3). The right pane (4) shows settings for that category. For a description of available settings, see Section 3.2.10.2.

Change preferences using the controls in the right pane. Click the Reset button (6) to undo latest changes. Click the OK button (8) to save the preferences, activate the profile, and close the editor. Active preferences take effect immediately and remain persistent across sessions of PEAXACT.

If you select another profile while the current one has changes, you will be asked to save the changes first.

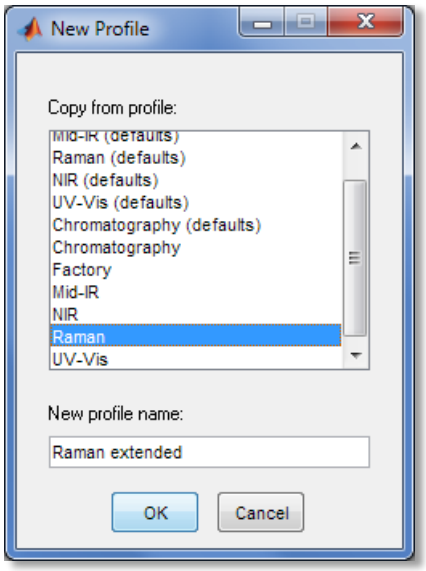
## Save, Import, Export

Click the Save button (7) in order to save changes to file. PEAXACT saves all profiles into a default directory. In order to export profiles to other directories, or to import profiles from an external directory to the default profile directory, use the Export or Import buttons (6).

Please note that saving a profile does not automatically activate it. You have to click OK (8) in order to activate a profile.

## New Profile

Click the New Profile button (2). In the next dialog (see Figure) you can select an existing profile (or pre-defined defaults) as a source profile to copy from. Enter a new profile name and confirm with OK. The new profile will be selected (but not activated) and can be further edited. Make sure to save changes even if you do not want to activate the profile.




**New Profile Dialog**

### Default Profiles

PEAXACT is delivered with pre-defined profiles

- Factory: a very general set of preferences
- Chromatography, Mid-IR, NIR, NMR, Raman, UV-Vis:
  adjusted preferences for different measurement techniques

Although these profiles can be edited like any other profile, they are special profiles which can always be reset to their initial settings by either deleting the profile or starting PEAXACT with options `-toolbox -defaults`. See user preferences file and command line options for additional help.

## 3.2.8 Working in Sessions

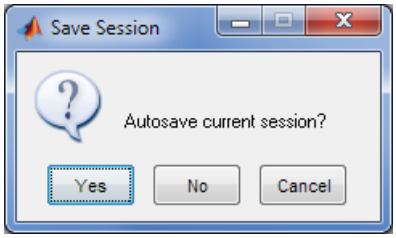
PEAXACT can save a snapshot of the current workspace into a session file. This is helpful for splitting up your work into different sessions.

### 3.2.8.1 New Session

| | |
|---|---|
| Menu bar: | File > New Session... |

Starting a new session clears the current workspace and resets the desktop to its initial state: all windows except the main window are closed, all model files are closed, and all data sets are unloaded.

You will be asked to automatically save the current session first.




### 3.2.8.2 Open Session

| | |
|---|---|
| Menu bar: | File > Open Session > Load Session File... |
| Menu bar: | File > Open Session > ... |

You can choose between loading a session from file (by using the File Dialog) and reloading a recently opened session.

When opening a session file, the workspace of the current session is replaced and all panels and controls are refreshed to restore the state of the opened session.

Caution: You will not be prompted to save the current session.

## Drag and Drop

A fast way to open a session file is by using drag & drop. Drag a single session file from the Windows Explorer (e.g.) and drop it into the Model Tree Panel or the Data Sets Panel. Note: if you drag & drop multiple files, PEAXACT ignores any session files among them and tries to open model files or data files instead.

## Command Line Alternatives

```
peaxact -toolbox -session <file>
peaxact -toolbox -restore
```

When starting PEAXACT from the command line, you can pass additional parameters to either open a certain session file or to restore the most recently saved session.

## What is Restored from Session Files?

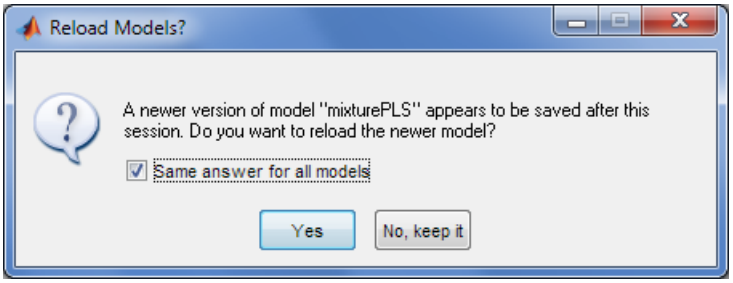
- *Master models* are restored exactly to the state when the session was saved. If at that time any model had unsaved changes, it will be marked unsaved again after loading the session. *Modeling history* is not restored, i.e. modeling steps from before the session was saved cannot be reverted using Undo/Redo.
- *Data sets* are restored exactly to the state when the session was saved, including filenames of data files. When you move or rename these files on your hard disk, filenames stored in the session still point to the old location and need to be updated after the session is restored (see below).
- All *component fits* of hard models to data sets are restored.
- *Selected and active elements* of the Model Tree Panel and Data Sets Panel are restored. Therefore, the Model Properties Panel and Plot Panel also are restored to their previous state.
- The *active user preferences profile* is not restored.

## Caution with Models Restored from Sessions!

A model restored from a session file must be treated with care, especially if

- multiple sessions exist containing the model at different building stages
- the model was modified and saved to file after the session was saved

In these cases it is possible to restore a model from a session file which in the meantime has been replaced by a newer version. You can choose whether to reload the newer model or continue working with the one contained in the session file.




### Update Missing Data Files

Data sets stored in a session file contain references (URIs) to data files from the moment the session was saved. When you move or rename these data files on your hard disk filenames stored in the session still point to the old locations. After reloading a session you may have to update these file references. PEAXACT informs you to update missing files when it tries to load data from such a file.

## 3.2.8.3 Save Session

| | |
|---|---|
| Menu bar: | File > Save Session As... |

Use the File Dialog to browse directories and select a filename for the session to be saved.

When saving a session file, the current workspace is written to the hard disk.

### Auto-Saving

The current session can be saved automatically to a special directory on PEAXACT shutdown. This behavior is controlled by a user preference.

### What is Stored in Session Files?

- All opened *master models* including changes made during runtime are saved to the session file. Please note that unsaved model changes are *not* saved to model files. Therefore, changed models will still be marked unsaved when the session is restored from file.
- Loaded *data sets*
- *Component fits*
- *Selected and active elements* of the Model Tree Panel and Data Sets Panel

# 3.2.9 License Management

| | |
|---|---|
| Menu bar: | Help > Licensing... |
| Menu bar: | Help > About PEAXACT |

From the Help menu you can open the License Activation Dialog and the License Information Window.

### License Activation

The License Activation Dialog can be used to change the current license file. The new license must be valid for the PEAXACT Toolbox.

### License Information

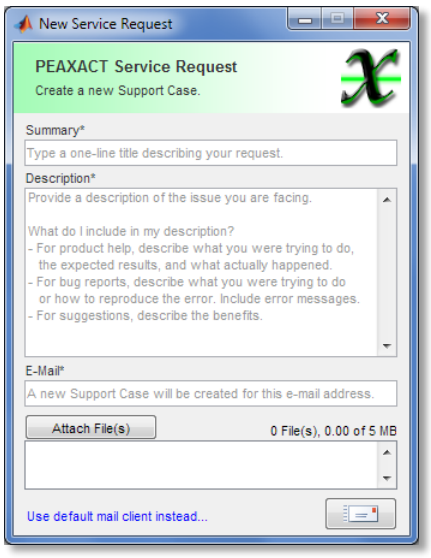
The License Information Window provides information about the activated license, e.g. which software modules you are allowed to use, for how long, and in which version.

# 3.2.10 Technical Support

## 3.2.10.1 Request Support

| Menu bar: | Help > Request Support... |
|---|---|

You can use the Service Request Dialog to send e-mails to Technical Support directly from within PEAXACT.




**Service Request Dialog**

You may want to make service requests for different purposes, e.g.:

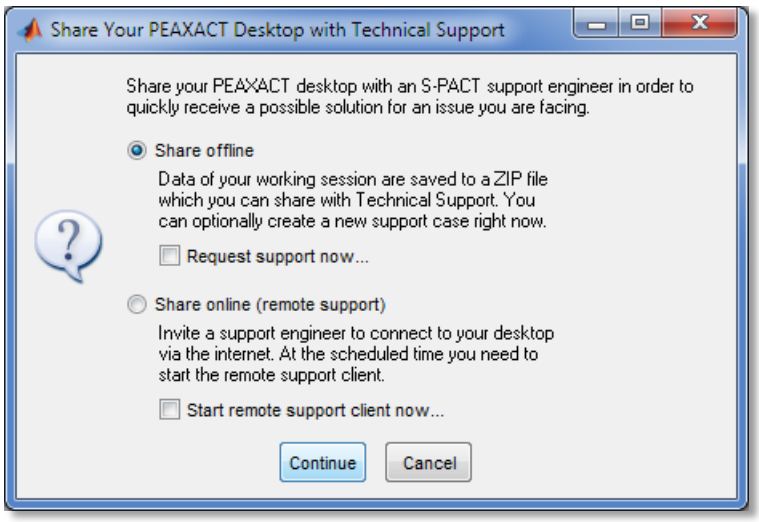
- asking for help
- reporting a bug / crash
- suggesting new features
- giving any kind of feedback

In each case, provide a subject and description of the issue you are facing with. You can optionally attach related files, e.g. a screen shot or a session file. When clicking the **Send** button, an e-mail is sent to support@s-pact.de, creating a new support case. If your computer is currently offline or uses protective methods to restrict internet access, you can click on **Use default mail client instead** to have the mail sent by your default e-mail application.

### 3.2.10.2 Share Desktop with Technical Support

| Menu bar: | Help > Share PEAXACT Desktop... |
|---|---|

You can share your PEAXACT desktop with an S•PACT support engineer in order to get a possible solution for an issue you are facing.




#### Share offline

If you choose this option and click **Continue**, data of your working session are saved as a ZIP file which you can send to Technical Support. Optionally tick the **Request support now...** checkbox to open the Service Request Dialog and create a new support case.

> Note: Technical Support can provide you with a secure upload link for securely transferring files up to 2 GB.

#### Share online (Remote Support)

If you choose this option and click **Continue**, the Service Request Dialog opens which can be used to invite a support engineer to connect to your desktop via the internet. At the scheduled time you need to start the remote support client.

The remote support client can be downloaded from www.s-pact.de/support/remote. Alternatively, you can tick the **Start remote support client now...** checkbox to download and start the client automatically when clicking **Continue**.


**Remote Support Client**

# 3.3 Settings

This Section explains PEAXACT settings and options. Settings can be found in two shapes. Firstly, as user preferences, i.e. as preferred or default options, and secondly, as model-specific settings, i.e. as customized options attached to individual model files.

Settings are grouped into the following categories, each of which is covered by a separate sub-section:

- General settings: basic options to miscellaneous topics
- Visualization settings: options controlling the appearance of the desktop
- Model settings: options related to models
- Report settings: options controlling the appearance of graphical and tabular reports

## 3.3.1 General Settings

### Application Data Directory

PEAXACT uses a special application data directory to save files during runtime. The directory cannot be changed, but for your convenience it is displayed in a text field of the User Preferences Editor and can be opened in the Windows explorer by clicking the button next to the field.

### Input / Output

**Delimiter character for CSV files**

Character to be used as delimiter for comma separated value (CSV) files.

Hint: You probably want to stay with the original comma but if you would like to open CSV files in Microsoft Excel, try switching the delimiter to semicolon if Excel does not recognize the comma as a delimiter.

**Default OPUS file block**

OPUS files may contain several different data blocks. Whenever PEAXACT opens an OPUS file, it only loads the default block.

Note: If an OPUS file to be loaded does not contain the default block, all blocks are loaded instead.

**Auto-save session on exit**

Auto-saving behavior of sessions during shutdown

| Value | Description |
|---|---|
| Never | Current session is never saved automatically |
| Ask | The user is always asked before auto-saving the current session |
| Always | Current session is automatically saved on shutdown |

**Keep latest N**

N specifies the number of auto-saved session files to keep on the hard disk.

### Model History

**Maximum number of undo steps**

The number specifies how many of the most recent modeling steps can be undone.

Note: You can disable the model history by setting the number to zero. This improves speed of some modeling actions, especially on slower computers.

## 3.3.2 Visualization Settings

**Reverse x-axis direction**

| Value | Description |
|---|---|
| Enabled | The x-axis in the Plot Panel is increasing from right to left. This is typical for IR, Raman, and NMR spectra. |
| Disabled | The x-axis is increasing from left to right. This is normal for chromatograms. |

**Ignore excluded ranges for auto-scaling of y-axis**

| Value | Description |
|---|---|
| Enabled | Y-axis limits are adjusted automatically to fit only the relevant measured signal, i.e. any signal that is not excluded. |
| Disabled | The y-axis is scaled automatically to fit the full measured signal. |

**Allow mouse modification of excluded ranges**

Excluded ranges are displayed as gray patches in the Plot Panel. These patches typically overlap other plots such as integration range plots or peak plots.

This option specifies the behavior when clicking on an excluded range patch.

| Value | Description |
|---|---|
| Enabled | Excluded range patches could be modified interactively using the mouse. In this case it may be difficult to interactively modify other plots which are overlapped by patches. |
| Disabled | Excluded range patches cannot be modified interactively. Instead, it is easier to click and modify plots which are overlapped by patches. |

**Show hard model baseline as separate plot**

| Value | Description |
|---|---|
| Enabled (recommended) | A separate plot for the baseline model is displayed in the Plot Panel. The plot can be used to interactively modify the baseline. |
| Disabled | The baseline model is still calculated but it is not displayed as a plot. Instead, the baseline model is subtracted from the active sample and the corrected sample is displayed in the Plot Panel. This needs to get used to but has the advantage of the sample signal being shifted towards zero, which sometimes may be favorable for modeling. |

## 3.3.3 Model Settings

Note: All settings described in this Section are both, user preferences and model-specific settings, depending on whether you open the User Preferences Editor or the Model Settings Editor.

In the User Preferences Editor, model settings are treated as *New Model Defaults*, i.e. they are copied as initial values to newly created models. Afterwards, the Model Settings Editor can be used to customize settings for each model individually. As for this Section, it makes no difference whether settings are considered as user preferences or model-specific settings.

### 3.3.3.1 Miscellaneous

**Model Report Directory**

The suggested default directory for saving reports (e.g. analysis results). The directory can be specified as absolute path (starting with a drive letter or double slash) or relative path. The latter will be resolved to a path relative to the model's directory. For more details see Section Report File.

### 3.3.3.2 Pretreatment Model

Some pretreatments modify x-data (wavenumber, frequency) while others alter y-data (intensities). Data Pretreatments are applied before any analysis in the following order:

- Resampling
- Global range
- Baseline correction
- Smoothing / Derivatives
- Standardization
- Excluded ranges

**Resampling**

Affects x-data

| Value | Description |
|---|---|
| None | Disabled |
| Thinning | Data thinning reduces the number of data points. A *thinning factor* of N keeps every N-th data point. The thinning factor can be any decimal number greater than or equal to 1. |
| Equidistant points | Resamples data at a specified number of new x-values. A *number of points* of N recalculates y-values at N equally spaced x-values within the visible range by linear interpolation. |

| Value | Description |
|---|---|
| Reference x-axis | Resamples data at new x-values provided by a specified *reference file*. Y-values are re-calculated at the same x-values of the reference sample by linear interpolation. Reference x-values are permanently stored in the model. |

**Global range**

Affects x-data

The global range defines the smallest and largest x-value. Only signal in between is considered for analysis. This is useful for cutting off noisy or unwanted signal at the edges.

**Excluded ranges**

Affects x-data

Excluded ranges are defined by a lower and upper x-value. Signal in between is not considered for analysis. This is useful for cutting off local signal artifacts.

**Baseline correction**

Affects y-data

Baseline correction subtracts a background signal from measured data.

| Value | Description |
|---|---|
| None | Disabled |
| Offset subtraction | Subtracts a constant intensity value such that the non-excluded data is shifted to zero |
| Straight line subtraction | Subtracts a straight line that is drawn between the two y-data points at the range bounds |
| Linear fit subtraction | Subtracts the straight line that lies closest to the sample's y-data, but always below it |
| Advanced baseline correction (experimental) | Tries to identify parts of the data that is not peaks and subtracts it. See also *baseline smoothness* |

**Baseline smoothness**

Affects y-data

The baseline smoothness value only applies when baseline correction is set to "advanced". The greater the value, the smoother the detected baseline.

**Smoothing / Derivatives**

Affects y-data

Smoothing and derivation are operations to emphasize the information content of the signal. While smoothing is meant to remove disturbing noise, derivation separates overlapping peaks into distinct signals. See also *smoothing filter*.

| Value | Description |
|---|---|
| None | Disabled |
| Smoothing only | Removal of noise by smoothing the signal |
| 1st order derivative | Numerical first derivation. Inflection points in the original signal are resolved to peaks |
| 2nd order derivative | Numerical second derivation. Shoulders of overlapping peaks in the original signal are resolved to peaks |

**Smoothing filter**

Affects y-data

A smoothing filter is required for smoothing and derivation. The smoothing filter N defines the size of a moving window. The data point in the window's center gets smoothed by taking its neighboring points into account. The greater N becomes the more neighboring points from more far away of the window's center are incorporated.

If you have problems getting your data smoothed sufficiently even with the maximum filter length of N=199 you should consider data thinning or other resampling options to reduce the total number of data points.

**Standardization**

Affects y-data

Standardization makes data of different scales comparable.

| Value | Description |
|---|---|
| None | Disabled |
| Min-Max-Normalization | Scales intensity to values between 0 and 1 |

| Value | Description |
|---|---|
| Vector Normalization | Centers intensities and scales by mean square error (MSE) such that values have a mean of 0 and a 2-norm of 1 (i.e., the Euclidean length becomes 1). |

### 3.3.3.3 Hard Model

**Baseline polynomial order**

| Value | Description |
|---|---|
| Linear | The baseline is a linear polynomial |
| Quadratic | The baseline is a quadratic polynomial |

**Default peak type**

The default peak refers to a mathematical function used for newly added peak models. This setting is currently restricted to a single type.

| Value | Description |
|---|---|
| Pseudo-Voigt | The Pseudo-Voigt function is a linear combination of Gaussian and Lorentzian functions:<br>$V = \alpha \left[\beta \cdot \exp\left(-\ln(2)\frac{(x-\delta)^2}{\gamma^2}\right) + (1-\beta)\frac{\gamma^2}{(x-\delta)^2-\gamma^2}\right]$<br>$\alpha$ = peak maximum at the peak's center position<br>$\beta$ = Gaussian-Lorentzian-ration between 0 and 1<br>$\gamma$ = half width at half maximum (HWHM)<br>$\delta$ = center position<br>x = independent variable<br>The area under the peak curve is:<br>$A = \alpha \cdot \gamma \left\lfloor \frac{\beta}{\sqrt{\ln(\ln(2)/\pi)}} + (1-\beta)\pi \right\rfloor$ |

### 3.3.3.4 Parameter Constraints

Parameter constraints limit values of hard model parameters when being adjusted automatically during model fitting. However, parameter constraints can easily be violated by manually setting parameter values outside the constrained range or by tightening constraints such that previous model parameters are then out of bounds. Make sure to not accidentally violate parameter constraints manually.

Caution: Model fitting may not work correctly for hard models with violated parameter constraints.

**Peak Type Selection**

Select a peak type for which parameter constraints should be set

## Peak Fitting

For peak fitting, only width and shape parameters can be constrained. Position and amplitude parameters are unconstrained.

**Min / Max**

Absolute lower and upper bounds on individual parameters, i.e. parameter values are guaranteed to be in between min and max.

## Component Fitting

For component fitting, all peak parameters can be constrained as well as the component shift parameter. Component weights are unconstrained.

**Variation**

Relative bounds on parameters, i.e. parameter values are allowed to vary relative to their starting value (in both directions). Also parameters still are guaranteed to be in between min and max (see above).

**In Percent?**

If yes, variation of parameters is measured in percent. If no, variation is measured in absolute parameter units. This setting is uneditable.

**Constant peak intensity**

This constraint results from the physical observation, that for a constant concentration, the peak intensity (peak area) is constant even if the peak shape changes, e.g. due to molecular interactions. Therefore, peak parameters amplitude, width, and shape are coupled. This should be considered when changing the bound constraints for these parameters. This setting is uneditable.

# 3.3.3.5 Fitting Options

## Peak Fitting (Auto-fit)

**Auto-fit mode**

| Value | Description |
|---|---|
| Regular (Sequential) | Peak models are added and fitted one by one. The next peak model will be added at the position of the largest positive residuals. This method is slow because all peaks are fitted every time a new peak is added, but sometimes this leads to a better accuracy of the final fit. |

| Value | Description |
|---|---|
| Fast (Simultaneous) | Peak models are added and fitted in groups. Peak positions are determined by means of a peak search. This method is fast and often leads to more physically reasonable results. |

**Weighted peak search**

| Value | Description |
|---|---|
| Enabled | Residuals between model and sample are weighted by the absolute sample intensity to take into account that noise typically scales with signal intensity, i.e. residual noise on top of a large peak might be even larger than small residual peaks. This way, "noisy peaks" are suppressed and the next peak model is more likely to be added at the position of a real peak. |
| Disabled | The next peak model is added at the position of the largest positive difference between model and sample. |

## Component Fitting

### Fitting Mode

The fitting mode controls the extent of model flexibility (free parameters besides component weights) during component fitting as part of the following analyses:

- component fitting analysis
- prediction analysis and validation analysis using an IHM calibration

The fitting mode is not used in the following cases in which component fitting is also performed:

- HMFA analysis: here, the component fitting mode is always "no interaction"
- automatic parameter adjustments during hard modeling: here, component fitting is controlled by a separate choice of parameters

Component fitting modes are named after the extent of molecular interaction among components in the mixture. Molecular interaction is supposed to be responsible for peak variations.

| Value | Description |
|---|---|
| No Interaction | Only component weights and baseline parameters are adjusted. This is a linear fit and thus very fast. |
| Low Interaction | Component weights (W), component shifts (S), and baseline (B) parameters are adjusted. The component shift parameter shifts all peaks of a component at once. |

| Value | Description |
|---|---|
| Medium Interaction | WSB parameters and individual peak positions are adjusted. The actual number of individual peaks being shifted is controlled by the "No. of Considered Peaks" option (see below). |
| High Interaction | WSB parameters and all peak parameters are adjusted. The actual number of individual peaks being shifted and deformed is controlled by the "No. of Considered Peaks" option (see below). |
| Very High Interaction | All model parameters are adjusted simultaneously. This fit is very time-consuming and can only be recommended for very small models. |

**Initialization of Fit**

Component fitting is an iterative process. This setting can be used to change starting values of model parameters.

| Value | Description |
|---|---|
| none | Component fitting starts with current parameter values. |
| linear fit | A linear fit of component weights and baseline parameters precedes the actual component fit. Exception: no preceding linear fit is performed when component fitting mode is set to "no interaction" which is a linear fit itself. |

**No. of Considered Peaks**

This setting only takes effect if component fitting modes "medium interaction" or "high interaction" are selected. In both cases, the number of actually adjusted peaks is limited by this setting. The fitting algorithm only considers the most important peaks, i.e. peaks which have the largest influence on improving the fit. Increasing the number of considered peaks increases computational time but not necessarily increases the accuracy of the fit because also insignificant peaks will be fitted. As a rule of thumb, the number of considered peaks should be as high as the number the significant/large peaks in each component model.

**Component Weight Threshold**

Component models with a component weight below the threshold will not be adjusted at all during component fitting. The threshold value applies if one would scale the model to a maximum y-value of 1. The value could then be interpreted as a fraction of 1.

### 3.3.3.6 Optimizer Options

Optimizer options are for advanced tuning and should be modified by experienced users only.

## Linear Fitting

Linear fitting options apply to component fits (no interaction) and HMFA component analysis. It is recommended not to change the default settings.

**Max. Iterations**

Stopping criterion. Optimization stops after a maximum number of iterations.

**Objective function tolerance**

Termination tolerance. Optimization stops only if objective function value drops below the tolerance. The value should be between $1e^{-10}$ and $1e^{-15}$.

## Nonlinear Fitting

Nonlinear fitting options apply to peak fits and component fits (low to very high interaction).

**Max. Iterations**

Stopping criterion. Optimization stops after a maximum number of iterations.

**Objective function tolerance**

Termination tolerance. Optimization stops only if objective function value drops below the tolerance. The value should be around $1e^{-6}$.

**Parameter tolerance**

Termination tolerance. Optimization stops only if progress in improving parameters drops below tolerance. The value should be around $1e^{-6}$.

**Constrained fitting algorithm**

This option only applies to constrained optimization problems, e.g. components fits with interaction high and very high.

| Value | Description |
|---|---|
| Active-set | Moderate progress during all iterations. This algorithm typically stops due to small progress. Default termination tolerances should be used. |
| SQP | Thorough optimization. Good progress at the beginning, steady progress at the end. This algorithm typically stops when the maximum number of iterations is reached. Therefore, the maximum number of iterations should be chosen carefully. |

**Constraint violation tolerance**

Termination tolerance. Optimization stops only if constraints are violated by less than the tolerance. The value should be around $1e^{-2}$. This option only applies to constrained optimization problems, e.g. components fits with interaction high and very high.

### 3.3.3.7 Calibration Model

#### Validation

Validation is a sub-step of calibration to validate the predictive quality of a model. In general, the kind of validation that can be performed depends on the provided data. PEAXACT will always calculate the root mean squared error of calibration (RMSEC) from the training samples. If independent test samples are provided, the root mean squared error of prediction (RMSEP) is calculated as well. Cross-validation is performed according to user preferences.

**Cross-validation (CV)**

| Value | Description |
|---|---|
| None | Cross-validation is disabled (not recommended). |
| Leave-N-out | Leave-N-out cross-validation is performed on the training samples by simulating a set of independent samples. The set of all training samples is partitioned into K subsets, each leaving N samples out. For each subset, a new calibration is performed and the left-out samples are predicted. After repeating this for all K subsets, a mean prediction error is calculated: the root mean squared error of cross-validation (RMSECV).<br><br>Cross-validation is slow because it involves K calibration steps. However, it is a convenient method because it does not require additional test samples. |

**N**

Specify N, the number of samples to leave out in each subset during cross-validation.

**Max. CV subsets**

Specify K, the maximum number of subsets during cross-validation. K takes precedence over N if leave-N-out would result in more subsets than K. In that case N is automatically increased to result in K subsets at most.

## 3.3.4 Report Settings

Report settings apply to tabular and graphical reports generated by report generators such as the Data Inspector, the calibration report generator, or analysis report generators.

#### Plot Appearance

**Font Style**

Choose font family, font decoration, and font size for text in graphical reports

**Visible Elements**

| Value | Description |
|---|---|
| Title | Show/Hide titles of graphical reports. The title is always displayed above the plot. |
| Axis Labels | Show/Hide descriptive labels at each axis. |
| Axis Ticks | Show/Hide axis tick labels. |
| Legend | Show/Hide legend (if any). |

**Sample X/Y-Label**

Default axis labels for sample plots

## Table Appearance

**Include rows with general information**

| Value | Description |
|---|---|
| Enabled | Tabular reports start with a few rows containing general information like the report creation time or the model file the analysis is based on (if any). |
| Disabled | The report consists of analysis results only. |

**Include columns with data set features**

| Value | Description |
|---|---|
| Enabled | The report's result table is expanded to also include all available features of the data sets being analyzed. |
| Disabled | The report consists of analysis results only. |

## Figure Export Options

**Image resolution (dpi)**

This option only applies when exporting images to a file.

| Value | Description |
|---|---|
| Auto | A default resolution for exporting graphical reports to file. The actual value depends on the output file format. |

| Value | Description |
|---|---|
| 72, 100, 150, 200, 300, 600 | A fixed resolution (in dots per inch) for exporting graphical reports. A larger value improves image quality, but also increases file size.<br>Note: The value is ignored when copying graphical reports to clipboard because the clipboard has a limited size. |

**Export scalable vector format**

This option only applies when copying images to clipboard using the Copy-tool from the toolbar.

| Value | Description |
|---|---|
| Enabled | Graphical reports are copied as vector images. |
| Disabled | Graphical reports are copied as raster images (bitmaps). |

# 3.4 Model File Management

## 3.4.1 About Models

### What are Models?

A model is a set of parameters used by PEAXACT algorithms to solve analysis tasks, e.g. predicting concentrations from measured spectra. A PEAXACT model is also called ***master model***. The master model contains sub-models which group parameters:

- Data pretreatment model: parameters for data manipulation
- Integration model: parameters for peak integration
- Hard model: parameters of hard models
- Calibration model: parameters for quantitative analyses

Almost everything you do with PEAXACT is related to models. You use PEAXACT to create and edit models, load and save models, and apply models to data for analysis purposes.

### Modeling and Analysis

The lifetime of a model passes at least two stages:

- First, an analysis method must be chosen and appropriate model parameters must be set. This process is called modeling. The amount of information that has to be provided during modeling depends on the analysis method. For instance, calibration methods require a calibration model; hard modeling methods require a hard model.
- After the model has been created, it can be utilized for data analysis.

## Browsing the Model Structure

The Model Tree Panel is a graphical control designed for browsing the master model's structure. It quickly gives an overview of all the sub-models and their content, e.g.

- What kind of data pretreatment is used?
- Does an integration model exist? How many integration ranges are defined?
- Does a hard model exist? Which components are modeled?
- Is the model calibrated?

## Getting Access to Model Parameters

The model tree can be used to select any element of the master model and then get displayed additional information about that element in the Model Properties Panel.

For most model elements, the Model Properties Panel gives access to model parameters and their values. For the Master Model item  , the Model Property Panel displays master model properties; a summary of the most important parameters in a tabular report.

## Model File

The model file is a physical file on the hard disk which is written by PEAXACT and stores all master model parameters.

## Exporting Model Reports

| Menu bar: | File > Export > ... |
|---|---|

Model reports are tabular or graphical representations of model parameters. Model reports can be re-created at any time because the necessary information is persistently stored within the model file. PEAXACT can export reports to different output formats such as XLS or PDF.

The following model reports are available:

**Model Summary**
: This report contains the same information as shown in the Model Properties Panel which is a summary of model parameters.

**Model Details**
: For documentation purposes it is recommended to export the model details report. It's a full report of all model parameters.

**Hard Model XY Data**
: This report contains x-values and y-values of the hard model. This data can be used to reproduce plots of the hard model similar to those shown in the Plot Panel.

**Plot**
: This report exports the current content of the Plot Panel to an image file format.

## 3.4.2 New Model

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | File > New Model |
| Hot key: | CTRL + N |
| Context menu: | Model Tree Panel > New Model |

A new master model is added to the Model Tree Panel.

### Model Name

The new model's name is "Untitled" by default. If an opened model already has this name, a consecutive number is added.

A model is always named after its filename. To rename a model in PEAXACT, save it under a new filename. To rename a model outside PEAXACT, simply rename the file. Caution: Do not change the file extension!

### Initial Settings

Each time a new model is created, its initial settings are copied from the then-active user preferences profile. You can modify settings of each model individually later.

By default, a new model has no integration model, no hard model, and no calibration model. It only contains a pretreatment model because pretreatment settings are inherited from the user preferences profile.

## 3.4.3 Open Model

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | File > Open Model... |
| Hot key: | CTRL + O |
| Context menu: | Model Tree Panel > Open Model... |
| Drag & drop: | from Explorer to Model Tree Panel |

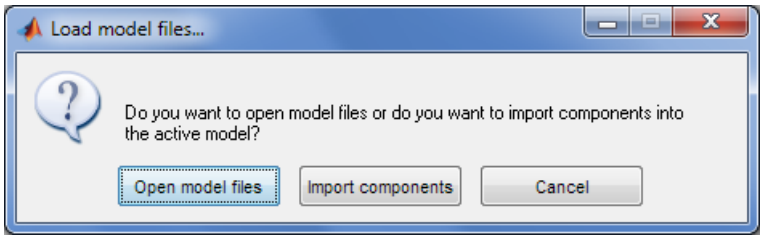
The File Dialog can be used to browse directories and select one or more model files to be opened. Choose a file filter from the pull-down list to localize certain model files. Opened models are added to the Model Tree Panel. The active model does not change.

### Drag and Drop

A fast way to open model files is by using drag & drop. Drag any files from the Windows Explorer (e.g.) and drop them into the Model Tree Panel. Hint: If your directory contains many different files, you can simply select all files and drop them to the Model Tree Panel.

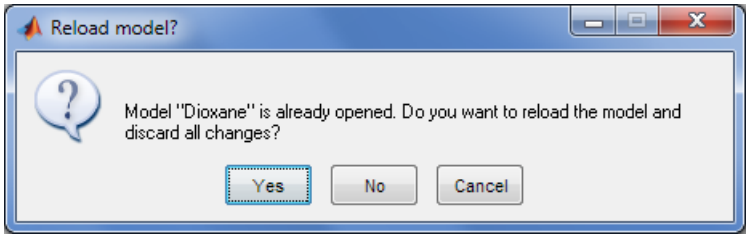
PEAXACT automatically filters out model files. You may also drag & drop directories in order to load all model files from the directory.

If any model is active while dropping files to the Model Tree Panel, you will be asked whether to open these files or to import components into the active model.




### Reload Open Model

If you open a model file that is already opened, you will be asked whether to reload the model or discard all unsaved changes (if any).



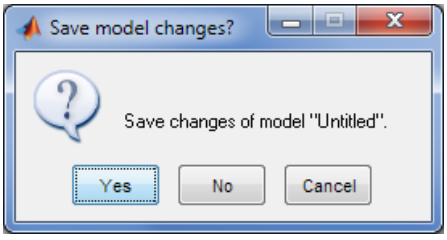

## 3.4.4 Close Model

| | |
|---|---|
| Menu bar: | File > Close Model |
| Hot key: | CTRL + W |
| Context menu: | Model Tree > Master Model Item > Close |

The active model file is closed. The model is removed from the Model Tree Panel and the next available model becomes active.

### Unsaved Changes

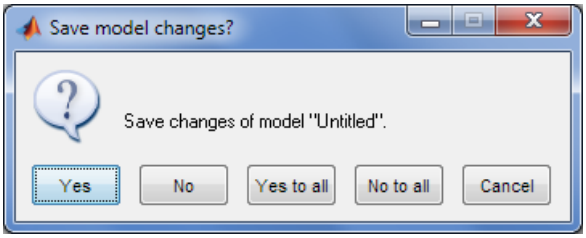
If the active model contains unsaved changes, you will be asked to save the model first. Otherwise, the model file is closed directly.



## 3.4.5 Close All

| | |
|---|---|
| Menu bar: | File > Close All |

All model files are closed. This is a fast way to clear the Model Tree Panel. If any master model contains unsaved changes, you will be asked whether to save them or not. To speed things up, you can choose the same answer for all models.




The empty item is activated next in the Model Tree Panel and the Plot Panel is refreshed accordingly.

### Closing Models on Shutdown

All models will be closed when quitting PEAXACT. Respectively, you get the chance to save unsaved changes.

## 3.4.6 Save Model

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | File > Save Model |
| Hot key: | CTRL + S |
| Context menu: | Model Tree > Master Model Item > Save |

The active master model is saved to its file. If the model has never been saved before, PEAXACT falls back to "Save As".

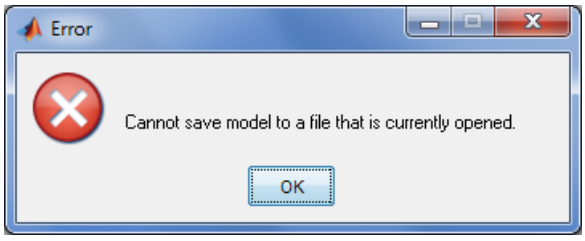
## 3.4.7 Save Model As

| | |
|---|---|
| Menu bar: | File > Save Model As... |
| Context menu: | Model Tree > Master Model Item > Save As... |

The File Dialog can be used to browse directories and select or enter a filename for the model to be saved.

### Overwriting Existing Files

As a precaution, you cannot save a master model to a file of another model that is currently opened in PEAXACT. You have to close the other model before overwriting its file.




## 3.4.8 Duplicate Model

| | |
|---|---|
| Menu bar: | File > Duplicate Model |
| Context menu: | Model Tree > Master Model Item > Duplicate |

An exact copy of the active master model is created. The new model is added to the Model Tree Panel and activated. The duplicate gets the name of the source model plus a consecutive number.

### Using Duplicates

Using duplicates is useful if you already have a populated model and want to compare the effects of different model settings without changing the original model each time.

# 3.5 Data Set Management

## 3.5.1 About Data Sets

### What are Data Sets?

A ***data set*** is a collection of data that belongs together. In PEAXACT, each data set consists of exactly one sample and optional features.

One may think of the data set as a row of a table. The table's columns then represent different pieces of information. The full table of all data sets is referred to as ***data table***.

### The Sample URI

Data sets only contain references to samples. The samples' xy-data are stored in data files. Therefore, each reference simply points to a sample in a file. The reference is composed of a filename and – because the file may contain more than one sample – an access ID. Together, filename and access ID represent a *uniform resource identifier (**URI**)* which unambiguously identifies a specific sample. Learn more about URIs in Section Data Table File.

Example URI:

```
C:\Program files\S-PACT\PEAXACT\data\HMFA\analysis\001min.CSV#1
```

## Data Set Name

In order to make it easier to refer to a certain data set, each one is given a ***data set name***. The name is a shortened URI, namely the filename's base name and the access ID. Data set names are often used in PEAXACT, e.g. in report dialogs, in result files, or as entries in the Data Sets Panel.

Example Data Set Name:

```
001min.CSV#1
```

## Features

The term ***feature*** refers to any kind of information that is optionally attached to a data set. Each feature is denoted by a name and corresponds to a column in the data table with the feature name being the column's title.

Data set features are used in several ways:

- coloring and arranging 2D and 3D plots in the Data Inspector
- sorting data sets in the Data Sets Panel
- plotting results vs. features (e.g. in analysis reports)
- changing the treatment of data sets by assigning special tags
- providing reference values for model calibration and validation

Features are either predictive, i.e. they can be calibrated and later predicted for unknown samples, or non-predictive. Non-predictive features are recognized by the following feature names: Usage, Time, Timestamp, Location. These features are treated specially:

**Usage**

The Usage feature influences the treatment of data sets during modeling and analysis. Values are confined to a predefined list:

| Value | Description |
|---|---|
| unspecified | No special treatment; data set is used in a normal manner. |
| train | Data set is explicitly labeled as training sample for calibration. This tag only affects calibration. It has no effect otherwise. |
| test | Data set is explicitly labeled as test sample for calibration. This tag only affects calibration. It has no effect otherwise. |

| Value | Description |
|---|---|
| ignore | Data set is ignored during calibration and analysis. |

**Time / Timestamp**

Values of the Time and Timestamp features are interpreted as measurement times. While the Time feature can hold any numeric value, the Timestamp feature must be a date (YYYY-MM-DD), a time (hh:mm:ss), or both (YYYY-MM-DD hh:mm:ss). Measurement times can be utilized for the initialization of components fits.

**Location**

Values of the Location feature are interpreted as measurement locations. Measurement locations can be utilized for the initialization of components fits and can be used for the regression of location-dependent calibration models.

### Data File and Data Table File

Both, data files and data table files are physical files on the hard disk. Data files contain xy-data of one or multiple samples and are typically generated by measurement software. Data table files contain at least one data set and are typically created by the user, either with PEAXACT or with Excel.

## 3.5.2 Load Data Files

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | Data > Load... |
| Hot key: | CTRL + L |
| Context menu: | Data Sets Panel > Load... |
| Context menu: | Model Tree > Hard Model Item > Load Associated Sample |
| Drag & drop: | from Explorer to Data Sets Panel |

The File Dialog can be used to browse directories and select one or more data files to be loaded. Choose a file filter from the pull-down list to localize certain data files.

### Drag and Drop

A fast way to load data files is by using drag & drop. Drag files from the Windows Explorer (e.g.) and drop them into the Data Sets Panel. Hint: If your directory contains many different files, you can simply select all files and drop them to the Data Sets Panel. PEAXACT automatically filters out data files. You may also drag & drop directories in order to load all data files and data table files from the directory.

### Creating Data Sets from Data Files

When loading a data file, PEAXACT scans its content for samples, creates URIs, and generates a data set for each sample. URIs are created automatically from the filename, the sample's number within the file, and optionally a block name (OPUS-files) or frame number (other 4D data files). See also Section Data Table File for further information about URIs.

The data set is added to the Data Sets Panel only if

- it is a new data set, i.e. none of the loaded data sets has the same URI
- in case of OPUS-files, the block name matches the block name filter (see user preferences)

## 3.5.3 Load Data Tables

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | Data > Load Table... |
| Hot key: | CTRL + T |
| Drag & drop: | from Explorer to Data Sets Panel |

The File Dialog can be used to browse directories and select data table files to be loaded.

### Drag and Drop

A fast way to load data tables is by using drag & drop. Drag files from the Windows Explorer (e.g.) and drop them into the Data Sets Panel. You may also drag & drop directories in order to load all data files and data table files from the directory.

### Loading Data Sets from Data Tables

When loading a data table file, PEAXACT scans it for Sample URIs, checks if they are valid, and generates data sets from each table row. If a data set is not loaded yet, i.e. if there is no data set in the Data Sets Panel with the same Sample URI, then it is added to the Data Sets Panel. Existing data sets are updated, i.e. new information loaded from the table is merged with the existing information.

## 3.5.4 Unload Selected Data Sets

| | |
|---|---|
| Menu bar: | Data > Unload Selected Data Set(s) |
| Hot key: | CTRL + U |
| Context menu: | Data Sets Panel > Unload Selection |

All selected data sets are instantaneously unloaded and removed from the Data Sets Panel.

Component fits to any of these data sets are removed along with the data sets.

If any selected data set was active, the next available data set is activated automatically and the Plot Panel is refreshed accordingly.

## 3.5.5 Unload All

| | |
|---|---|
| Menu bar: | Data > Unload All |

All data sets are unloaded at once. This is a fast way to clear the Data Sets Panel.

All component fits are removed as well.

The empty item is activated next in the Data Sets Panel and the Plot Panel is refreshed accordingly.

## 3.5.6 Save Data Table

| | |
|---|---|
| Menu bar: | Data > Save Data Table... |
| Menu bar: | Data > Data Inspector... > Data Table Editor > |
| Context menu: | Data Sets Panel > Save as Data Table... |

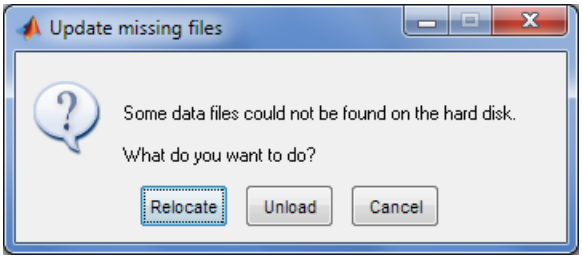
All selected data sets are exported to a data table file. The File Dialog can be used to browse directories and select or enter a filename.

## 3.5.7 Update Missing Files

| | |
|---|---|
| Menu bar: | Data > Update Missing Files |
| Data Sets Panel: | Activate data set whose data file is missing |

All data sets are checked for URIs pointing to missing files on the hard disk. If any are found, the File Update Dialog appears (see Figure below) which can be used to either relocate these file or unload affected data sets.

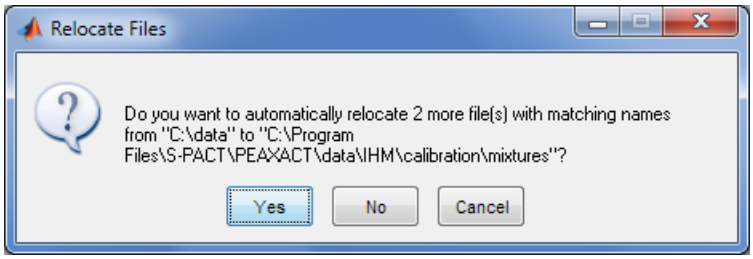
URIs happen to point to missing files if data files are moved or renamed on the hard disk while being loaded in PEAXACT or if a session file is restored which refers to data files that have been moved or renamed since the session file was saved.



### Relocate

Relocating comprises two steps:

- Relocating the next missing file, i.e. select a new target directory and target filename using the File Dialog.
- If possible, automatically relocating more missing files with matching names from the source directory to the selected target directory.




Note: Relocating data files is fast and easy if files were only moved from one directory to another directory. Try to avoid moving data files from one directory to many different directories. Also avoid renaming data files.

Both steps execute until either all missing files are relocated or the user cancels.

### Unload

Unloading removes all data sets which are pointing to missing data files.

## 3.5.8 Edit Data Sets

Editing data sets refers to adding, removing, and modifying features of data sets. It is not possible to edit a data set's sample directory or filename unless the file location is invalid (in which case you should update missing files).

### 3.5.8.1 Edit Data Sets with PEAXACT

PEAXACT comes with its own integrated Data Table Editor which is part of the Data Inspector.

| | |
|---|---|
| Menu bar: | Data > Data Inspector... > Data Table Editor |
| Menu bar: | Analysis > Data Inspector... > Data Table Editor |
| Context Menu: | Data Sets Panel > Data Inspector > Data Table Editor |

In addition, the context menu of the Data Sets Panel offers shortcuts to setting the Usage feature and fully removing all features. Both actions are applied to selected data sets.

| | |
|---|---|
| Context Menu: | Data Sets Panel > Set Usage > ... |
| Context Menu: | Data Sets Panel > Remove Features |

### 3.5.8.2 Edit Data Sets with Microsoft Excel

Microsoft Excel can be used as an external editor for data sets. You can use Excel to create and edit data table files and then load these data tables in PEAXACT. Because Excel is a powerful spreadsheet editor this is a convenient way to edit data sets. However, keep in mind that you have to load the data table file in PEAXACT in order to import its changes.

Please consider the data table formatting rules. Otherwise, PEAXACT may not be able to read the file correctly. In particular, you have to save the Excel file with file extension `.xls` (Excel 97-2003-workbook). Despite these rules, you are free to use all Excel functionalities like formulas, diagrams, filters, conditional formatting, and so on.

### 3.5.8.3 Common practice

A typical procedure of editing data sets involves creating an initial data table file with PEAXACT and then editing the file with Excel:

- Load data files in PEAXACT
- Select all data sets in the Data Sets Panel and open the Data Inspector
- Save the data table (this creates a data table file that contains Sample URIs)
- Use Excel to edit the data table file (e.g. add concentrations)
- Load the final data table file in PEAXACT to import features to existing data sets.

## 3.5.9 Sort Data Sets

| Context Menu: | Data Sets Panel > Sort Selection by |
|---|---|

Sorting is applied to selected data sets. Data sets can be sorted by name, usage, or any other available data set feature.

#### Sort order

Unsorted data sets are sorted in ascending order (a to z, 0 to 9) when sorted once, and in descending order when sorted a second time. The sort order has no effect on any analysis results but sometimes on how the results are displayed.

## 3.5.10 Create Mean Sample

| Menu bar: | Data > Create Mean Sample |
|---|---|

A mean sample can be calculated if:

- at least 2 samples are selected in the Data Sets Panel (otherwise the feature is disabled)

- all selected data sets have valid data files (otherwise you will be prompted to update missing files)
- all selected samples have identical x-axes (otherwise an error is displayed)

The mean sample is the sum of these samples divided by their number.

You will be prompted to enter a filename for saving the sample. Finally, a new data set is added to the Data Sets Panel.

### Data Pretreatment

If a master model is active, its pretreatment model is applied to all selected samples before the mean sample is calculated.

> Note: When applying data pretreatments the new sample will be manipulated permanently. When you activate the new sample in the Data Sets Panel while the model is still active, the model's pretreatment settings are temporarily applied again for visualization. The result may be unexpected in case of smoothing, derivation, or standardization.

## 3.5.11 Create Representative Sample

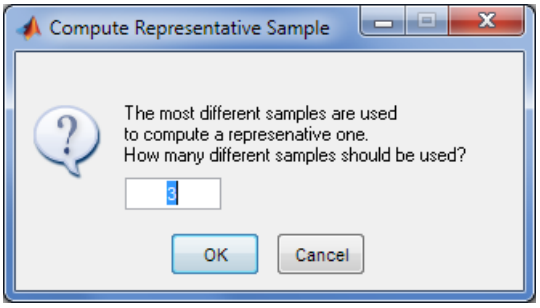
| Menu bar: | Data > Create Representative Sample |
|---|---|

A representative sample is calculated if:

- at least 2 samples are selected in the Data Sets Panel (otherwise the feature is disabled)
- all selected data sets have valid data files (otherwise you will be prompted to update missing files)
- all selected samples have identical x-axes (otherwise an error is displayed)

The representative sample is similar to the mean sample but it is only calculated from the most different samples among the selected samples. The representative sample typically will contain all peaks of all samples in a very pronounced manner. Therefore, it is a suitable sample for creating a representative hard model for HMFA.

Selection of the most different samples is done automatically. You just have to enter the number of samples to consider. The number should correspond to the number of chemical compounds in the data.




You will be prompted to enter a filename for saving the sample. Finally, a new data set is added to the Data Sets Panel.

### Data Pretreatment

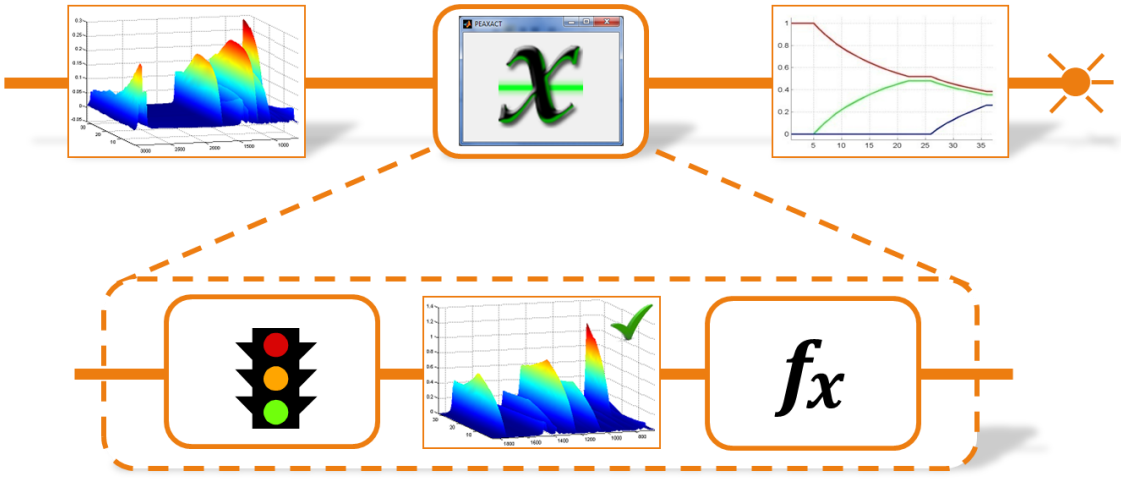
If a master model is active, its pretreatment model is applied to all selected samples before the representative sample is calculated.

> Note: When applying data pretreatments the new sample will be manipulated permanently. When you activate the new sample in the Data Sets Panel while the model is still active, the model's pretreatment settings are temporarily applied again for visualization. The result may be unexpected in case of smoothing, derivation, or standardization.

## 3.6 Analysis

Analysis of data is a process of data transformation by inspecting and modeling data with the goal of highlighting useful information.

PEAXACT covers two kinds of analysis steps: ***data-driven analysis*** and ***model-based analysis***. Data-driven analysis comprises visual data inspection and plausibility tests of data. Model-based analysis requires a mathematical model which has multiple facets depending on individual analysis methods.



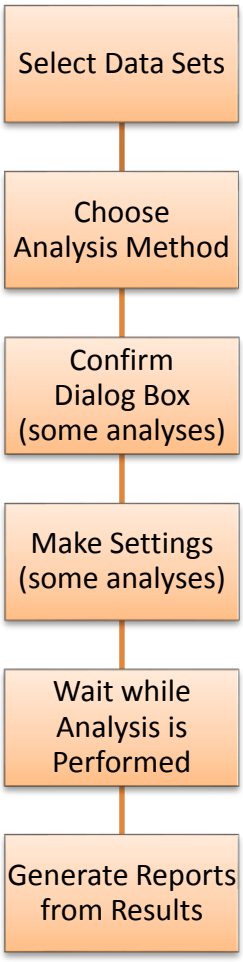

**Conversion of measurements to target quantities in two steps: data-driven analysis and model-based analysis**

The first sub-section of this chapter provides information common to all analyses in PEAXACT. Following sub-sections are dedicated to the specifics of the Data Inspector for visual data inspection and various model-based analyses. See also Section 3.7 to learn more about modeling.

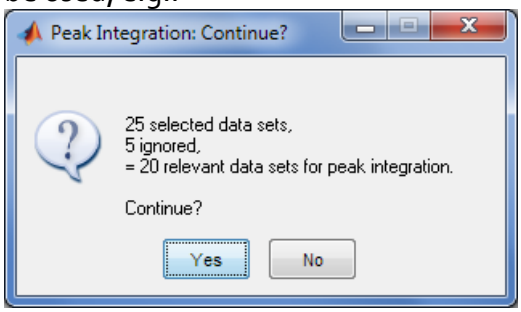
## 3.6.1 General Information

Although the designated use of different analyses naturally differs among each other, there is a common way in PEAXACT how these analyses are applied (workflow) and how results are presented (reporting).

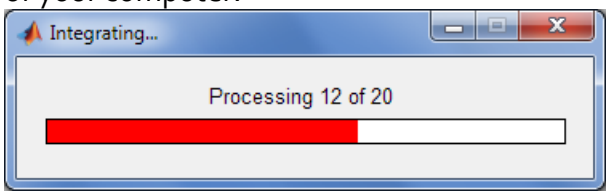
### 3.6.1.1 Typical Analysis Workflow




(1) Select data sets to be analyzed in the Data Sets Panel.
(2) Choose analysis method from menu Analysis.
(3) The confirmation dialog is a short notice on data sets to be used, e.g.:




(4) Some analyses need additional input.
(5) Processing time depends on the chosen analysis, on the number of selected data sets, and on the performance of your computer.



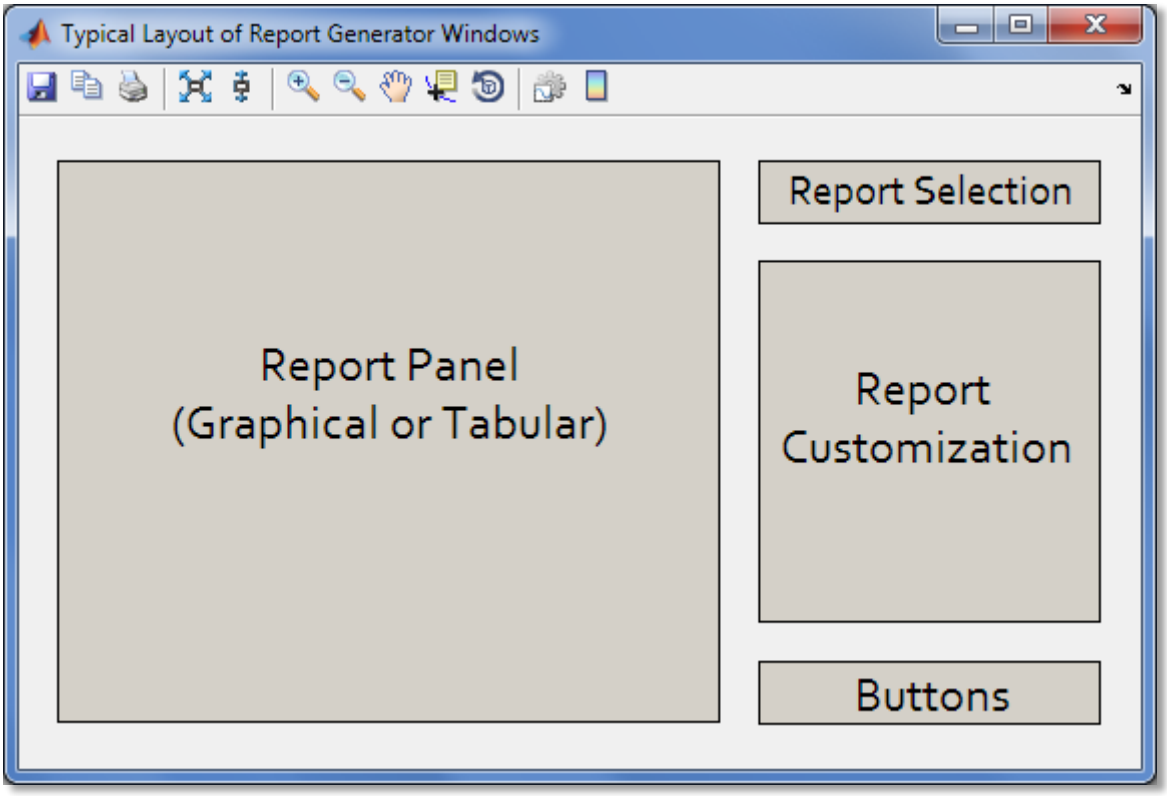

(6) Analysis results are displayed in a Report Generator Window.

**General Analysis Workflow**

See Chapter 5 for full workflow diagrams of individual analyses.

### 3.6.1.2 Report Generator Window

The Report Generator Window is a graphical tool for visualizing and exporting data of any kind, e.g. analysis results. All report generators have a similar layout and common functionalities which are explained in this section. Variations from the typical layout and custom features are explained in separate sections dedicated to the individual report generator.




**Typical Layout of Report Generator Windows**

#### Toolbar

Export report as shown in the Report Panel to file. The exact list of available file formats depends on the type of report, but in general, graphical reports can be exported to image files while tabular reports can be exported to text or spreadsheet files.
Copy report to clipboard. Graphical reports are copied as rendered images or vector images (see Report Settings) while tabular reports are copied as text.
Open Print Preview.
Open User Preferences Editor

**Graphical Reports only**
Full zoom, Full y-zoom
Zoom tools (2D plots only), zoom in or out on a plot
Pan tool (2D plots only), move the view of a graph
Data Cursor tool, read data directly from a graph by placing datatips in a plot
Rotate tool (3D plots only), rotate the orientation of a graph
Show/Hide color bar

> Note: The exact set of available toolbar icons depends on the specific Report Generator Window. However, same icons have the same meaning each time.

Zoom, Pan, Data Cursor and Rotate are interactive tools for graph and plot manipulation. Before using them, you need to enable these tools, either by clicking the toolbar icon, or by using the short keys Q, W, E, R, and T. Pressing the Escape key disables all tools.

## Report Selection & Customization

The report to be displayed in the Report Panel can be chosen from a pull-down list located at the top-right of the Report Generator Window. Each individual report generator has a very specific list of available reports.

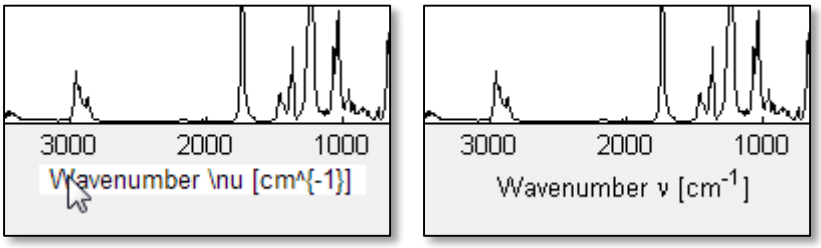
Most reports can be customized in various ways using some controls located at the right of the window.

### Changing Report Appearance

The appearance of reports (e.g. font size, visible elements of graphs) is controlled by several adjustable preferences. You can reach the Preferences Editor from the toolbar.

### Editing Titles and Labels

All titles and labels of graphical reports can be edited when clicking on them.

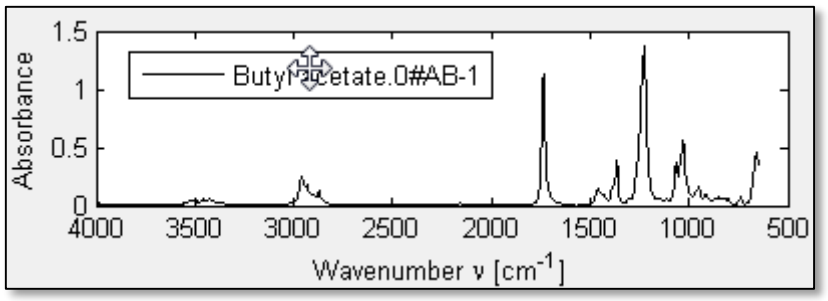



Note that your text is processed by a TeX interpreter, i.e. you can use `^` for superscripts, `_` for subscripts, `{ }` for grouping. Also, you can use special commands for Greek letters (`\alpha`, `\beta`, `\gamma`, `\Gamma`, ...). When finished editing, press the Escape key or click anywhere outside the text field.

> Note: User edited titles and labels are replaced by default text each time the selected report changes. Therefore, it is recommended to customize titles and labels as a final step before exporting the report to a file.

### Positioning of Legends

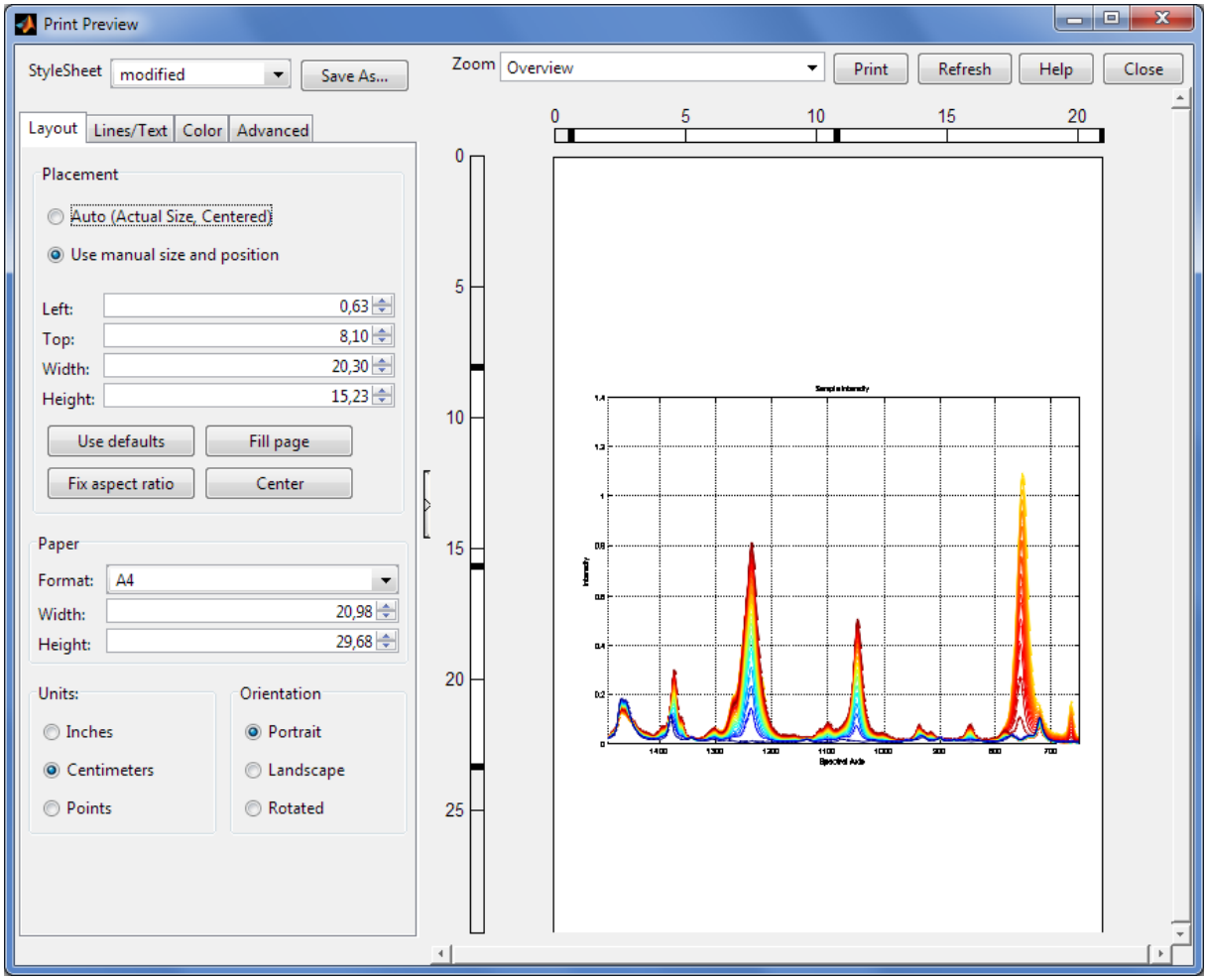
Legends in graphical reports are positioned automatically to least interfere with the plots. However, you can reposition legends interactively by dragging them to the desired location.




### 3.6.1.3 Print Preview Dialog

The Print Preview is a dialog box showing the graphical report as it will print. A scaled version of the report displays in the right-hand pane of the window.

Use the Print Preview dialog box, shown below, to control the layout and appearance of graphical reports before sending them to a printer or print file.




**Print Preview Dialog**

## 3.6.2 Data Inspector

| | |
|---|---|
| Menu bar: | Data > Data Inspector... |
| Menu bar: | Analysis > Data Inspector... |
| Context Menu: | Data Sets Panel > Data Inspector |

The *Data Inspector* is a graphical tool designed for editing and visualizing data sets. For this purpose, the Data Inspector contains two tabs:

- *Data Table Editor* for editing features of data sets
- *Data Plotter* for visualizing samples and features in various 2D and 3D plots.

The Data Inspector does not require a model; it operates on data sets alone and can therefore be used for data-driven analysis. You typically use the Data Inspector to visually inspect measured samples and manage features before you start with modeling.

**Data-driven analysis – data inspection and plausibility tests**

To learn more about data sets and data tables, see Section About Data Sets.

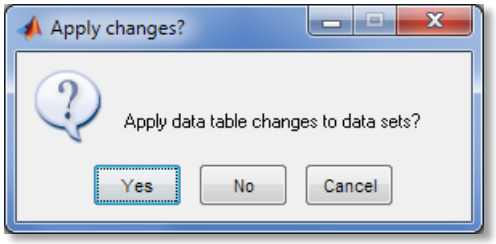
### Starting the Data Inspector

Each time you open the Data Inspector, all data sets selected in the Data Sets Panel are loaded into the Data Inspector and the Data Table Editor and the Data Plotter are updated accordingly. To load different data sets, change the selection in the Data Sets Panel and open the Data Inspector again.

Note: Restarting the Data Inspector refreshes the content of the Data Table. Any unsaved changes made in the Data Table Editor will get lost.

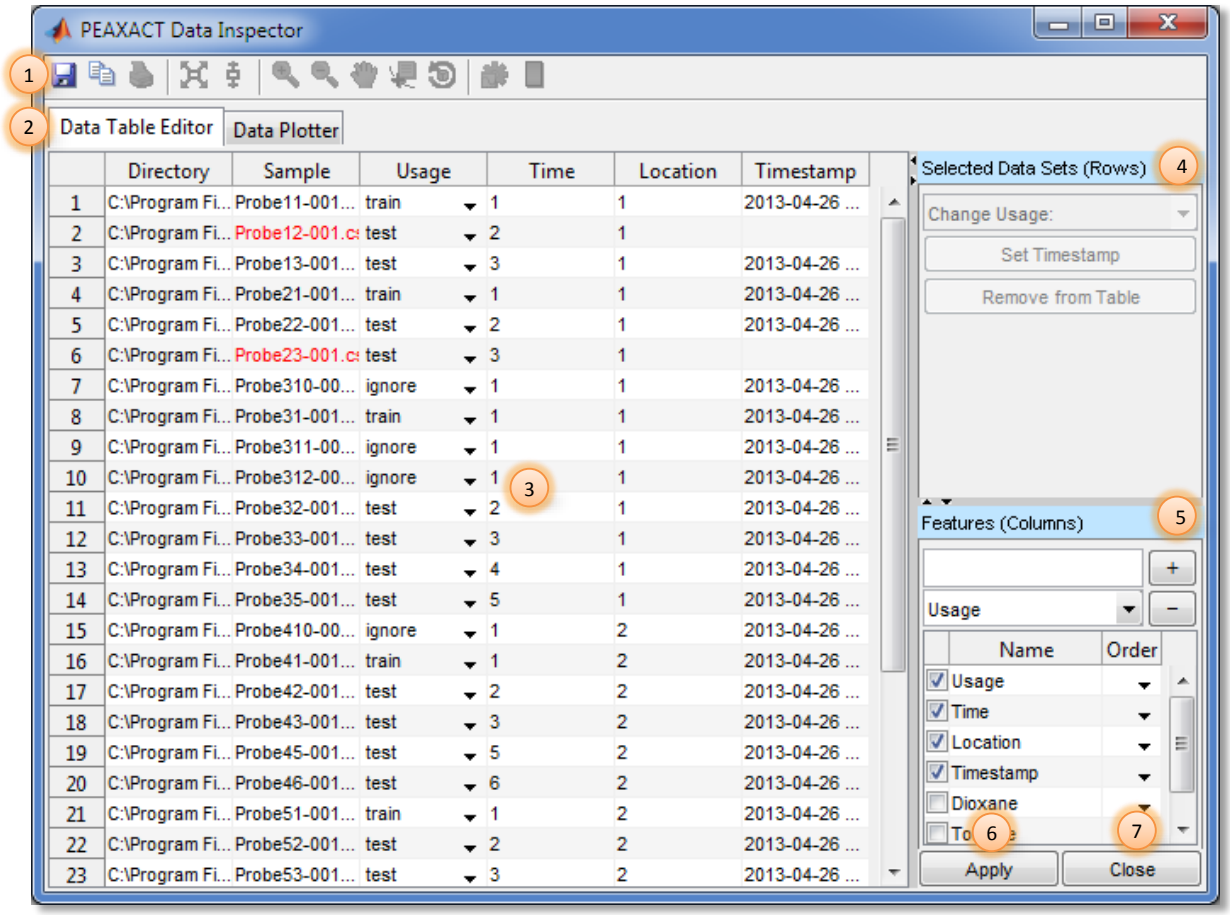
### Closing the Data Inspector

When closing the Data Inspector while the Data Table Editor has unsaved changes, you will be asked whether to apply these changes to data sets.



### 3.6.2.1 Data Table Editor

The *Data Table Editor* is a tool for modifying features of data sets. Each feature is represented by a column in the data table. The first two columns are not editable; they contain the sample's directory and filename.




**Data Inspector – Data Table Editor**

(1) Toolbar
(2) Tabs
(3) Data Table, consists of columns:

- Directory: path of data file (not editable)
- Sample: data set name (not editable), printed in red if file is missing
- any number of feature columns (editable)

(4) Actions applicable to selected rows
(5) Actions applicable to feature columns: add, remove, hide, show, rename, arrange
(6) Apply button: copy table changes to data sets
(7) Close button: close Data Inspector

#### Export Options

The Toolbar of the Data Table Editor contains two export tools.

 Export table to a Data Table File.
 Export table to clipboard

In both cases, the content of columns Directory and Sample is merged and exported as a single column representing Sample URIs. For more information about Sample URIs see Section Data Set Management.

> Note: When exporting the Data Table to a file, all URIs pointing to subdirectories of the table file location are converted to relative paths.

## Edit Table

You can edit the table by clicking into a table cell and modifying the value. Use the Enter (Escape) key to accept (discard) your changes. Press the Tab key or the arrow keys to accept changes and move the cursor to an adjacent cell.

> Note: Changing the table does not yet change the underlying data sets. You have to click the Apply button to copy tables changes back to data sets.

Except for the following special columns, all values must be numeric. If an invalid value is entered it is discarded and the previous value is restored.

**Directory**
: This column is not editable. It shows the directory of the sample file.

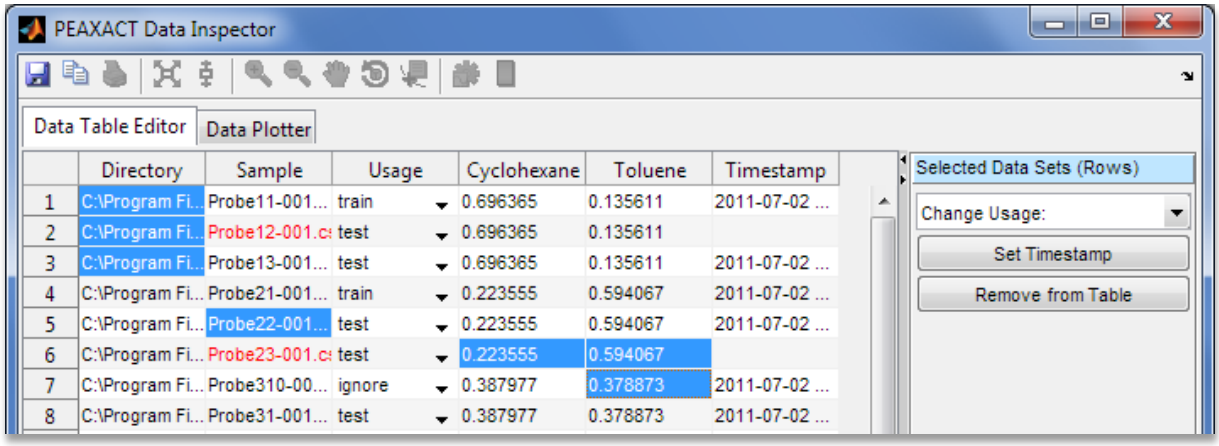
**Sample**
: This column is not editable. It displays the data set name (sample filename and access ID). If the value is printed in red, the file does not exist in the given directory.

**Usage**
: Values of the Usage feature are confined to a predefined list and can be chosen from a pull-down menu.

**Timestamp**
: Values must be entered as dates (YYYY-MM-DD), times (hh:mm:ss), or a combination of both (YYYY-MM-DD hh:mm:ss).

## Alter Selected Table Rows

From the top-right panel of the Data Table Editor you can apply some actions to all selected data sets of the table. A data set counts as selected if any cell in the corresponding row is selected (see Figure). It is recommended to make the selection in the Directory column or Sample column because clicking on these cells will not activate the cell for editing.




**Selecting Data Sets in the Data Table Editor**

### Change Usage

Choose a value from the "Change Usage" pull-down menu in order to set the same value in all selected data sets. The Usage column is added if it is not visible.

### Set Timestamp

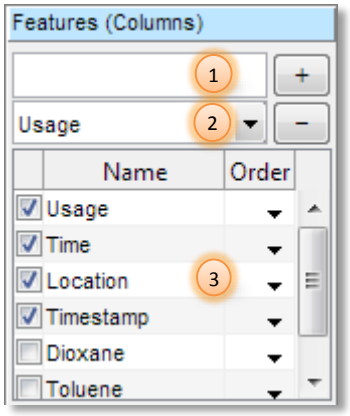
Click the "Set Timestamp" button in order automatically set the value of the Timestamp column. The column is added if it is not visible. Existing values will be overwritten. The value is read from the data file content if available, or from the file system otherwise.

### Remove Rows

Click the "Remove from Table" button in order to remove all selected data sets from the table. Data sets are only removed from the Data Inspector but not from the Data Sets Panel of the desktop.

## Alter Table Columns

From the bottom-right panel of the Data Table Editor you can perform some actions to modify table columns.




(1) Text field for adding new features
(2) Pull-down menu for removing existing features
(3) Feature list (list of visible and invisible features).

The table layout is controlled by the feature list (3). You can show (hide) a table column by placing (removing) the checkmark of the corresponding feature.

Each time the Data Inspector is opened, the feature list is automatically updated to only show available features.

**Show / Add Colum**
The simplest way of adding a column is by placing a checkmark in the feature list in order to show a hidden feature. However, if the feature name you wish to add does not exist in the list, you can enter a new name in the text field (1) and press the "+" button afterwards. The name is added to the feature list (3) and a new column is added to the table.

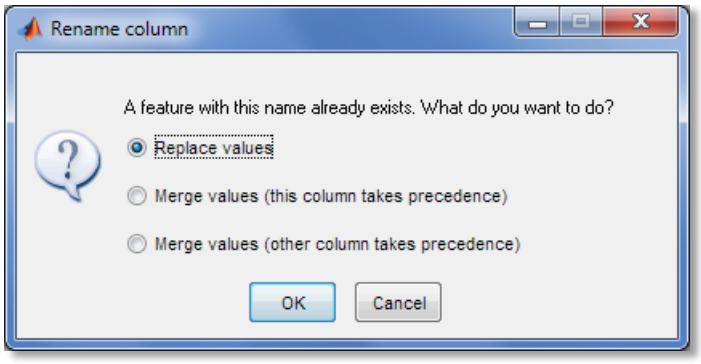
Feature names must be unique, i.e. you cannot add the same feature twice. If you try to add a feature that already exists in the feature list but is currently hidden, the column is simply made visible again.

Names of predictive features are not case sensitive; names of non-predictive features are. See Section About Data Sets for a list of non-predictive feature names. If you enter a non-predictive feature name it is automatically converted to upper case syntax (e.g. "Timestamp" instead of "timestamp").

**Hide / Remove Column**
By unchecking its checkbox in the feature list columns can be hidden temporarily. In order to permanently remove columns, choose a feature name from the pull-down menu (2) and press the "-"-button afterwards. The corresponding column will be removed from the table and from the feature list (3).

**Rename Column**
You can change a feature's name in the feature list (3).

If you try to rename a feature to another name in the list, you will be asked whether to replace or merge column values.



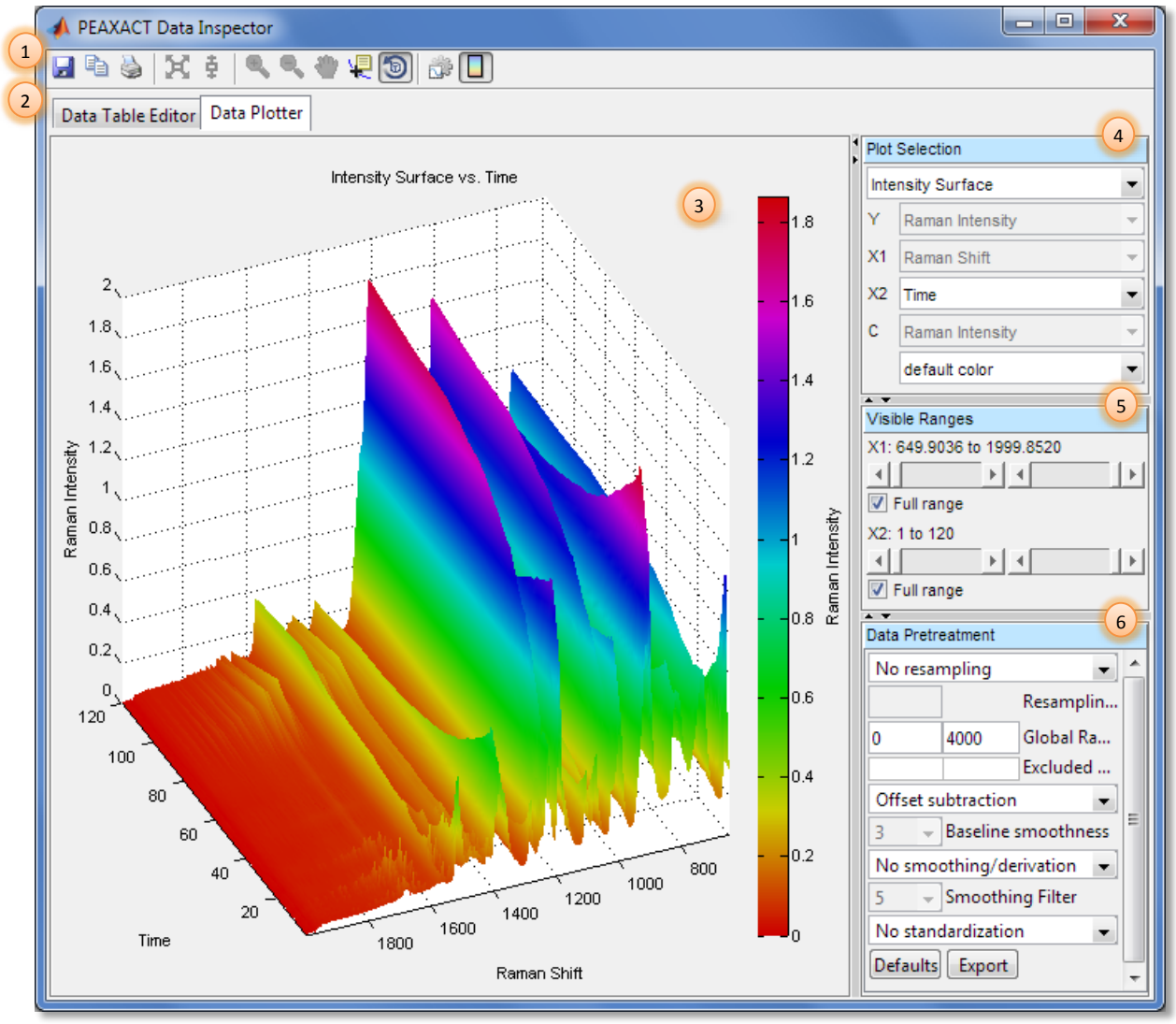

Note: You can neither change the name of feature "Usage" nor rename any other feature to "Usage"

**Re-arrange Columns**
The order of columns can be changed in the feature list (3).

## 3.6.2.2 Data Plotter

The ***Data Plotter*** is a report generator. It generates customizable graphical reports of xy-data and data set features.

The layout of the Data Plotter matches the typical layout of Report Generator Windows. See Section 3.6.1.2 for a description of functionalities common to all report generators. This Section focuses on specifics of the Data Plotter Report Generator.



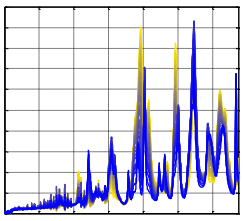

**Data Inspector – Data Plotter**

(1) Toolbar
(2) Tabs
(3) Report Panel
(4) Plot Selection Panel
(5) Visible Range Panel
(6) Data Pretreatment Panel

## Available Reports

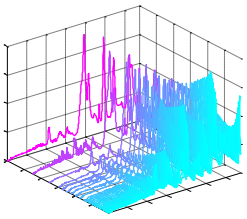
The Data Plotter generates the following graphical reports which can be chosen from the Plot Selection Panel:

**Samples (2D)**



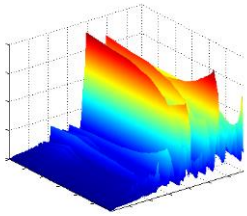
2D-plot of xy-data of samples, overlapping
y-axis: Signal Intensity
x-axis: Spectral Axis
colored by: selectable feature

**Samples (3D)**



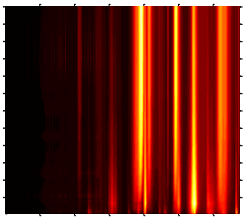
3D-plot of xy-data of samples, arranged along the x2-axis
y-axis: Signal Intensity
x1-axis: Spectral Axis
x2-axis: selectable feature
colored by: selectable feature

**Intensity Surface**



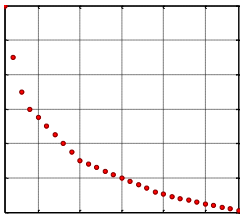
3D-surface plot of signal intensity
y-axis: Signal Intensity
x1-axis Spectral Axis
x2-axis: selectable feature
colored by: Signal Intensity

**Intensity Contour**



2D-contour plot of signal intensity
x1-axis: Spectral Axis
x2-axis: selectable feature
colored by: Signal Intensity

**Features**

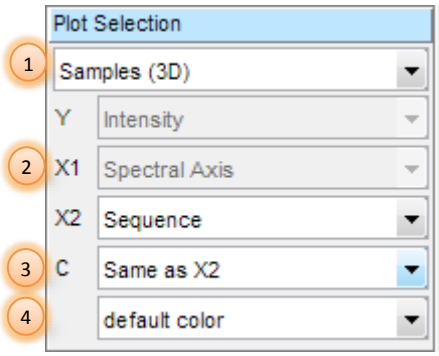


2D-plot of any two features of data sets.
x-axis: selectable feature
y-axis: selectable feature
pre-defined color

## Report Customization

### Axes and Colors

The Plot Selection Panel contains controls for changing axes variables and colors.

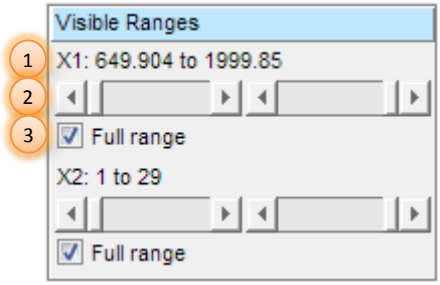



(1) Select report
(2) Choose axis variable from list of available features. Labels X (or X1/X2) and Y correspond to axes.
(3) Choose coloration variable from list of available features.
(4) Choose color scheme from list of popular color schemes. The "default color" depends on the selected report and for some reports also on the selected axis variable.

Note: The actual number of visible and editable pull-down controls in the Plot Selection Panel depends on the chosen report.

### Visible Ranges

The Visible Ranges Panel contains controls for adjusting visible limits of axis and coloration variables.

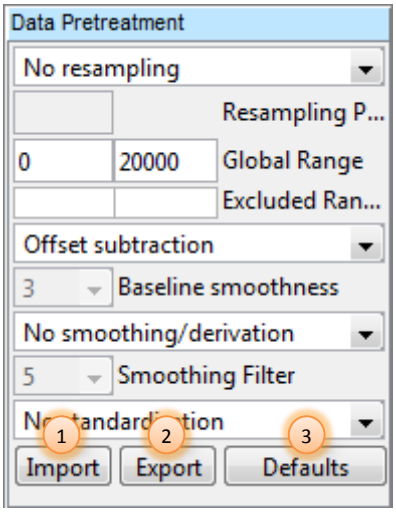



(1) Visible range of variable
(2) Move sliders to change visible range of variable
(3) Show full range if control is ticked; show adjusted range if not ticked

Note: The actual variables shown in the Visible Ranges Panel depend on the chosen report.

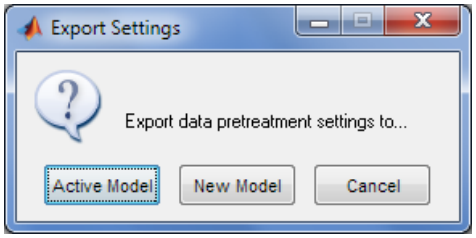
### Data Pretreatment

The Data Pretreatment Panel of the Data Plotter contains controls for data manipulations. The impact of these settings can directly be previewed in the Report Panel.



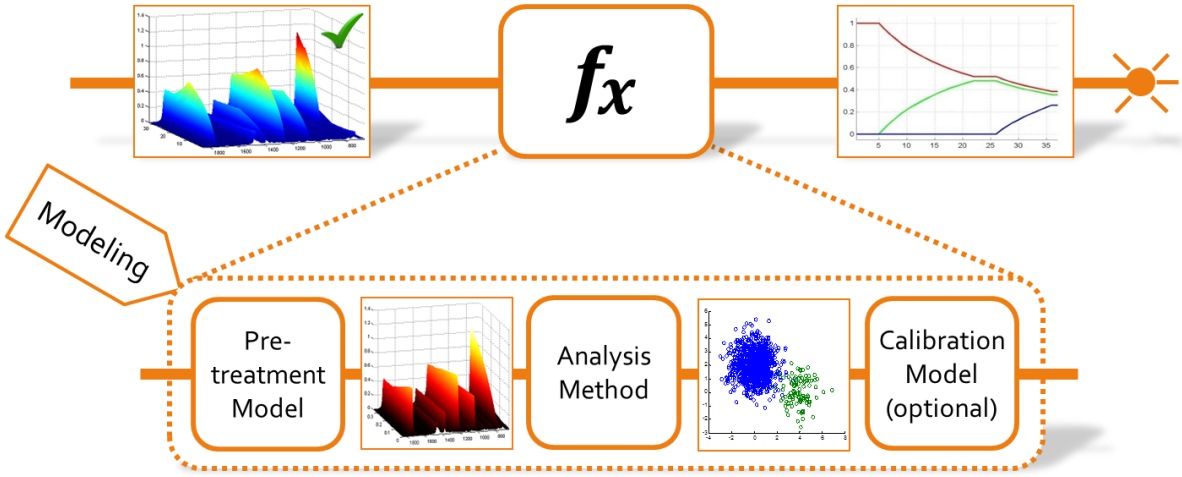

(1) Import Button: Import pretreatment settings from active model
(2) Export Button: Export pretreatment settings to active model or to new model
(3) Defaults Button: Reset settings to active user preferences profile

Pretreatment settings used by the Data Plotter have the same meaning as in the Pretreatment Model. However, the Data Plotter keeps its own instance of settings. You can exchange settings between the Data Plotter and the active model by using the Import (1) and Export (2) buttons. On export you can also choose to create a new model.




## 3.6.3 Model-based Analysis

Model-based analysis refers to the conversion of measured data to quantities of interest. This Section explains the application of several analyses which make use of or require a model. See Section 3.7 to learn more about modeling.



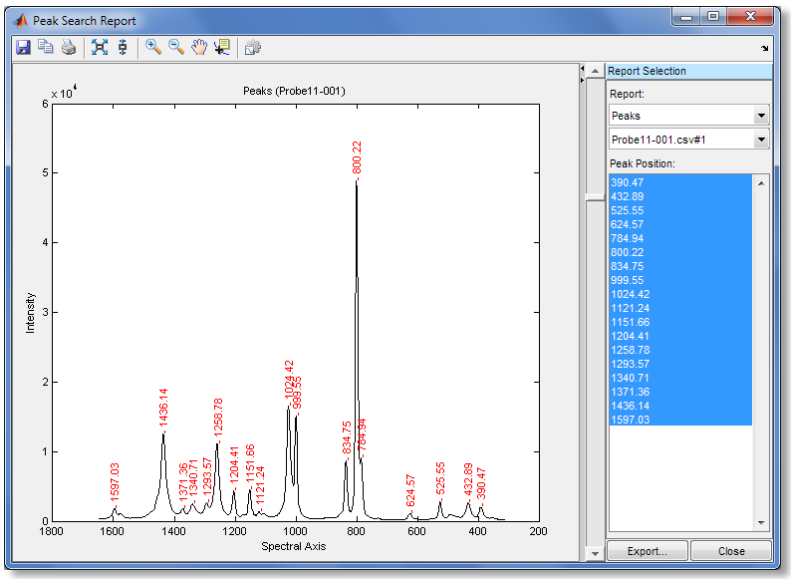

**Model-based analysis – conversion of measurements to target quantities**

## 3.6.3.1 Peak Search

| Menu bar: | Analysis > Peak Search |
|---|---|

Peak search is applied to selected data sets. A model is not required for this analysis, but if one is active, data pretreatments are applied.

Peak search analysis finds local maxima in the signal and displays corresponding x- and y-values. Results are displayed in a Report Generator Window.



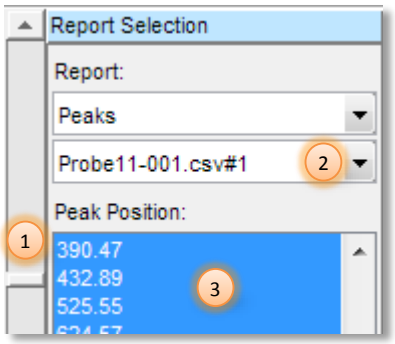

**Peak Search Report Generator**

### Available Reports

**Peaks**
Positions of peaks plotted (in red) into chart of selected sample (black)

**Report Table**
Table with numeric values of peak positions and peak intensities

## Report Customization

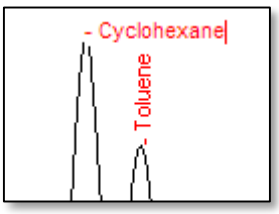



**Peak Detection Limit**
Move the slider (1) up and down to increase and decrease the detection limit of peaks.

**Sample Selection**
The report shows results for a single sample at a time. Select the visible sample from the pull-down menu (2).

**Peak Selection**
By default, all detected peaks are shown in the graphical and tabular report. Use the listbox (3) to select and show specific peaks only.

**Edit Peak Label**
Left-click on the red peak label to start editing. Note that labels will be reset each time you change the peak detection limit, the sample, or the peak selection.




## 3.6.3.2 MCR-ALS

| | |
|---|---|
| Menu bar: | Analysis > MCR-ALS > Component Number Analysis |
| Menu bar: | Analysis > MCR-ALS > Component Analysis |

Multivariate Curve Resolution (MCR) Alternating Least Squares (ALS) is applied to selected data sets. A model is not required for this analysis, but if one is active, data pretreatments are applied.

Multivariate Curve Resolution designs a group of techniques which intend the recovery of the pure response profiles (spectra, pH profiles, time profiles, elution profiles ...) of the chemical

constituents or species of an unresolved mixture when no prior information is available about the nature and composition of these mixtures.

## The Bilinear Model of MCR

MCR requires the experimental data to be explained reasonably well by a bilinear model using a limited number of components. The MCR bilinear model is usually written down as $D = CS^T$, where $D$ is the spectroscopic data matrix, and $S^T$ and $C$ are the matrices of the pure spectra and the related concentration profiles for each of the compounds (contributions) in the system. $C$ and $S^T$ are the small matrices of the bilinear model that contain profiles of the pure contributions (species, compounds) of the original data matrix and may change chemical meaning depending on the nature of the data

MCR-ALS is an algorithm that solves the MCR basic bilinear model using a constrained Alternating Least Squares algorithm. The constraints used to improve the interpretability of the profiles in $C$ and $S^T$ may respond to chemical properties of these profiles (e.g., non-negativity, unimodality, closure ...) or have a mathematical origin (e.g., local rank and selective windows, trilinear structure ...). The 'art' and expertise in using MCR-ALS stems from the proper selection and application of the constraints that are really fulfilled by the data.

## Component Number Analysis

| Menu bar: | Analysis > MCR-ALS > Component Number Analysis |
|---|---|

Component Number Analysis consists in the estimation of the number of unknown mixture constituents. Results are displayed in a Report Generator Window.

## Available Reports

### Eigenvalues of Principle Components

The eigenvalues of principle components indicate how much each principle component contributes to explaining the variance of the data matrix. The most likely number of components is reached when the next eigenvalue is small, i.e. when the next principle component does not significantly increase the explained variance any more. Note that the diagram only displays major principle components which explain up to 99.9% of the variance.

### Ratios

The ratios plot supplements the Eigenvalues-plot by showing the ratios of two successive eigenvalues. The most likely number of components is indicated by a large ratio.

### Factor Indicator Function

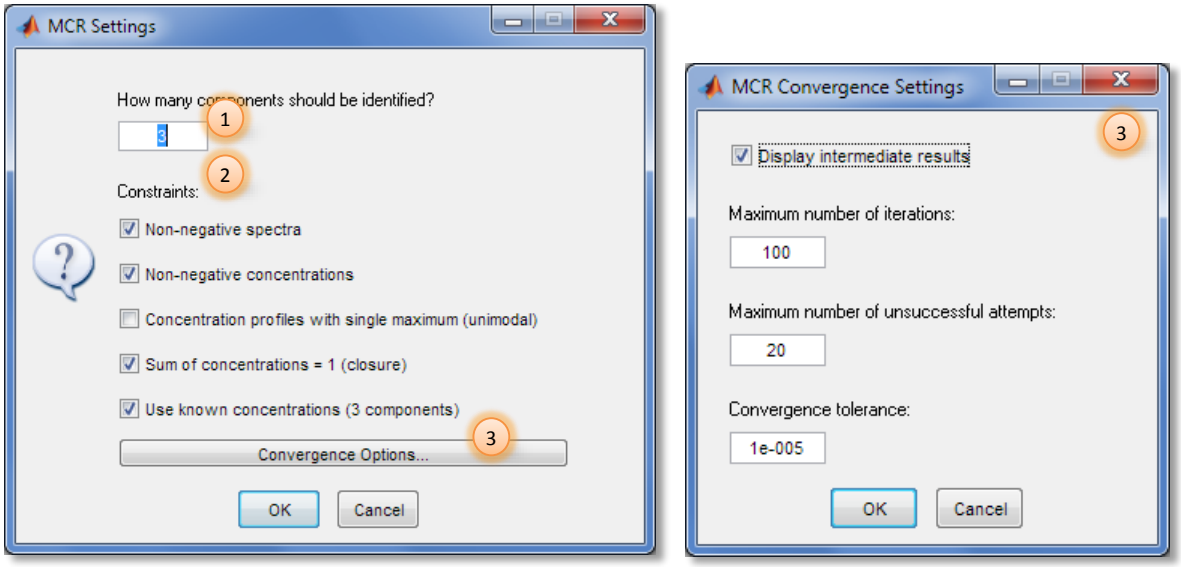
The factor indicator function reaches a minimum where the correct number of factors/components is expected. However, this only occurs when the error is random and fairly uniform throughout the entire data. [E. R. Malinowski, Analytical Chemistry, Vol. 49 (4), April 1977]. In real cases, the function is typically L-shaped for about the first 10 components and the most likely number of components is indicated by a more or less sharp bend in the curve.

## Component Analysis

| Menu bar: | Analysis > MCR-ALS > Component Analysis |
|---|---|

Component Analysis consists in the identification of unknown component spectra and the computation of concentrations for each component. Results are displayed in a Report Generator Window.

Additional input is required for this analysis:




The number of component to be identified must be specified (1) and optional additional constraints (2) to improve the outcome of the analysis:

**Non-negative spectra**

Enable this constraint if y-values of pure component spectra are positive. This is usually valid for spectroscopic and chromatographic data. However, it does not hold if you use a pretreatment model with derivatives.

**Non-negative concentrations**

Enable this constraint if concentration values are positive. This is usually valid for molar or weight concentrations/fractions.

**Concentration profiles with single maximum (unimodal)**

Enable this unimodality constraint if the shape of concentration profile has a single maximum. This also applies to profiles which are monotonously increasing or decreasing.

**Sum of concentrations = 1 (closure)**

Enable this closure constraint if component concentrations sum up to 1.

**Use known concentrations (N components)**

Enable this constraint to use feature values as known concentrations. Feature values could be provided for all or just some data sets. Any available information will

be exploited by MCR. See Section Data Table Editor for how to add features to data sets.

MCR-ALS is an iterative algorithm which uses convergence options (3) as stopping criteria:

**Display intermediate results**
Enable this checkbox to display graphical progress of component spectra and concentration profiles after each iteration.

**Maximum number of iterations**
Algorithm stops after a specified number of iterations

**Maximum number of unsuccessful**
Algorithm stops after a specified number of iterations without progress

**Convergence tolerance**
Algorithm stops after a specified accuracy is reached

## Available Reports

**Component Spectra**
Spectra of identified pure components

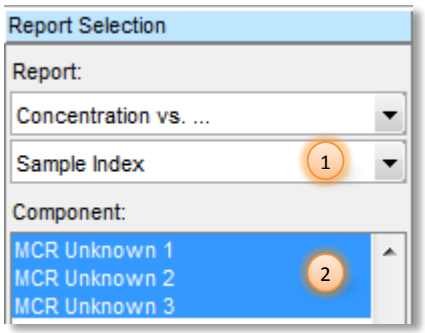
**Concentrations**
Estimated concentrations (dimensionless)

**RMS Spectral Residuals**
Euclidean mean error of identified component spectra fitted linearly to measured samples

**Table: Concentration Profiles**
Numeric values of computed concentrations

**Table: Component Spectra**
Numeric values of pure component spectra

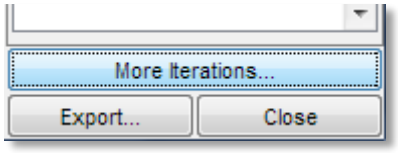
## Report Customization




**X-axis**
By default, concentrations and RMSE values are plotted against the sample index. You can change the x-axis to any available data set feature (1). See Section Data Table Editor for how to add features to data sets.

**Component Selection**
By default, plots of all pure component spectra and concentrations are shown in the graphical report. Use the list (2) for selective plotting. Note that RMSE values are independent of the selected component.

### More Iterations




You could carry out more iterations in order to improve results. When clicking the "More Iterations" button, the MCR-ALS Component Analysis starts at again, letting you choose different constraints if desired, but using previous results as starting values.

## 3.6.3.3 HMFA

| | |
|---|---|
| Menu bar: | Analysis > HMFA > Component Number Analysis |
| Menu bar: | Analysis > HMFA > Component Analysis |

Hard Modeling Factor Analysis (HMFA) is applied to selected data sets. A hard model is required for this analysis.

HMFA is a technique which intends the recovery of pure component spectra and concentration profiles of the chemical constituents of a mixture when no prior information is available about the nature and composition of these mixtures. It is akin to other factor analysis methods, e.g. Multivariate Curve Resolution (MCR), but results are closer to physics due to the usage of hard models.

### The Hard Model of HMFA

The hard model in HMFA is used to provide the factor analysis with as much information as possible about the spectral structure of the mixture. E.g., if any pure component spectrum is known in advance, a corresponding component model should be part of the hard model. All remaining peaks belonging to unknown components should be modeled in the first hard model's component. HMFA will decompose peaks of the first component model into new component models by analyzing the mixture spectra and identifying peaks that belong together.

If, for instance, all pure component spectra were unknown, you could initially create a hard model with all peaks contained in the first component model and let HMFA reassign them to new component models.

The identification of unknown component spectra works well if two requirements are met:

- The first component model must contain all the relevant peaks. No peaks must be missing, because only existing peaks can be reassigned. For creating such a representative model, it is recommended to use a representative sample of the mixture spectra for modeling.
- The mixture spectra must contain reasonable variation in the components' concentrations.

Eventually, HMFA fits the new hard model (which now contains all identified component models and all initially provided component models) to the mixture spectra in order to compute component weights. Please note that this kind of component fitting differs from regular Component Fitting Analysis in the following aspects:

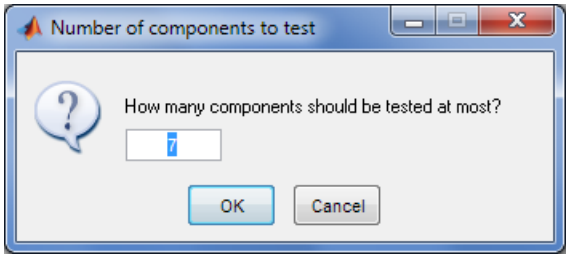
- The hard model is fitted linearly no matter the component fitting mode specified in the model settings.
- Fixed parameters are adjusted anyway.
- The baseline model is ignored; instead a linear baseline is fitted.

## Component Number Analysis

| Menu bar: | Analysis > HMFA > Component Number Analysis |
|---|---|

Component Number Analysis consists in the estimation of the number of unknown mixture constituents. A hard model with at least 2 peaks in the first component is required for this analysis. Results are displayed in a Report Generator Window.

You are prompted to enter the maximum number of components to get tested.




## Available Reports

### Average RMS Residuals of Component Fits

The average of RMS residuals of component fits is the best indicator for determining the component number because it is an a-posteriori figure of merit. A-posteriori means that a concrete model is actually created for each number of components and then tested for how good it explains the data by averaging the RMS residuals over all samples. The most likely number is reached when the next component does not significantly reduce the average RMS residuals any more. Note that the diagram also displays a vertical line which is an estimate of the lower bound.

**Eigenvalues of Principle Components**
The eigenvalues of principle components indicate how much each principle component contributes to explaining the variance of the data matrix. The most likely number of components is reached when the next eigenvalue is small, i.e. when the next principle component does not significantly increase the explained variance any more. Note that the diagram only displays major principle components which explain up to 99.9% of the variance.

**Ratios**
The ratios plot supplements the Average-RMS-residuals-plot and the Eigenvalues-plot by showing the ratios of two successive RMS-values and eigenvalues, respectively. The most likely number of components is indicated by a large ratio.

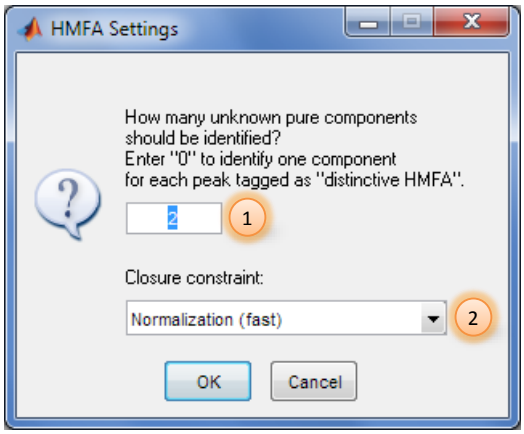
**Factor Indicator Function**
The factor indicator function reaches a minimum where the correct number of factors/components is expected. However, this only occurs when the error is random and fairly uniform throughout the entire data. [E. R. Malinowski, Analytical Chemistry, Vol. 49 (4), April 1977]. In real cases, the function is typically L-shaped for about the first 10 components and the most likely number of components is indicated by a more or less sharp bend in the curve.

## Component Analysis

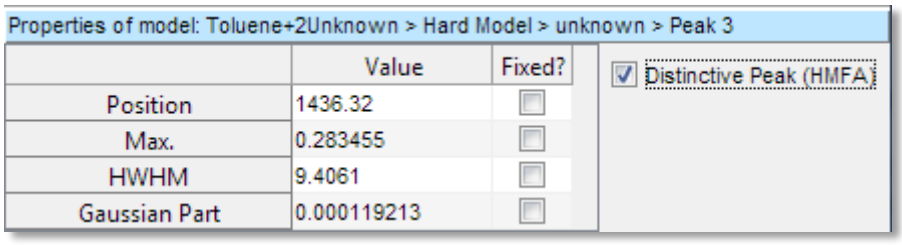
| Menu bar: | Analysis > HMFA > Component Analysis |
|---|---|

Component Analysis consists in the identification of unknown component spectra (if any) and the computation of concentrations for each component. Results are displayed in a Report Generator Window. If any unknown component is to be identified and you close the report generator window with the OK button, a new master model will be added to the Model Tree Panel. It will contain a hard model with all identified components as separate component models.

Additional input is required for this analysis:




Firstly, if the hard model's first component contains peaks, you have to enter the number of unknown component to be identified (1). Alternatively, you could tag some of the model's peaks as "distinctive HMFA" in the Model Properties Panel (see below) and enter "0". Then, a separate component will be identified for each tagged peak.




Secondly, you could choose between three kinds of closure constraints (2) for post-processing calculated component weights:

- none: computed component weights are not processed any further
- Normalization: for each sample separately, computed component weights are normalized such that the sum of weights will be one.
- Optimization: for all samples simultaneously, computed component weights are multiplied by common factors which are found by numerical optimization. Again, the sum of weights will be one for each sample.

## Available Reports

**Component Spectra**
Spectra of all identified pure components (if any) and all initially provided pure components (if any)

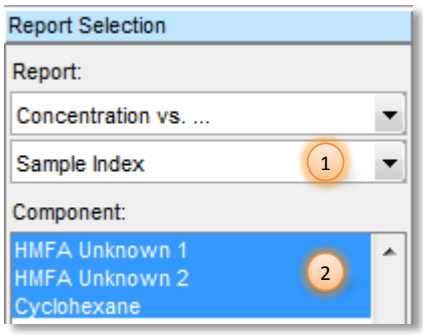
**Concentrations**
Calculated component weights processed according to chosen closure constraint. If no closure constraint is applied, these values are in fact component weights. Otherwise, values represent dimensionless concentrations.

**RMS Spectral Residuals**
Euclidean mean of differences between fitted model and measured samples

**Table: Concentration Profiles**
Numeric values of computed concentrations

**Table: Component Spectra**
Numeric values of pure component spectra

## Report Customization



**X-axis**
By default, concentrations and RMSE values are plotted against the sample index. You can change the x-axis to any available data set feature (1). See Section Data Table Editor for how to add features to data sets.

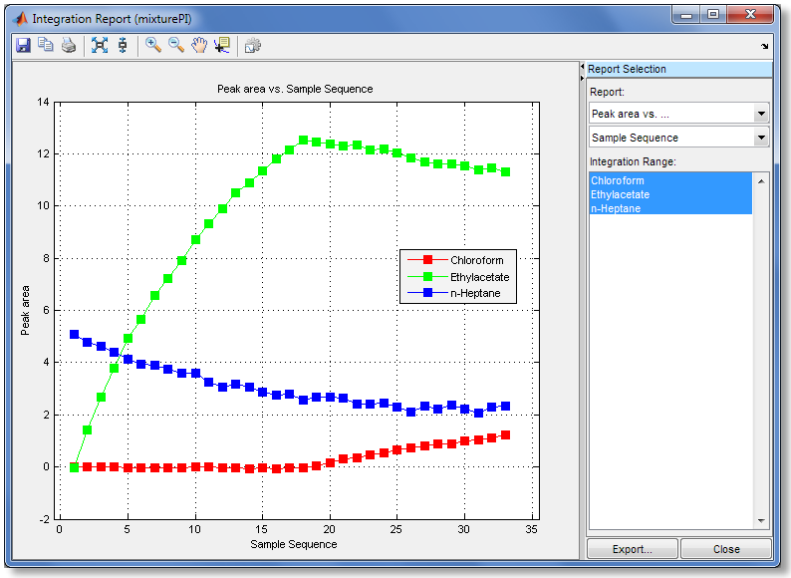
**Component Selection**
By default, plots of all pure component spectra and concentrations are shown in the graphical report. Use the list (2) for selective plotting. Note that RMSE values are independent of the selected component.

### 3.6.3.4 Integration

| Menu bar: | Analysis > Integration |
|---|---|

Integration is applied to selected data sets. A model with integration ranges is required for this analysis.

Integration analysis consists in the calculation of peak areas within each integration range for each sample. Results are displayed in a Report Generator Window.



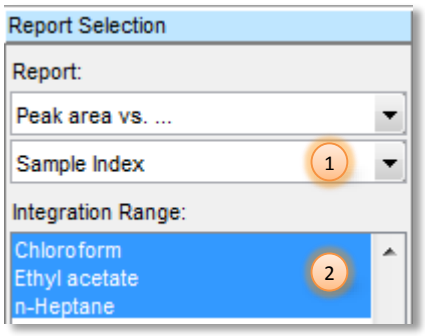

**Integration Report Generator**

## Available Reports

**Peak Area**
Calculated peak areas

**Report Table**
Table with numeric peak areas

### Report Customization




**X-axis**
By default, peak areas are plotted against the sample index. You can change the x-axis to any available data set feature (1). See Section Data Table Editor for how to add features to data sets.

**Integration Range Selection**
By default, plots of all integration ranges are shown in the graphical report. Use the list (2) for selective plotting.

## 3.6.3.5 Component Fitting

| Menu bar: | Analysis > Component Fitting |
|---|---|

Component fitting is applied to selected data sets. A hard model is required for this analysis.

Component fitting analysis consists in the calculation of component weights of the hard model for each sample. Results are displayed in a Report Generator Window.

Component fitting is done by a mathematical procedure called model fitting, in which the model's parameters are automatically adjusted until the model fits a certain measured spectrum. The number of model parameters being adjusted (besides component weights) is controlled by model setting, giving the model more or less flexibility to account for peak variations in the measured spectrum like shifts or shape changes.

### Available Reports

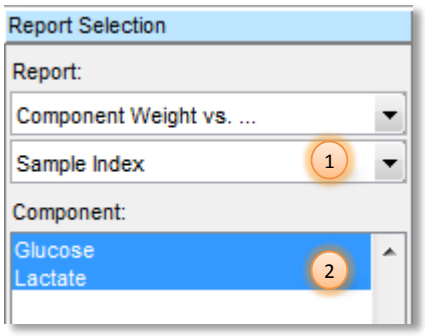
**Component Weight**
Calculated component weights

**RMS Spectral Residuals**
Euclidean mean of differences between fitted model and measured samples

**Table: Component Report**
Table with numeric component weights and RMSE values

**Table: Parameter Report**
Table with numeric values of all parameters of all component fits

### Report Customization




**X-axis**
By default, component weights and RMSE values are plotted against the sample index. You can change the x-axis to any available data set feature (1). See Section Data Table Editor for how to add features to data sets.

**Component Selection**
By default, plots of all components weights are shown in the graphical report. Use the list (2) for selective plotting. Note that RMSE values are independent of the selected component.

## 3.6.3.6 Component Fit Preview

| | |
|---|---|
| Toolbar: | (preview linear component fit) |
| Menu bar: | Analysis > Manage Fits > Preview Linear Fit |
| or | |
| Toolbar: | (preview component fit according to model settings) |
| Menu bar: | Analysis > Manage Fits > Preview Component Fit |

The component fit preview mode can be used to visually inspect the goodness of component fits in the desktop. A hard model is required and a data set must be active in order to enable the preview mode.

You can either display a linear component fit or the actual component fit as specified in the model setting. The former one is just provided for a quick preview; the latter is the important one and is typically referred to when speaking of "component fits".

When component fit preview mode is enabled, the Model Properties Panel displays properties of the component fit instead of the original hard model and the Plot Panel shows the component fit, i.e. the fit of the hard model to the active sample.

Note: Models cannot be modified while component fit preview is enabled.

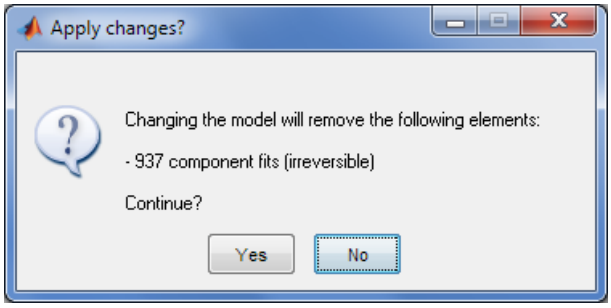
### Managing Component Fits

Component fits will be computed only once and then stored in memory. When enabling preview mode, a new component fit will be computed for the active sample if it does not exist.

Otherwise, a previously computed component fits will be reused. All data sets having component fits available are labeled with [C] in the Data Sets Panel.

Component fits will be removed, when

- modifying the hard model. You will be notified in advance:




- removing fits manually via menu Analysis > Manage Fits > Remove ...
- shutting down PEAXACT. You could save the session in order to preserve all component fits for the next time.

Component fits can be stored permanently, when

- saving the session
- exporting the current component fit as new hard model to a model file via menu Analysis > Manage Fits > Export Visible Fit as New Model...

### 3.6.3.7 Prediction

| Menu bar: | Analysis > Prediction |
|---|---|

Prediction is performed for selected data sets. A calibrated model is required for this analysis.

Prediction analysis consists in the calculation of features for each sample. Results are displayed in a Report Generator Window. Different results may be available depending on the calibration method.

#### Available Reports

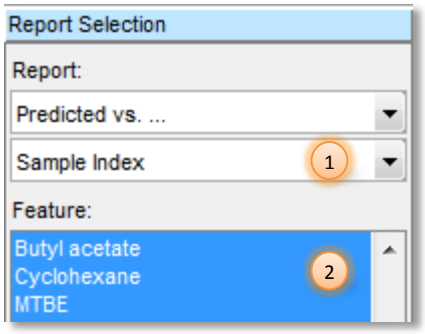
**Predicted Feature**
Calculated features

**RMS Spectral Residuals**
The Root Mean Squared Spectral Residuals measure how well the model explains the spectral data. RMSE values are only shown for data sets for which reference values are provided. See also calibration reports.

**Mahalanobis distance**
The Mahalanobis Distance tells how well a sample matches the training samples used for calibration. The Mahalanobis Distance can be used to spot spectral outliers. See also calibration reports.

**Report Table**
Table with numeric values of predicted features

## Report Customization




**X-axis**
By default, predicted features are plotted against the sample index. You can change the x-axis to any provided data set feature (1). See Section Data Table Editor for how to add features to data sets.

**Feature Selection**
By default, plots of all predicted features are shown in the graphical report. Use the list (2) for selective plotting.

## Command Line Alternative

Prediction analysis can also be started from the command line with options "`-toolbox -predict`" which has advantages over the graphical user interface:

- Simultaneous analysis with multiple models
- Results of all models are combined in a single report file
- Results will not be lost in case of a system crash
- Less memory-intensive
- Suited for analyzing many samples

Prediction analysis from the command line is performed for all loaded data sets using all opened calibrated models. In order to load models and data sets from the command line, you could:

- save a session file in advance (by using the user interface) which contains all your necessary models and data sets; then use command line options `-restore` or `-session <file>` in addition to `-predict`
- load additional files directly from the command line by adding filenames to the command, e.g. `peaxact -toolbox -predict "my model.pxm" c:\data\dataTable.xls`

Results are written to a spreadsheet
`<directory of first opened model>\Reports\<yyyy-mm-dd>-prediction_<#>.csv`
which is located in subdirectory `Reports` of the first opened model file.

> Note: Results are written to file directly after the analysis of a data set, i.e. results won't get lost in case of a system crash.

See also: Toolbox Command Line Options

### 3.6.3.8 Validation

> Menu bar: Analysis > Validation

Validation is performed for selected data sets. A calibrated model is required for this analysis. Selected data sets should (but do not have to) provide values for calibrated features which could be used as reference values for quantitative validation of the model's predictive capabilities.

Validation analysis consists in the calculation of features for each sample and the comparison with given feature values. Results are displayed in a Report Generator Window. Different results may be available depending on the calibration method.

The Validation Report Generator differs from the Prediction Report Generator in two ways:

1) Predicted values are shown independently for each feature and for each location (in case of a location-dependent calibration/prediction), i.e. each calibration model can be validated separately.
2) In addition to predicted values, reports display additional data points such as actual reference values, differences between predicted and actual values, and reference values used during calibration. This is useful for comparing results with the original calibration.

#### Available Reports

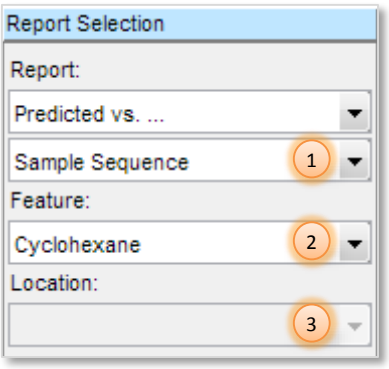
**Predicted vs. True**
Predicted feature values are plotted against the actual values for each data set for which such references are provided. In addition, reference values used during prediction are displayed.

**Predicted vs. ...**
Predicted feature values are plotted against a selectable feature. In addition, differences to the actual values are shown for each data set for which such references are provided.

**RMS Spectral Residuals vs. True**
**RMS Spectral Residuals vs. ...**
The Root Mean Squared Spectral Residuals measure how well the model explains the spectral data. RMSE values are only shown for data sets for which reference values are provided. See also calibration reports.

**Mahalanobis distance vs. True**
**Mahalanobis distance vs. RMS Spectral Residuals**
**Mahalanobis distance vs. ...**
The Mahalanobis Distance tells how well a sample matches the training samples used for calibration. The Mahalanobis Distance can be used to spot spectral outliers. See also calibration reports.

### Report Customization



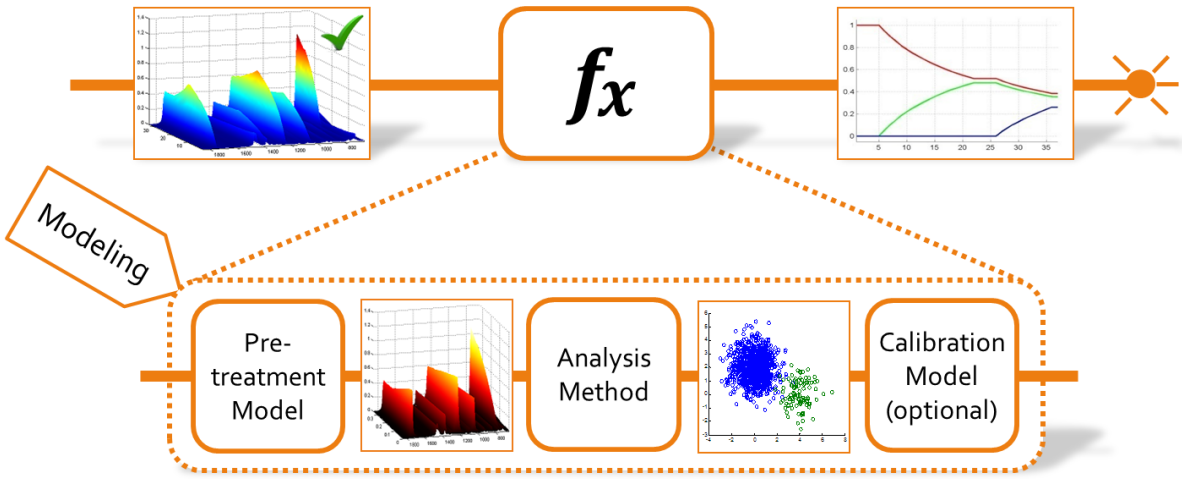

**X-Axis**
For some reports, you can change the x-axis to any provided data set feature (1). See Section Data Table Editor for how to add features to data sets.

**Feature Selection**
Use the pull-down menu (2) to display results for individual calibrated features.

**Location**
Use the pull-down menu (3) to display results for individual locations.

# 3.7 Modeling

Modeling is the process of creating a model for data analysis. It is subject to the specifics of a chosen analysis method. For instance, modeling for peak integration comprises the definition of integration ranges, while Indirect Hard Modeling comprises the creation of a hard model. Modeling also deals with data pretreatment and calibration.




**Modeling steps**

This Section explains modeling functionalities of PEAXACT. See Section 3.6.3 to learn more about model-based analysis.

## Nomenclature

**Model**

A mathematical model explains / describes / models data by means of functions and parameters

**Modeling**

The process of developing a model by choosing functions and initial parameters; functions are typically given when deciding on a certain analysis method.

**Analysis method**

A set of predefined functions for a certain analysis purpose

**Model-based analysis**

Applying the model to unknown data with the goal of determining parameters of interest

**Variable**

A model parameter which varies between samples (in contrast to a constant parameter); some variables are of interest for analysis because they correspond to / are related to the quantity of interest.

**Master Model**

A term used in PEAXACT for a model that consists of sub-models:

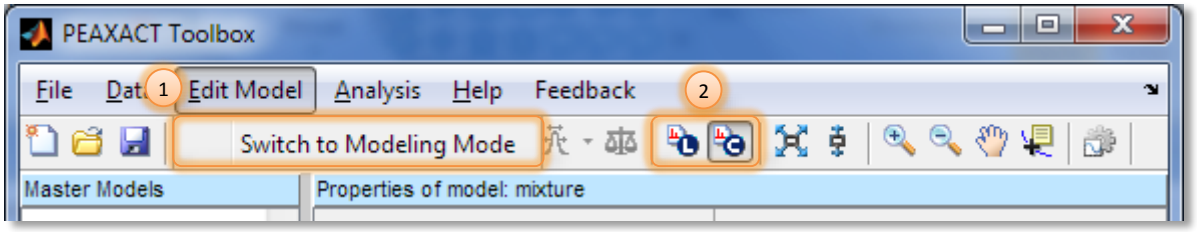
- Pretreatment Model: contains parameters for data manipulation
- Integration Model: contains parameters for peak integration
- Hard Model: contains parameters for peak fitting and component fitting
- Calibration Model: contains regression parameters

### Modeling Mode

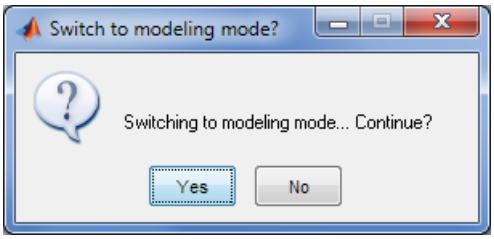
The desktop can be in either one of two states: the modeling mode or the component fit preview mode. The default mode is modeling mode. It is automatically enabled if no other mode is active. Component fit preview mode can be enabled as part of the Component Fitting Analysis.

> Note: Editing models is only possible in modeling mode.

If you encounter problems when trying to edit the model you may not be in modeling mode. In this case you can switch to modeling mode via **Edit Model > Switch to Modeling Mode** (1) or by actively deactivating the component fit preview mode (2).




Some actions automatically ask you to switch to modeling mode.




## 3.7.1 Undo & Redo

| | |
|---|---|
| Menu bar: | Edit Model > Undo / Redo |
| Hot key: | CTRL + Z / CTRL + Y |

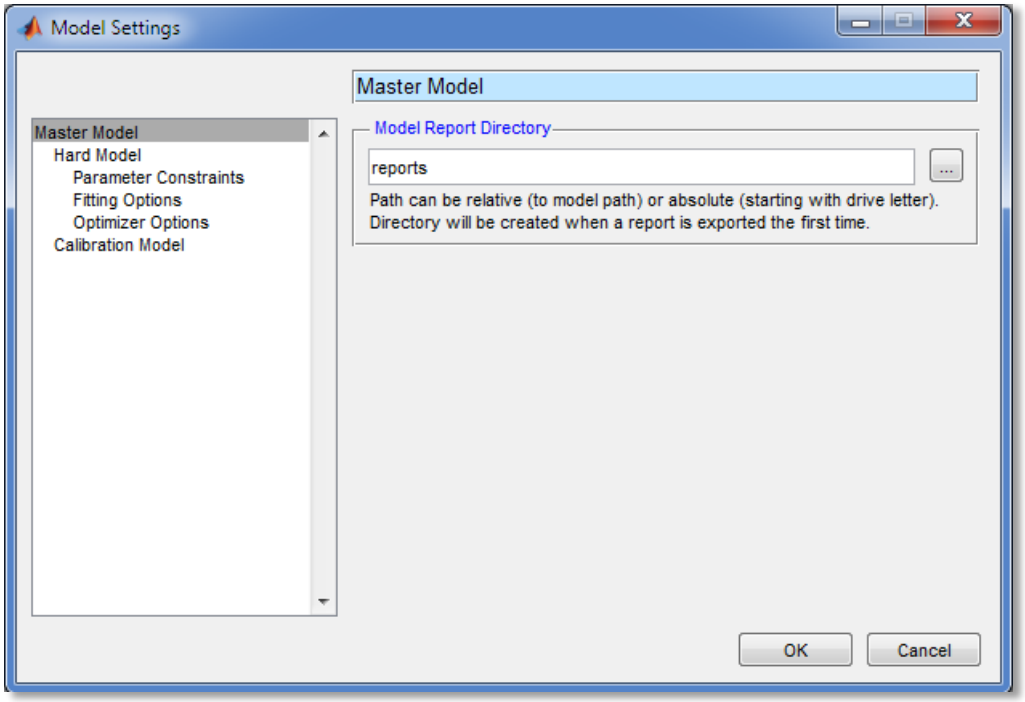
Each model keeps a history of its most recent changes. You can undo and redo these modifications step-by-step. The number of undo-steps is controlled by a user preference.

Undo/Redo is model-dependent, i.e. each master model has its own history.

## 3.7.2 Model-Specific Settings

| | |
|---|---|
| Toolbar |  |
| Menu bar: | Edit Model > Model Settings... |
| Context menu: | Model Tree > Master Model Item > Settings... |
| Properties Panel: | Click Model Tree > Master Model Item |

The Model Settings Editor looks similar to the User Preferences Editor. This is because model-specific settings originate from user preferences. Once copied from the user preferences profile, model settings can be customized independently from user preferences and separately for each model, i.e. each model contains its own individual settings. For a description of all model settings, see Section 3.3.3.




**Model Settings Editor**

## 3.7.3 Pretreatment Model

The data pretreatment model is a sub-model of the master model. It consists of a set of data manipulations which are applied to sample xy-data when the master model is active. Pretreatments defined by different master models never interfere because data are only manipulated temporarily as long the corresponding master model stays active.

For a description of all available pretreatment settings see Section 3.3.3.2.

### Modification

Properties Panel: Click Model Tree > Data Pretreatment Item

The pretreatment model can be modified by using the Pretreatment Model Properties Panel.

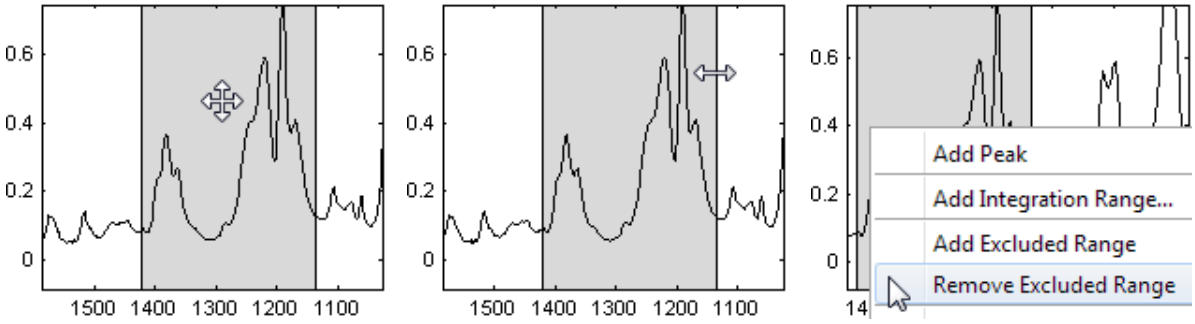
### Reset

Context Menu: Model Tree > Data Pretreatment Item > Reset

The pretreatment model can be reset to the active user preferences profile.

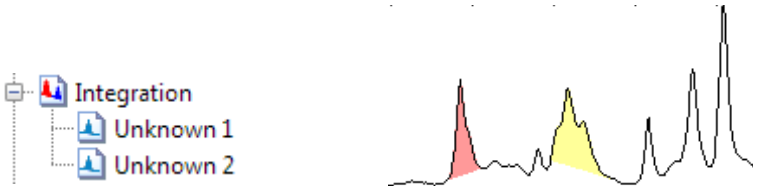
### Interactive Modification of Excluded Ranges

Interactive modification of excluded ranges in the Plot Panel is controlled by a user preference. If enabled, you can use the mouse to modify the bounds of excluded ranges. Drag the gray area with the left mouse button to shift the whole excluded range. Drag the boundary line to modify the boundaries. Right-click the area to open the context menu for more options.




## 3.7.4 Integration Model

The integration model is a sub-model of the master model. It consists of one or more integration ranges, each specifying parameters for calculating the signal area within a certain x-range.



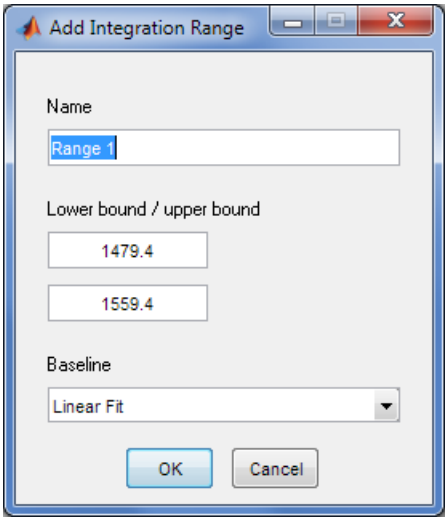

**Integration model displayed in Model Tree Panel (left) and Plot Panel (right)**

Integration models can be utilized for Integration Analysis and Peak Integration Prediction.

### 3.7.4.1 Add Integration Range

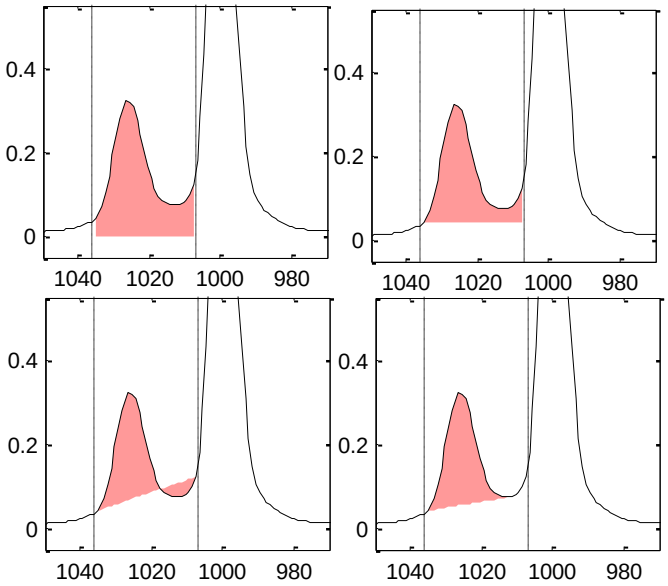
| | |
|---|---|
| Toolbar: |  |
| Menu bar: | Edit > Add Integration Range... |
| Context menu: | Model Tree > Integration > Add Range |
| Context menu: | Plot Panel > Add Integration Range... |

A new dialog appears in order to specify properties of the new integration range such as its name, its bounds, and the type of baseline.




The integration range should be named after the component whose signal is to be integrated. The name is also utilized during calibration. Lower bound and upper bound specify integration x-limits; baseline specifies a y-limit for the integrated area. You can choose from 4 baselines types:

- None: Baseline is zero; integration result may be negative if signal is negative
- Offset: Baseline is the smallest y-value within x-limits; integration result is positive
- Straight Line: Baseline is a straight line from lower to upper bound; integration result might be negative
- Linear Fit: Baseline is a straight line fitted to signal; integration result is positive




**Baseline types None and Offset (top), Straight Line and Linear Fit (bottom)**

Clicking OK adds a new integration range to sub-model "Integration" of the active master model and refreshes the Plot Panel accordingly.

### 3.7.4.2 Modify Integration Range

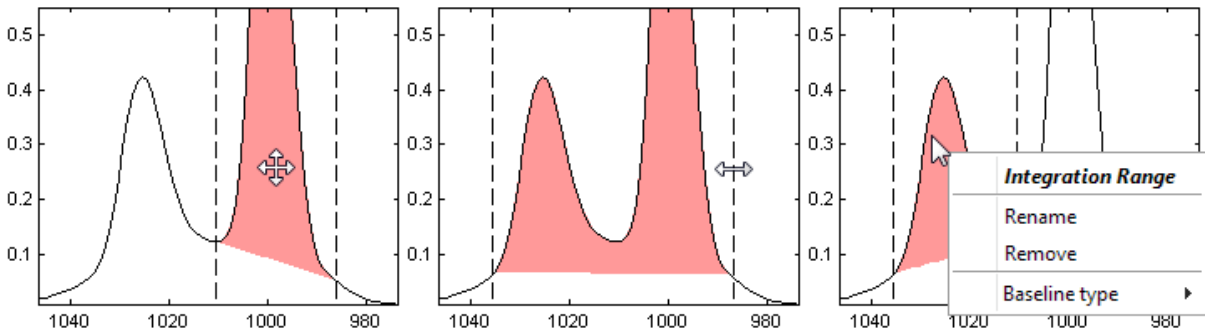
| | |
|---|---|
| Menu bar: | Edit Model > Rename Integration Range... |
| Context menu: | Model Tree > Integration Range Item |
| Context menu: | Plot Panel > Integration Range Patch |
| Properties Panel: | Select Model Tree > Integration Model Item |
| Properties Panel: | Select Model Tree > Integration Range Item |
| Plot Panel: | Interactive modification of bounds |

Integration ranges can be modified in several ways:

- From the menu bar you can rename the selected integration range.
- From the context menu (either in the Model Tree or in the Plot Panel) you can rename the integration range and change the baseline type.
- The Properties Panel can be used to modify all integration range properties (name, bounds, and baseline type).

#### Interactive Modification

In the Plot Panel you can use the mouse to interactively modify the bounds of an integration range. Drag the colored area with the left mouse button to shift the whole integration range. Drag the dashed line to modify the boundaries. Right-click the area to open the context menu for more options.




### 3.7.4.3 Remove Integration Range

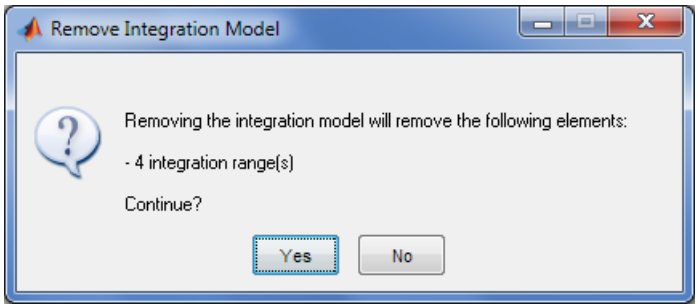
| | |
|---|---|
| Menu bar: | Remove Integration range |
| Context menu: | Model Tree > Integration Range Item> Remove |
| Context menu: | Plot Panel > Integration Range Patch> Remove |
| Keyboard: | Delete Key |

From the menu bar and from the context menu (either in the Model Tree Panel or in the Plot Panel) you can remove the selected integration range. You can also select an integration range item in the Model Tree Panel or its graphical representing in the Plot Panel and press the delete key.

### Remove All Integration Ranges

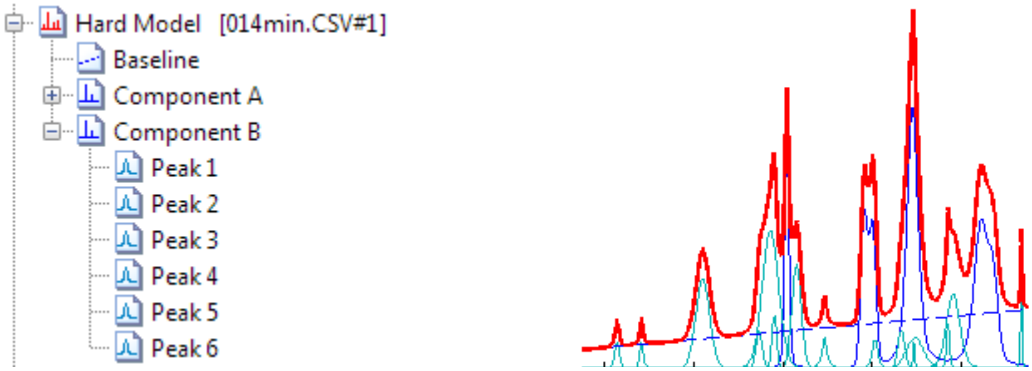
Context menu: Model Tree > Integration Model Item > Remove

You can remove all integration ranges at once by removing the full integration model.




## 3.7.5 Hard Model

The hard model is a sub-model of the master model. The hard model is a mathematical representation of a sample. The model contains parameters which are to be determined during modeling.




**Hard model displayed in Model Tree Panel (left) and Plot Panel (right)**

Hard models can be utilized for Hard Modeling Factor Analysis, Component Fitting Analysis, and Indirect Hard Modeling Prediction.

### Nomenclature

When working with hard models, you should be familiar with the following wording:

**Hard Model**

A hard model in general is a mathematical function which is derived from equations representing the physics behind an underlying process. In PEAXACT, the term hard model refers to a function derived from the physics of molecular spectroscopy; it is a mathematical representation of a mixture spectrum.

Physics tells us that a mixture spectrum is composed of superimposed peaks originating from the individual components in the mixture, with the components' con-

centrations being responsible for the peaks' intensities. This structural information is maintained in the hard model by means of a sum of peak-shaped curves. Those subsets of peak-curves that represent pure component spectra are referred to as component models; they are multiplied by concentration-related weight parameters.

In summary, the hard model consists of weighted component models which again consist of peak models. The hard model is completed by a simple baseline model.

**Component Model**

The component model is a mathematical representation of a pure component spectrum. It consists of peak models.

**Peak Model**

The peak model is a mathematical function with a peak-shaped profile. PEAXACT uses pseudo-Voigt functions. A pseudo-Voigt function is a linear combination of Gaussian and Lorentzian functions.

**Baseline Model**

The baseline model is a linear or quadratic polynomial.

**Component Weight**

The component weight is a parameter of the hard model. The hard model is a weighted sum of component models, i.e. each component model is multiplied with its component weight. Component weights correspond to concentrations; they are relevant parameters for calibration.

**Peak Parameters**

Peak parameters are the parameters of peak models, e.g. the position or width. These parameters are adjusted when fitting the hard model to measured samples in order to correct peak variations.

**Editable Component & Editable Peaks**

The first component model of the hard model is called "editable component" because peak models can only be added, removed, or modified in this component. Accordingly, peaks of the editable component are called "editable peaks".

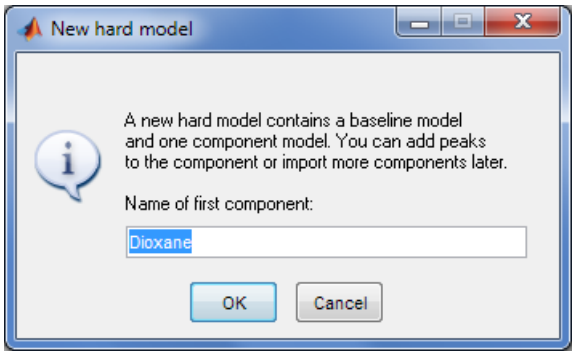
**Model Fitting, Peak Fitting & Component Fitting**

Model fitting is a general expression for a mathematical procedure in which the model's parameters are automatically adjusted until the model fits a certain measured sample. Peak fitting and component fitting are special kinds of model fitting which differ in the parameters being adjusted. Peak fitting refers to the adjustment of peak parameters of only editable peaks. Component fitting mainly refers to the adjustment of component weights. Errors in component weights can be reduced by simultaneously adjusting peak parameters to account for peak variations in the measured sample. The extent of peak adjustments during component fitting is controlled by model settings.

### 3.7.5.1 New Hard Model

| | |
|---|---|
| Menu bar: | Edit Model > New Hard Model |
| Context menu: | Model Tree > Hard Model Item > New |

A new dialog appears in order to specify the name of the first component model. The first component of each hard model called editable component because new peak models can only be added, removed, and modified for this component.




#### Name of Editable Component

If it is your intention to add peaks to the editable component, you should name it after the pure component to be modeled. If, however, you do not want to add peaks because you only intend to import other component models, you should name the editable component like "empty" or "-". You can always rename components later.

If a data set is active while the new hard model is being created, the name of the editable component is suggested to be the data set name.

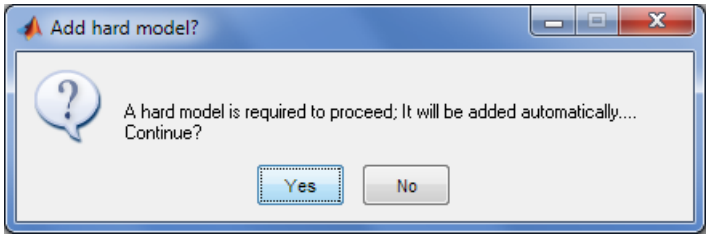
#### New Hard Model for Active Sample

| | |
|---|---|
| Context menu: | Data Sets Panel > New Hard Model for Active Sample |

When creating a hard model from the context menu of the Data Sets Panel the first component is automatically named after the active data set. If the currently active master model already has a hard model, or if no master model is active at all, a new master model is created first.

#### New Hard Model if Required

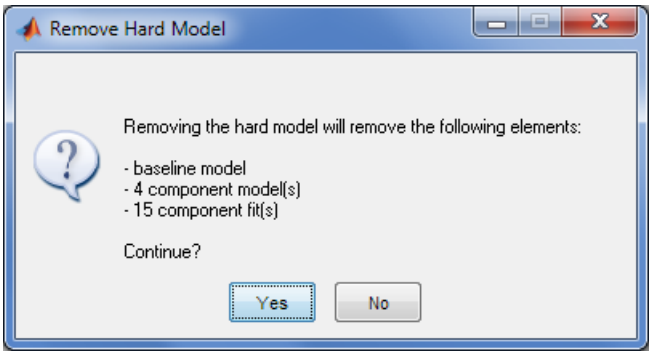
Some hard modeling actions automatically prompt you to create a new hard model if none exists.




### 3.7.5.2 Remove Hard Model

| | |
|---|---|
| Context menu: | Model Tree > Hard Model Item > Remove |

When removing a hard model, all its dependent models will be removed, too, namely the baseline model, component models, and also components fits of the hard model to data sets.




### 3.7.5.3 Rename Components

| | |
|---|---|
| Menu bar: | Edit Model > Rename Component |
| Context menu: | Model Tree > Component Model Item > Rename |
| Properties Panel: | Click Model Tree > Hard Model Item |

From the menu bar or the context menu of the Model Tree Panel you can rename the active component. In addition, when activating the Hard Model item in the Model Tree Panel, you can rename components in the Model Properties Panel.

Component models can get any names as long as they are unique within each hard model.

### 3.7.5.4 Remove Components

| | |
|---|---|
| Menu bar: | Edit Model > Remove Component |
| Context menu: | Model Tree > Component Model Item > Remove |
| Keyboard: | Delete Key |

From the menu bar and from the context menu of the Model Tree Panel you can remove the active component model. You can also select a component model item in the Model Tree Panel or its graphical representing in the Plot Panel and press the delete key.

Removing component models also removes the following dependent elements:

- calibration model (IHM)
- component fits

Note: You cannot remove the editable component.

### 3.7.5.5 Import Components

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | Edit Model > Import Component(s)... |
| Context menu: | Model Tree > Hard Model Item > Import Component(s)... |
| Drag & Drop: | from Explorer to Model Tree Panel |

The File Dialog can be used to browse directories and select one or more model files to import component models from. Choose a file filter from the pull-down list to localize certain model files. A master model must be active in order to import components.

All non-empty components of the source models are copied to the active model. If no hard model exists yet, you will be prompted to create one first.

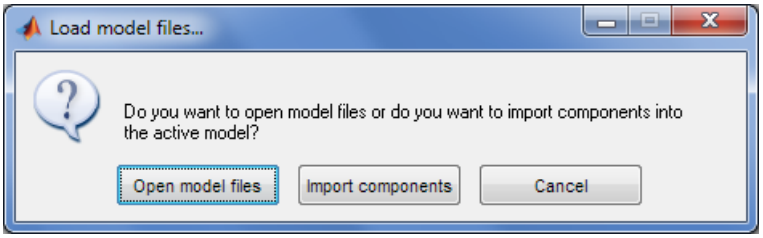
Names of imported components that already exist will automatically be numbered consecutively.

When importing components, the following elements become obsolete and are removed:

- calibration model (IHM)
- component fits

#### Drag & Drop

A fast way to import component models is by using drag & drop. Activate the target model in the Model Tree Panel, then drag model files from the Windows Explorer (e.g.) and drop them into the Model Tree Panel. You will be asked whether to open these files or to import components into the active model.




#### Component Equilibration

Imported component models are equilibrated, which means that

- component weights are set to 1, and
- peak amplitudes of each component are scaled such that the maximum peak amplitude is in the same order of magnitude as the active sample (or 1 if no sample is active)

Component equilibration is reasonable because otherwise parameters of different component models could have totally different orders of magnitude.

> Caution: Do not import components while a sample is active which is completely unrelated to the hard model.

### 3.7.5.6 Export Components

| | |
|---|---|
| Menu bar: | Edit Model > Export Component... |
| Context menu: | Model Tree > Component Model Item > Export... |

Each component model can be exported to a new model file. The new master model will contain a hard model with just the one component model. In addition, the data pretreatment model is copied to the new model, too. The File Dialog can be used to browse directories and select or enter a filename for the model to be saved. By default, the component name is suggested as filename.

Exporting of components is not required if the hard model only contains a single component. In that case, simply save the master model.

#### Overwriting Existing Files

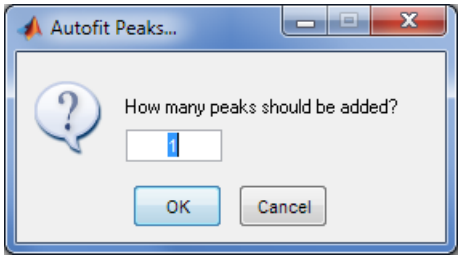
As a precaution, you cannot save the model to a file of another model that is currently opened in PEAXACT. You have to close the other model before overwriting its file.

### 3.7.5.7 Auto-fit Peaks

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | Edit Model > Autofit Peaks... |
| Context menu: | Model Tree > Hard Model Item > Autofit Peaks... |
| Context menu: | Model Tree > First Component Model Item > Autofit Peaks... |

Auto-fitting of peaks is a combination of adding new peaks to the editable component and automatically adjusting editable peaks such that the model fits the active sample.

You are prompted to enter the number of new peaks. Start with a rough count of peaks in the sample and add more peaks later if required.




Note: New peaks will always be added to the editable component.

The peak fitting algorithm is controlled by model settings. In general, first the position of the next peak to be added is determined by analyzing the spectral residuals, then a new peak model is added to the editable component, and finally the model is fitted to the active sample according to the peak fitting procedure.

When fitting new peaks, the following elements become obsolete and are removed:

- calibration model (IHM)
- component fits

### 3.7.5.8 Adjust Parameters Manually

Manual modification of hard model parameters becomes reasonable when automatic parameter adjustment does not work as expected. Then, minor manual modifications could remedy these modeling problems. However, manual adjustments should always be followed by automatic parameter fitting in order to further improve the model.

There are several ways to modify hard models manually. In each case the following elements become obsolete and are removed:

- calibration model (IHM)
- component fits

#### Changing Parameter Values

The following parameter values can be modified manually using the Model Tree Panel:

- all baseline parameters
- peak parameters of the editable component

The following parameters cannot be modified manually:

- component parameters (component shift and component weight)
- peak parameters of other components than the editable component

Select the hard model item in the Model Tree Panel or its graphical representation in the Plot Panel in order to display the element's parameters in the Model Properties Panel. Use the table in the Properties Panel to change parameter values.

## Fixing Parameters

All model parameters but component weights can be fixed, which means that parameter values are not allowed to get changed during model fitting. Still, parameters are allowed to get changed manually.

Select the model item in the Model Tree Panel or its graphical representation in the Plot Panel in order to display the element's parameters in the Model Properties Panel. Use the check boxes in the Properties Panel to fix or free parameter values. Baseline parameters and peak parameters can also be fixed all at once from the element's context menu.

> Important note: The position of peak models is influenced by two parameters, the peak position parameter and the component shift parameter. While the former is individual for each peak, the latter is the same for all peaks of a component. Fix both parameters if you do not want peak models to get shifted.

Fixing parameters is useful for reducing model complexity (because less parameters have to be adjusted during model fitting), but at the same time model flexibility is reduced, too, because less peak variations in the sample can be explained by the model.

## Adding Peaks

| | |
|---|---|
| Context menu: | Plot Panel > Add Peak |

To add a peak model manually, right-click at the desired peak position in the Plot Panel and select **Add Peak** from the context menu. The new peak model has its position and maximum at the clicked point and gets some suggested values for the other peak parameters.

> Note: New peaks will always be added to the editable component.

## Removing Peaks

| | |
|---|---|
| Menu bar: | Edit Model > Remove Selected Peak(s) |
| Context menu: | Model Tree > Peak Model Item > Remove Selected Peak(s) |
| Context menu: | Plot Panel > Peak Plot > Remove Selected Peak(s) |
| Keyboard: | Delete Key |

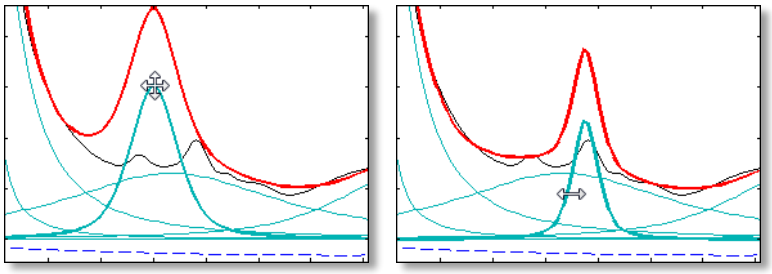
You can remove currently selected peak models from the hard model. Select a peak model item in the Model Tree Panel or its graphical representing in the Plot Panel first. In the Plot Panel you could even select multiple peaks while the SHIFT-key is pressed.

Only peaks of the editable component can be removed.

## Modifying Peaks Interactively

Peaks models can be modified interactively by using the mouse:

- Height and position: Grab the top of an editable peak plot and move the mouse up / down to change the height and move it to the right / left to change the position
- Width: Grab one side of an editable peak plot and move the mouse to the right / left



Caution: You should always respect specified constraints for the peak width parameter when modifying peaks manually. Otherwise, model fitting might not work as expected or might not work at all.

## Grouping Peaks

| | |
|---|---|
| Context menu: | Plot Panel > Peak Plot > Peak Grouping > Group / Ungroup |

Grouped peaks behave like a single new peak with new shape. Select two or more peak plots in the Plot Panel (left-click on the plots while the SHIFT-key is pressed) and then right-click on one of the plots to open the context menu. From the context menu you could either add all selected peaks to a new group or release peaks from the group. When clicking on a grouped peak in the Plot Panel, the whole group is selected. When clicking on a grouped peak in the Model Tree Panel, only the clicked peak is selected.

Because of the following impacts, peak grouping is relevant for component fitting only:

- The intensity of grouped peaks (peak area) is kept constant as a sum, but not for each peak individually (see also: Fitting Options).
- If a parameter (e.g. position) of any peak is considered for automatic adjustment, the same kind of parameter of all peaks in the group will be fitted, too.

Peak grouping is particularly useful to model asymmetric measured peaks by multiple symmetric peak models.

Caution: Peak grouping should only be used for fine tuning the model.

## Switching Baseline Polynomial between Linear and Quadratic

| | |
|---|---|
| Menu bar: | Edit Model > Model Settings > Hard Model > Baseline order |
| Context menu: | Model Tree > Baseline Model Item > Quadratic Baseline |
| Context menu: | Plot Panel > Baseline Plot > Quadratic Baseline |

The baseline model can be a linear or quadratic polynomial. The linear baseline has two parameters (offset and slope), the quadratic baseline has three parameters (offset, slope, and quadratic factor).

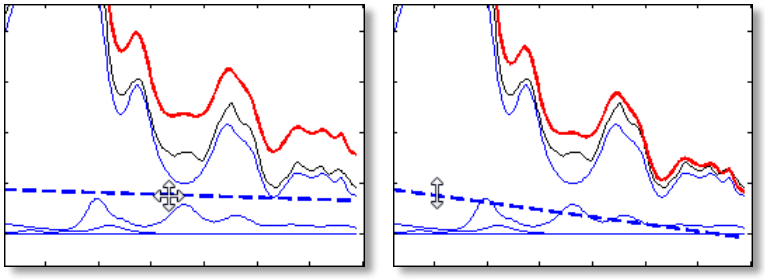
The baseline model should be choses as simple as possible because it is partly redundant to baseline corrections specified in the data pretreatment model.

> Caution: Consider switching to "quadratic" as a last resort. Try to work around the issue by choosing different data pretreatments such as range selection and baseline correction.

### Modifying Baseline Interactively

If the baseline model is displayed as a separate plot in the Plot Panel (see user preferences), it can be modified interactively by using the mouse. In case of a linear baseline model:

- Offset: Grab the center of the baseline plot and move the mouse up / down
- Slope: Grab the left or right fifth of the baseline plot and move the mouse up / down



In case of a quadratic baseline model, baseline parameters are determined by means of three points the baseline goes through. You can adjust the baseline parameters by moving these three points:

- Center point: Grab the center of the baseline plot and move the mouse in any direction
- Leftmost point, rightmost point: Grab the left or right fifth of the baseline plot and move the mouse up / down

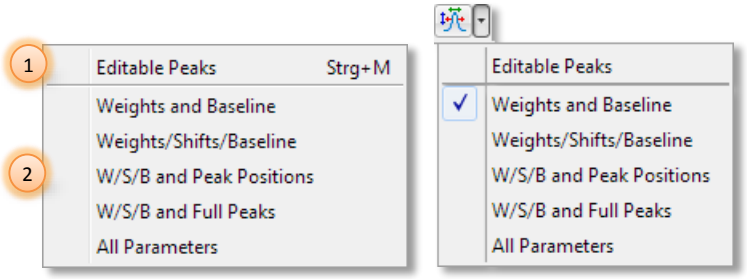
## 3.7.5.9 Adjust Parameters Automatically

| | |
|---|---|
| Toolbar: |  |
| Menu bar: | Edit Model > Adjust Hard Model > ... |

The model's parameter can be adjusted automatically such that the model fits a certain measured spectrum. Model fitting is an important step because the hard model is intended to represent a mixture spectrum well. A good fit will provide good initial parameter values for subsequent analyses.

There are two kinds of model fitting: peak fitting and component fitting. They differ in parameter constraints being applied and in the set of parameters being adjusted. From the

menu bar (left figure) or from the toolbar (right figure) you can choose these different parameter sets:




**Parameter sets of peak fitting (1) and component fitting (2)**

When using the toolbar, the parameter set can be changed by clicking on the arrow next to the icon. The most recently used parameter set is remembered and is used when clicking the toolbar icon.

## Peak Fitting

Peak fitting is suited for creating new hard models from the scratch. Basically, it is an unconstrained fit (except for bounds on peak width and shape parameters) of editable peaks and therefore, it is very qualified when major parameter adjustments are required, e.g. when adding new peaks.

In detail, peak fitting covers:

- adjustment of all parameters of all editable peaks
- adjustment of all baseline parameters
- adjustment of component weights of all components but the editable component (if any)

## Component Fitting

Component fitting is suited for fine tuning parameters of an existing hard model. Basically, it is a fit of component weights but other parameters can be adjusted, too, according to the chosen parameter set. These additional parameters are fitted with respect to tight constraints. In particular, peak parameters can only vary with respect to constant peak intensities (peak areas). Because of these constraints, component fitting is only qualified for moderate parameter adjustments.

> Note: Component fitting is also performed during some analyses, but while the analysis never changes the hard model but just evaluates fitted parameters, during modeling the hard model is actually changed in that adjusted parameter values become new initial values.

Adjusting model parameters by means of component fitting is not necessarily required at all, but it can be reasonable in order to get better initial parameter values.

### Repeated Model Fitting

When conducting peak fitting or component fitting repeatedly it is likely you will see the fit improving slightly each time. There are several reasons for this, e.g.:

- Stopping criteria for model fitting are measured relative to the starting point. The starting point of each new fit is the previous fit.
- Relative parameter bounds might be reached during one fit. Each new fit has new bounds.

Caution: Repeated component fitting might significantly deform peaks. This could even result in poorer initial parameter values for component fits to be performed during subsequent analysis.

## 3.7.6 Calibration Model

The calibration model is a sub-model of the master model. It is a functional relationship between values calculated from samples (e.g. component weights) and predictive features (e.g. concentrations). The functional relationship is established by regression.

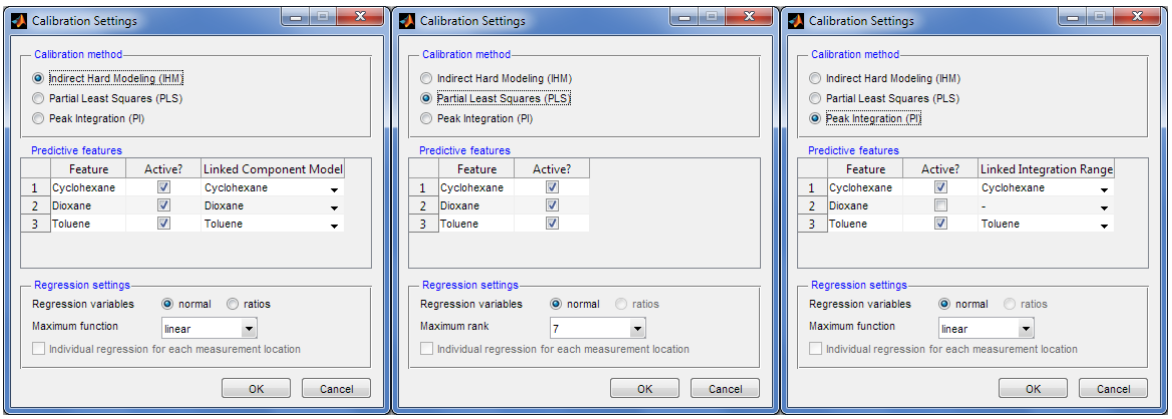
### 3.7.6.1 New Calibration Model

| | |
|---|---|
| Toolbar |  |
| Menu bar: | Edit Model > Calibration... |
| Context menu: | Model Tree > Calibration Model Item > Calibrate... |

Calibration is performed for selected data sets. All selected data sets must provide a common set of predictive features. Data sets with usage set to "train" or "unspecified" will be used for training the calibration model. Data sets with usage set to "test" will be used for validating the calibration model. Data sets with usage set to "ignore" will not be considered.

Calibration begins with the specification of settings. Results will then be displayed in a Report Generator Window. Finally, you have to choose and accept a specific calibration model.

## Calibration Setup




**Calibration Settings Dialog for calibration methods IHM, PLS, and PI**

The Calibration Settings Dialog contains three panes for three kinds of decisions you have to make:

1) Which calibration method should be used?
2) Which predictive features should be calibrated?
3) Which settings should be used for regression of the calibration model?

### Calibration Method

You can choose between three different calibration methods. Depending on your choice, the Calibration Settings Dialog changes its appearance to provide proper controls for each method.

- Indirect Hard Modeling (IHM): available if the master model contains a hard model
- Partial Least Squares (PLS): always available
- Peak Integration (PI): available if the master model contains an integration model

### Feature Selection

The table shows a list of available predictive features. Use the checkboxes to activate/deactivate features for calibration.

For IHM and PI, each active feature must be linked with a component model or integration range, respectively, in order to provide a connection between features and model parameters such as component weights or peak areas.

### Regression Settings

First select the kind of regression variables:

- normal: available for all calibration methods; regression of the calibration model is performed separately for each feature
- ratios: available for IHM and PI if the sum of active feature values is the same constant (closure constraint) for each data set; regression of the calibration model is done for pairwise ratios of features and by incorporating a closure constant.

Then, choose the maximum polynomial degree of the regression function (IHM and PI) or the maximum rank (PLS). A separate model will be calibrated for each function / rank up to the selected maximum.

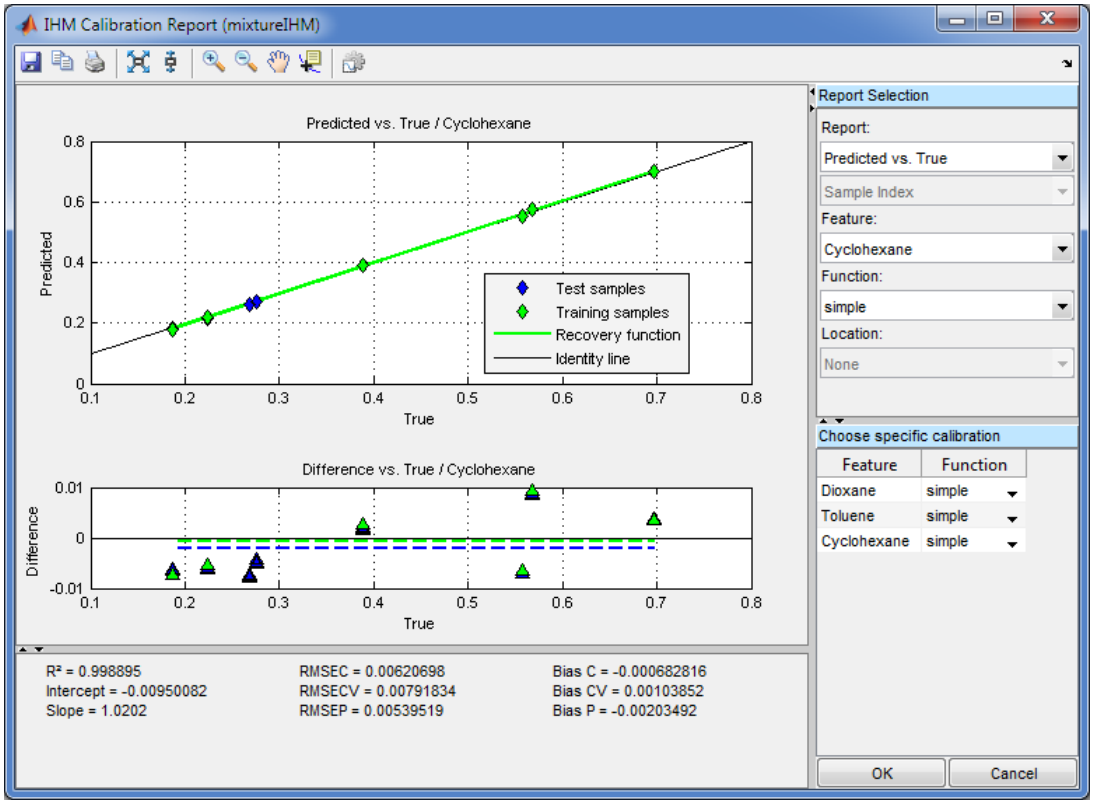
- Regression function: the regression function is used by univariate calibration methods (IHM, PI) to establish the functional relationship between component weights / peak areas and predictive features. A special function "simple" can be chosen; it is a linear function without intercept.
- Rank: the rank is used by multivariate calibration methods (PLS). It is the number of multivariate factors used to explain the variance of the data.

If the "location" feature is provided for data sets, you could optionally enable "individual regression for each measurement location". In this case, a separate model will be calibrated for each location.

Finally, start calibration by clicking the OK button.

## Calibration Results

Results are displayed in a Report Generator Window. Different results may be available depending on the calibration method.




**Calibration Report Generator Window**

## Available Reports

### Predicted vs. True

Predicted feature values $x_{predicted}$ are plotted against the actual values $x_{true}$ for each sample. Samples are color-coded by their usage (green = training samples, blue = test samples).

The plot can be used determine the predictive accuracy of the model and to spot feature outliers.

$x_{predicted}$ are calculated from samples $Y$ using the calibration model $f(Y, K)$

$$x_{predicted} = f(Y, K)$$

The structure of function $f(Y, K)$ depends on the calibration method (IHIM, PI, PLS) and on the regression settings.

A best-fit line though the training points – the recovery function – is added to the chart.

$$x_{recovered} = intercept + slope \cdot x_{true}$$

In case of a perfect calibration, intercept and slope of the recovery function would be zero and one, respectively, i.e. it would match the identity line. The deviation from the identity line can be used as an indicator for the predictive capability of the calibration model.

**Differences vs. True**

Differences between predicted and actual values are plotted for each sample.

$$x_{difference} = x_{predicted} - x_{true}$$

Typically, differences will be distributed around zero. The plot can be used to determine shifts and drifts in the calibration by analyzing the distribution pattern.

A dashed line – the average bias – is added to the chart. The average bias is the mean of the differences.

$$bias = \frac{1}{N}\sum_{i}^{N} x_{difference,i}$$

The average bias is shown for training samples (bias C, green) and test samples (bias P, blue). It can be interpreted as the systematic error of prediction, i.e. predictions are off by this value on average.

**Predicted vs. …**

Predicted feature values are plotted against a selectable feature.

**R² vs. Function / Rank**

R² is the fraction of variance in $x$ which is explained by the calibration model. A value of 1 corresponds to 100%. The plot gives an overview of R² for different regression functions / ranks.

$$R^2 = 1 - \sum_{i}^{N} \frac{\left(x_{predicted,i} - x_{true,i}\right)^2}{\left(x_{true,i} - \bar{x}_{true}\right)^2}$$

$x$ denotes feature values (e.g. concentrations) of training samples.
$\bar{x}_{true}$ is the mean $x_{true}$ of all training samples.
$N$ is the number of training samples.

R² could be negative if the calibration model predicts features worse than a constant function would do. Obviously, such models should be discarded.

**RMSEC vs. Function / Rank**
RMSEC is the Root Mean Square Error of Calibration. It is the average deviation between predicted and actual feature values of the training samples.

$$RMSEC = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(x_{predicted,i} - x_{true,i}\right)^2}$$

**RMSECV vs. Function / Rank**
RMSECV is the Root Mean Square Error of Cross-Validation. It is the average deviation between predicted and actual feature values when doing a cross-validation of training samples. Use this report to determine the regression function / rank with the least error if no independent test samples are available.

$$RMSECV = \sqrt{\frac{1}{N}\sum_{i}^{n}\left(x_{predicted\ CV,i} - x_{true,i}\right)^2}$$

$x_{predicted\ CV,i}$ is calculated internally and is not displayed in any report.

**RMSEP vs. Function / Rank**
RMSEP is the Root Mean Square Error of Prediction. It is the average deviation between predicted and actual feature when doing a test-set validation of independent test samples. Use this report to determine the regression function / rank with the least error.

$$RMSEP = \sqrt{\frac{1}{N'}\sum_{i}^{N'}\left(x'_{predicted,i} - x'_{true,i}\right)^2}$$

$x'$ denotes feature values of independent test samples.
$N'$ is the number of test samples.

**RMS Spectral Residuals vs. True**
**RMS Spectral Residuals vs. …**
The Root Mean Squared Spectral Residuals measure how well the model explains the spectral data. It can be used to spot spectral outliers. The smaller the value the better the model explains the spectral signal.

$$RMS\ Res = \sqrt{\frac{1}{P}\sum_{i}^{P}\left(y_{reconstructed,i} - y_{measured,i}\right)^2}$$

$y$ denotes spectral intensities.
$y_{reconstructed}$ are spectral intensities reconstructed by the model. For IHM this is the fitted hard model; for PLS it is the reconstruction using $R$ PLS loadings. Spectral residuals are not available for PI models.
$P$ is the number of spectral data points.
$R$ is the rank of the PLS model.

**Mahalanobis distance vs. True**
**Mahalanobis distance vs. RMS Spectral Residuals**
**Mahalanobis distance vs. …**
The Mahalanobis Distance tells how similar a sample is to the set of training samples and can be used to spot spectral outliers. It is a measure of the multivariate distance from the centroid comprised of the PLS scores of all training samples. The Mahalanobis distance is not available for IHM and PI models.

$$MD_i = \sqrt{s_i \cdot (S^T \cdot S)^{-1} \cdot s_i^T \cdot (R-1)}$$

$s_i$ is the row vector of $R$ PLS scores of sample $i$ for which the MD should be calculated.
$S$ is the $N$-by-$R$ matrix of row-wise PLS scores of all training samples.
$R$ is the PLS rank.
$N$ is the number of training samples.
Superscript $T$ denotes the transposed matrix.
Superscript $-1$ denotes the inverse matrix.

**Calibration curve**
The calibration curve is the actual regression function of the calibration model. It can only be displayed for univariate calibration (IHM and PI).

$$x_i = Y_i \cdot K$$

$x$ denotes features of training samples.
$Y_i$ is the design matrix of component weights (IHM) or peak areas (PI) of sample $i$.
$K$ is a vector of regression coefficients (calibration constants).
Note that the regression function is linear in $K$, but could be nonlinear in component weights or peak areas depending on how the design matrix is built. For instance:

$$x_i = [\omega_i^0 \quad \omega_i^1 \quad \omega_i^2] \cdot [k_0 \quad k_1 \quad k_2]^T$$

would be quadratic in component weights $\omega$.

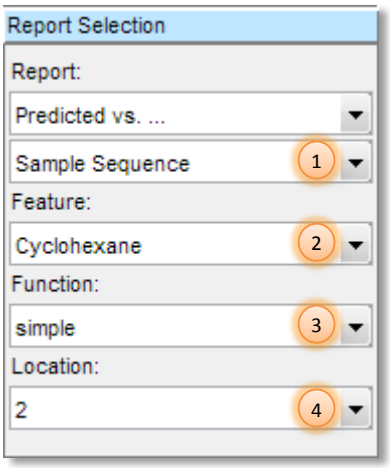
PEAXACT uses inverse calibration (regression of $x$ on $Y$) to calculate K. It has been shown that inverse calibration yields more precise predictions of unknowns than classical calibration [Tellinghuisen2000].

**PLS Loadings**
PLS loadings are the actual regression data from the PLS algorithm, and they allow you to examine how PLS "sees" the data; that is, how it is modeling the samples. PLS loadings take the form of spectra because the x-axis is the same as of the training samples. Typically, at higher ranks, loadings look more and more noisy. As the rank increases, more of the noise in the system is being modeled. Use this report to determine the maximum rank which does not model noise.

$$y_{reconstructed,i} = s_i \cdot L^T$$

$y_{reconstructed,i}$ is the reconstruction of spectrum $i$.
$s_i$ is the row vector of $R$ PLS scores of spectrum $i$.
$L$ is a $P$-by-$R$ matrix of column-wise loadings. Loadings are calculated from training samples using the SIMPLS algorithm.
$R$ is the rank of the PLS model.
$P$ is the number of data points.

### Report Customization




**X-Axis**
For several reports, you can change the x-axis to any provided data set feature (1). See Section Data Table Editor for how to add features to data sets.

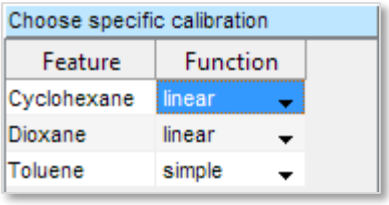
**Feature Selection**
Use the pull-down menu (2) to display results for individual calibrated features.

**Function / Rank**
Use the pull-down menu (3) to display results for individual regression functions / ranks.

**Location**
Use the pull-down menu (4) to display results for individual locations.

### Accepting a Specific Calibration

In order to finalize calibration, you have to decide on a specific regression function / rank for each calibrated feature. Make your choice in the bottom-left pane of the Report Generator Window.




By clicking the OK button of the Report Generator Window, your choice is accepted and the calibration model is stored in the master model. Furthermore, because calibration marks the final step of modeling, you are prompted to export the model summary report to your hard disk.

## 3.7.6.2 Remove Calibration Model

Context menu: Model Tree > Calibration Model Item > Remove

Removing the calibration model will remove all calibrated features. Please note that you cannot independently remove single calibrated features.

#### 3.7.6.3 Rename Calibrated Feature

| | |
|---|---|
| Context menu: | Model Tree > Calibrated Feature Item > Rename |
| Properties Panel: | Click Model Tree > Calibration Model Item |

Note: This operation applies to after a calibration model has been created. If you would like to change feature names prior to calibration, use the Data Table Editor.

From the context menu of the Model Tree Panel you can rename the selected calibrated feature. In addition, when activating the Calibration Model item in the Model Tree Panel, you can rename features in the Model Properties Panel.

Calibrated features can get any names as long as they are unique within the calibration model.

### 3.7.7 Modeling Assistant

| | |
|---|---|
| Menu bar: | Edit Model > Assistant... |

The modeling assistant is a step-by-step dialog for creating calibrated models. After starting the assistant it will guide you through the calibration process.

Please note that the assistant is not fully featured yet, i.e. several modeling features are not supported. So far, the assistant should be seen as a demonstration tool.

# 4 CHROM USER INTERFACE

## 4.1 Introduction

### 4.1.1 Product Overview

PEAXACT Chrom is a graphical user interface for the automatic classification of large numbers of gas chromatograms by comparison with a reference model. Functionality covers the whole process of sample classification, namely data visualization, modeling of reference samples with the goal of establishing validated tolerances, as well as reporting of results.

Features include:

- Step-by-step assistant for interactive data modeling
- Data pretreatment (baseline correction)
- Hard Modeling of reference chromatograms
- Tolerances for certain classification criteria
- Classification of samples using a simple ***classification code***
  - -3 Error: Reference components outside measured x-range
  - -2 Error: Peaks shifted out of measured x-range
  - -1 Error: Unhandled exception
  - 0 OK
  - 1 Baseline problem (currently undefined)
  - 2 Retention time shift too large
  - 3 Missing component
  - 4 Additional component
  - 5 Both, missing and additional components
- Command line usage for batch classification

### 4.1.2 Starting the Chrom User Interface

#### The Step-by-step Assistant

After starting the PEAXACT program choose the Chrom user interface from the Start Dialog.

Alternatively, you can run PEAXACT with startup option `-chrom` which directly starts the Chrom user interface.

### Command Line Usage

In addition to PEAXACT startup options, the Chrom user interface has individual command line options.

Parameters in square brackets are optional. Angle brackets represent placeholders which must be replaced by specific values. The vertical bar `(|)` separates alternative arguments.

```
PEAXACT -chrom
PEAXACT -chrom -help
PEAXACT -chrom [-builder | -analyzer | -classify]
        [<modelFile>] [<dataFileFilter1> [<dataFileFilter2> [...]]]

-help       display help for command line usage
-builder    start PEAXACT Chrom for interactive model building
-analyzer   start PEAXACT Chrom for interactive sample classification
-classify <modelFile> <dataFileFilter*> perform non-interactive
            classification of all data files matching <dataFileFilter*>
            using <modelFile> as reference model
```

Additional notes:

- `<dataFileFilter>` may contain the wildcard character (`*`) to match multiple files, e.g. `c:\my data\2012*.cdf`
- Filenames must be quoted if they contain spaces, e.g. `"c:\my data\file A.pxm"`

## 4.1.3 Quitting the PEAXACT Program

To quit the PEAXACT program, click the close box or the Close button in the main window. You may need to cancel running operations first in order to do this.

PEAXACT closes after prompting you to save any unsaved files.

### Enforced shutdown

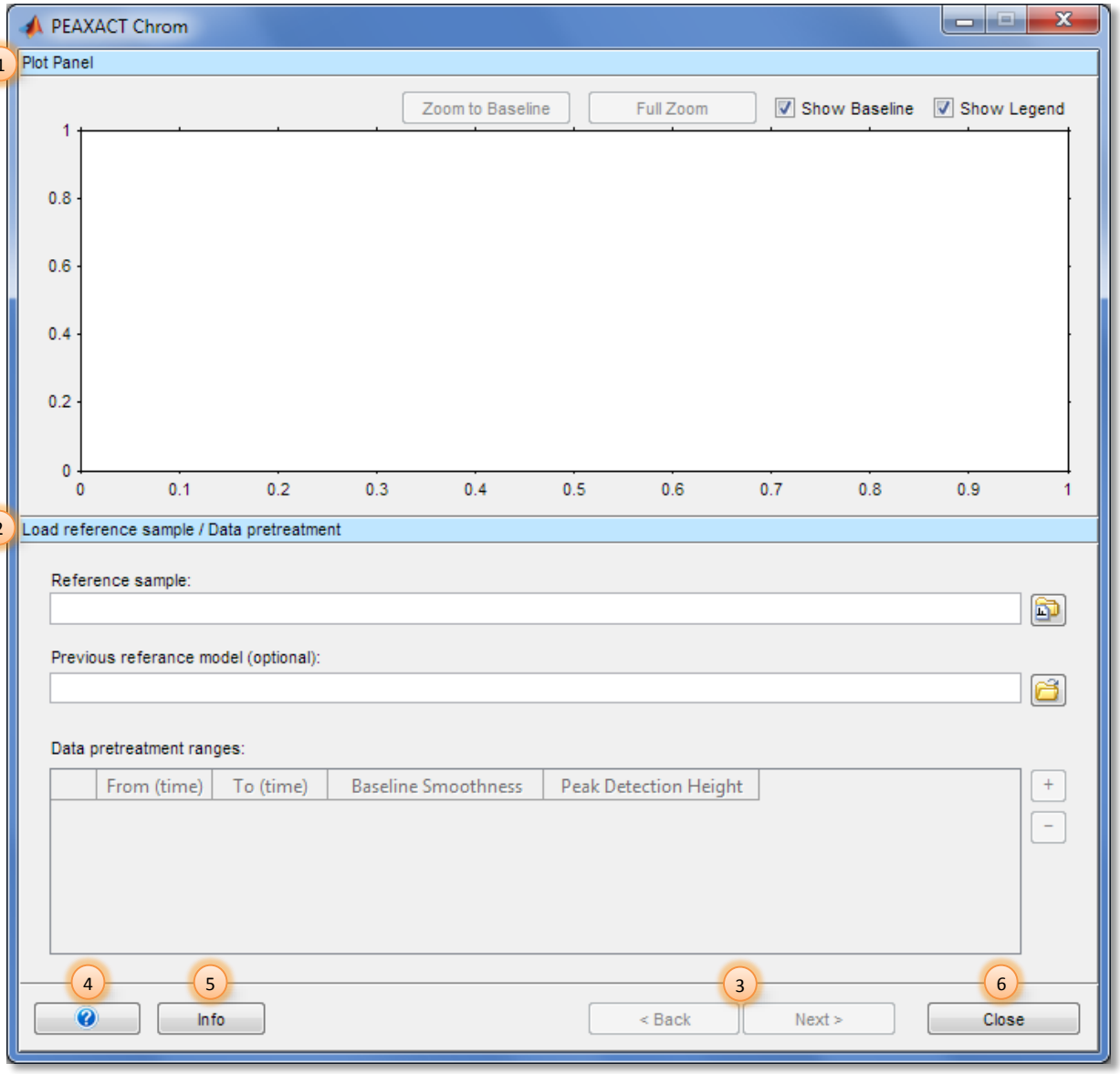
To force quitting the PEAXACT program at any time, click the close box in the Status Window.

> Caution: When closing the Status Window you will not be prompted to confirm quitting and all unsaved changes will be discarded.

# 4.2 Graphical User Interface

## About the Graphical Assistant

The graphical user interface of PEAXACT Chrom is a step-by-step assistant that guides you through the modeling and classification process.




**PEAXACT Chrom user interface**

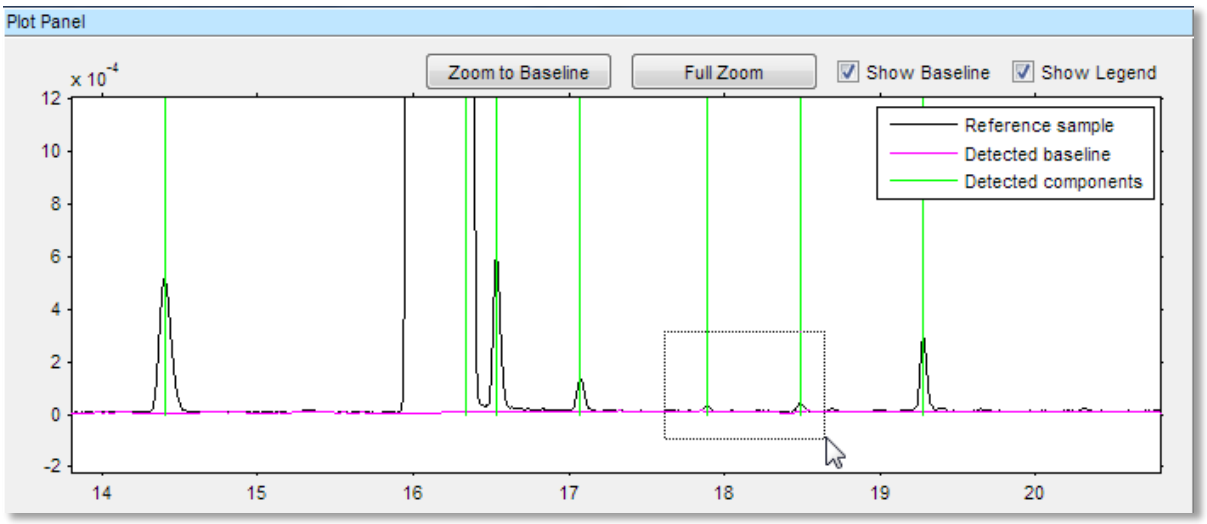
(1) Plot Panel: The Plot Panel (upper part of the window) is used for graphical output. It also offers a way to interactively edit the model by using the mouse.

(2) Properties Panel: The Properties Panel (lower part of the window) changes its appearance with each step. The panel displays controls (such as edit fields, pull-down menus, and tables) to modify parameter values.

(3) The Next button remains grey (inactive) until you have provided enough information to proceed to the next step. The Back button can be used to go back to the previous step if necessary.
(4) Opens this user manual (PDF file)
(5) Displays additional information about PEAXACT
(6) Quit PEAXACT

### Plot Panel

The Plot Panel offers some features for data visualization and zooming.




**Plot Panel**

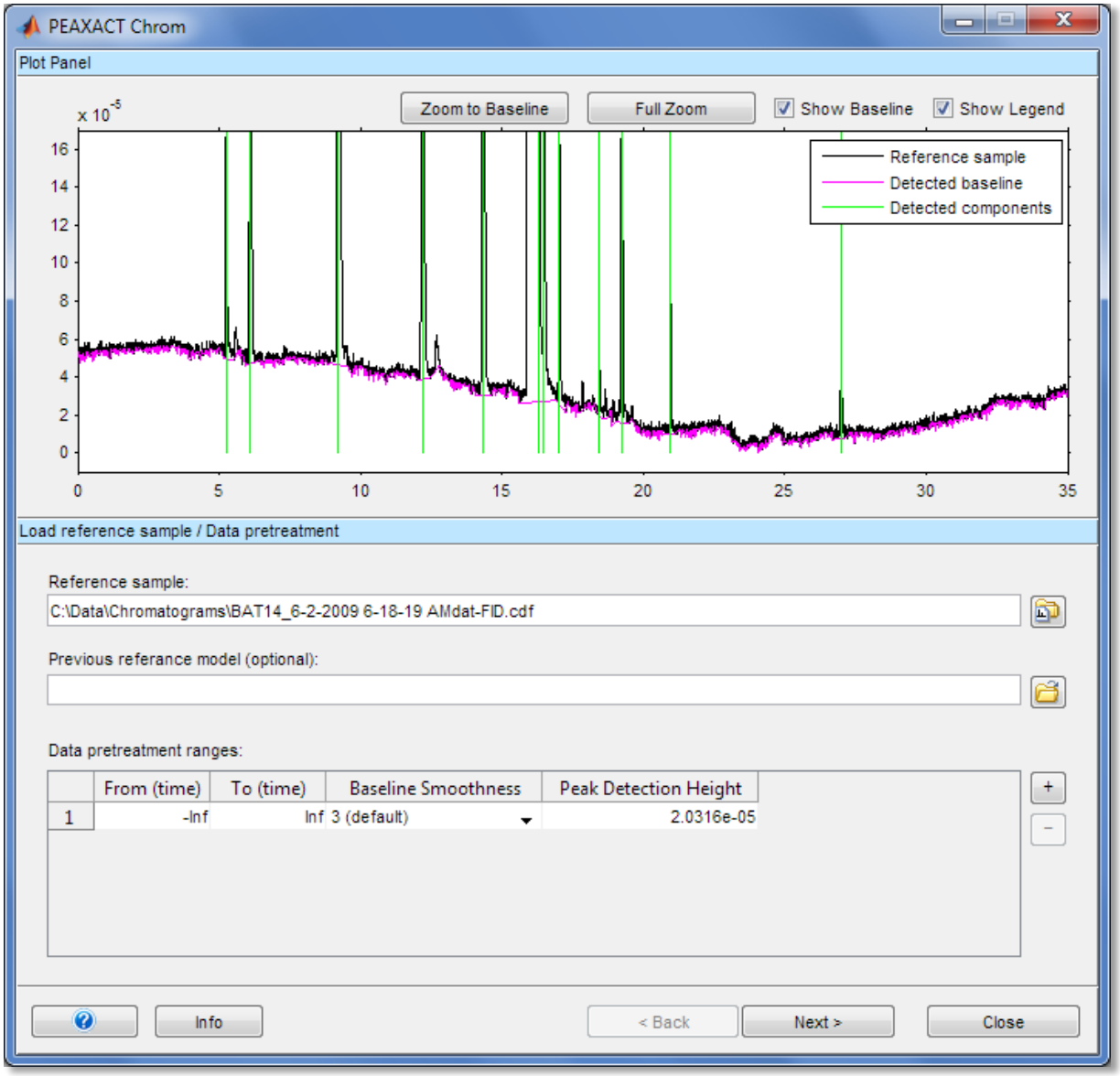
- To **zoom in,** left-click into the figure, hold the button down and drag the mouse to select a rectangular area. When you release the button, this area will be shown in detail. This operation can be repeated.
- To **zoom out,** right-click on the Figure area or left-click while holding the CTRL key down. This operation can be repeated.
- To move the visible region, right-click into the figure and hold the button pressed while moving the mouse.
- **Zoom to Baseline** button: zoom in on baseline. This is helpful to inspect the baseline and to see tiny peaks with just one click.
- **Full zoom** button: zoom out to get a full view on the data
- **Show Baseline** checkbox: show or hide a plot of the detected baseline
- **Show Legend** checkbox: show or hide a legend

## 4.3 Modeling

The goal of modeling is to create a ***peak model*** (also called hard model) of a reference chromatogram. For classification of chromatograms, the model will be fitted and compared to measured samples.

## 4.3.1 Step 1: Reference Sample & Pretreatments

The goal in step 1 is to load a reference chromatogram and set-up data pretreatments such that all relevant reference peaks will be detected. Start by clicking the button next to the **Reference sample** text field in order to load a reference file. Optionally, you can load a **Previous reference model** by clicking the button. The model may be useful in later steps.
The reference sample is displayed in the Plot Panel along with the detected baseline (if **Show Baseline** is checked) and detected peak positions.




**Reference data (black), detected baseline (purple) and peaks (green). Baseline and peak detection according to data pretreatments defined for range [-inf, +inf].**

Baseline and peak detection are controlled by two parameters:

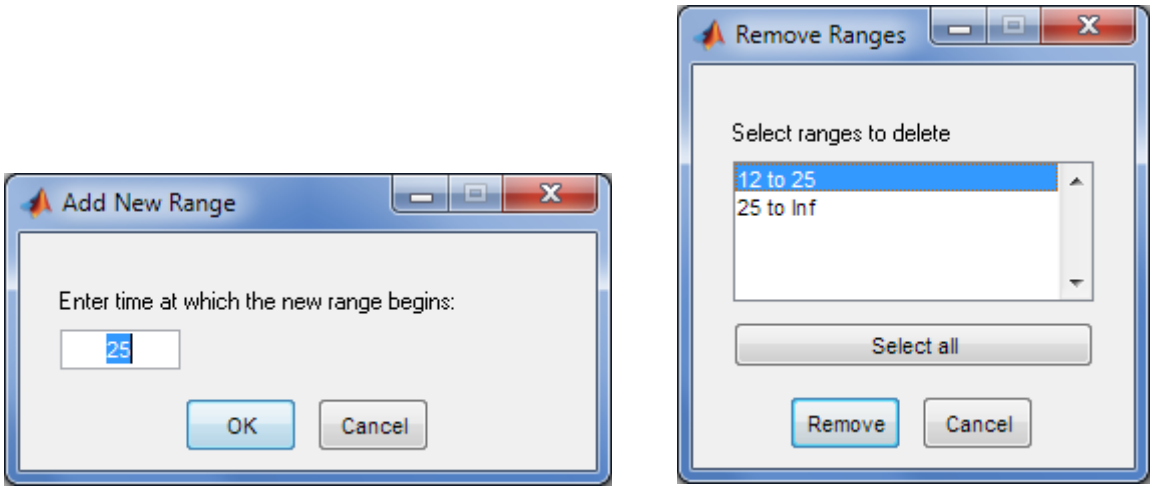
- **Baseline smoothness level**: an integer between 1 and 5. The larger the value the smoother the detected baseline. The default value is 3.

- **Peak detection height:** A value defining the minimum height a peak must have to count as peak. When loading a new reference sample, this value is determined automatically from the y-data of the sample.

Because the result of peak detection is relevant for modeling and classification, you should consider adjusting both parameters using the **table** in the Properties Panel. Typically, the initial baseline smoothness level doesn't need to be changed, but the minimum peak detection height should be set that small that all significant components will be detected, and that large that insignificant components will not be detected.

Changes in the pretreatment settings are immediately applied to the reference sample (indicated by the **Busy...** message at the bottom of the window) and the Plot Panel is refreshed.

By default, baseline smoothness and peak detection height apply to the full x-range from minus infinity (-Inf) to plus infinity (Inf). If you are not satisfied with the results you can split the range into new **data pretreatment ranges** and define individual settings for each. For example, you can set the minimum peak detection height to a large value in order to ignore all peaks in that region. Click the [+] button and enter the retention time at which the new range should begin; click the [−] button to remove ranges. The first range will always start at time –Inf and the last range will always end at time +Inf.



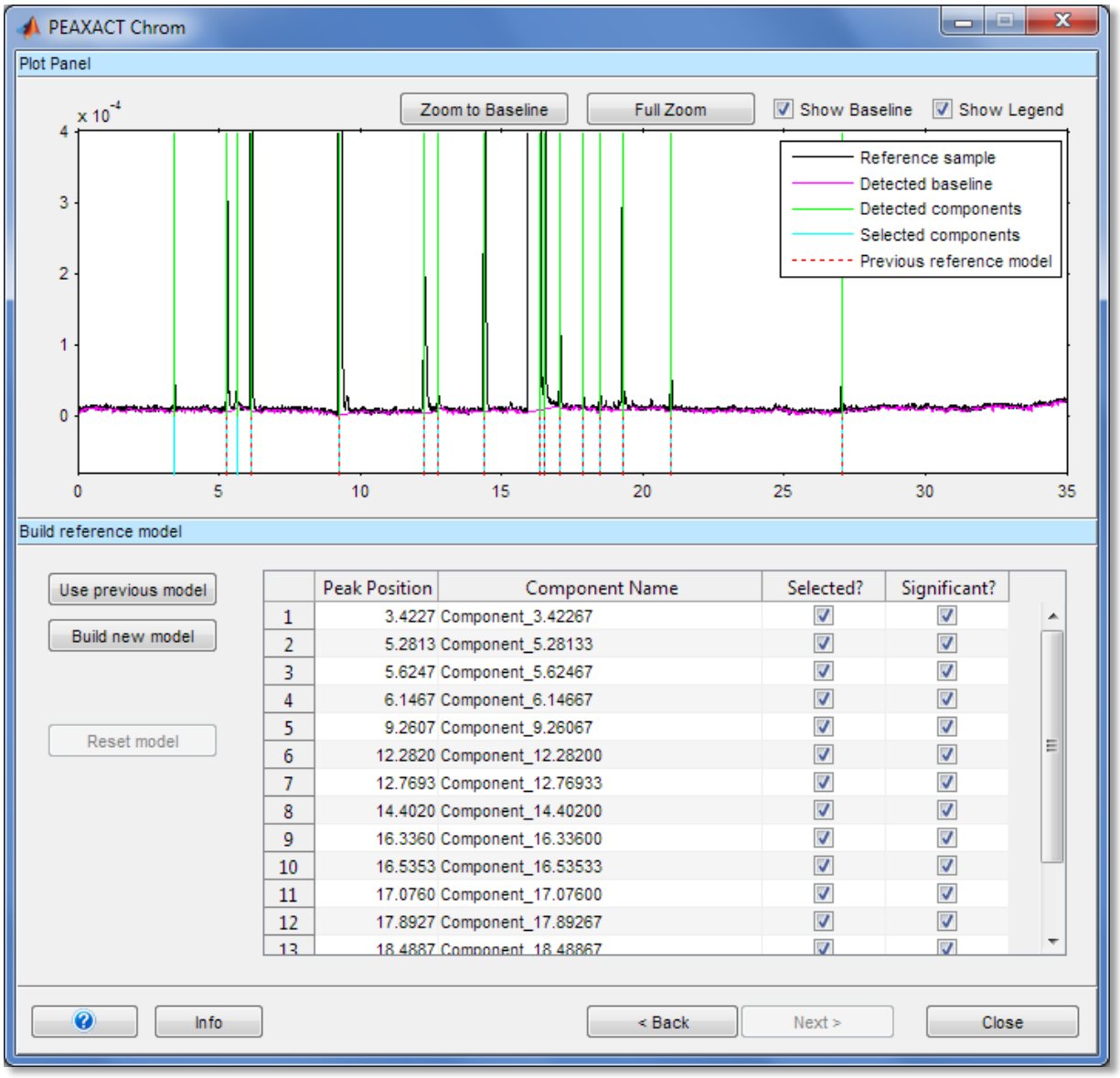

**Dialog boxes for adding (left) and removing (right) data pretreatment ranges**

**CAUTION:** You should use as less ranges as possible! Because same pretreatments chosen in step 1 will later be applied for classification, it is more likely that using many pretreatments removes artifacts you actually want to test for.

Proceed with to step 2 after all pretreatments have been made.

## 4.3.2 Step 2: Building a Reference Model

The second step deals with generating the peak model. You can either re-use the previous reference model by clicking the **Use previous model** button or you can create a new model by clicking the **Build new model** button.




**Step 2: Building a reference model**

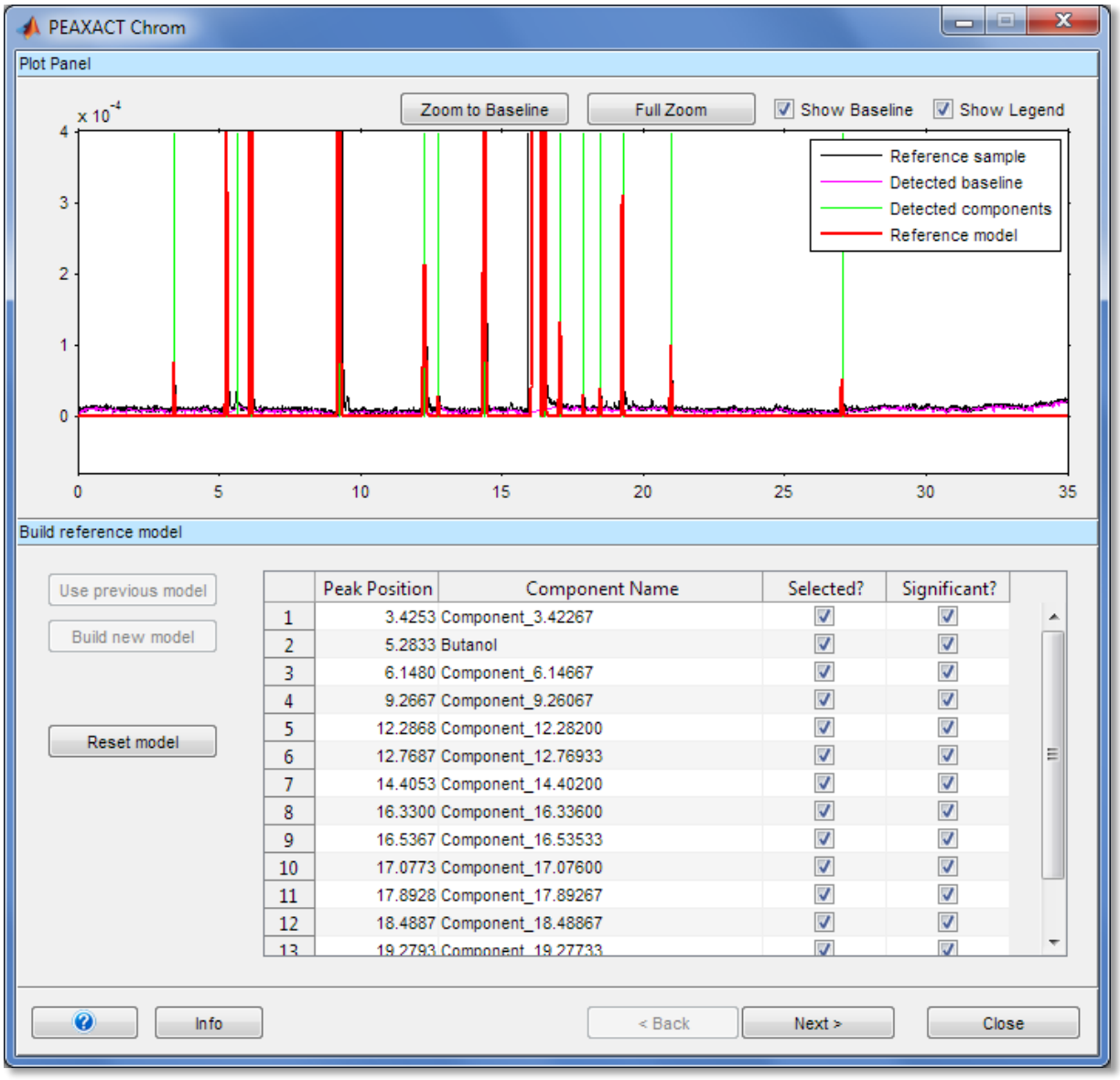
The **table** in the Properties Panel can be seen as a "recipe" for creating a new model. The table contains a row for each peak detected in the reference sample as well as 4 columns:

- **Peak Position**: retention time of detected peak (unchangeable)
- **Component Name**: you can name each peak, e.g. after a chemical compound
- **Selected?** Choose which peak should be modeled and which peaks shouldn't.
- **Significant?** Modeled peaks marked as significant must occur in a chromatogram in order not to be classified as "missing component"; insignificant peaks are optional.

For visual assistance, each selected peak is marked in light blue in the Plot Panel. Unselected components are marked dark blue. You can also toggle the "selected" state of each component by clicking on its corresponding blue line in the Plot Panel. If a previous reference model is loaded, peak positions of the model are shown as dashed red lines.

In order to complete modeling, you can either press the **Use previous model** button or the **Build new model** button. This will display the model in the Plot Panel and makes it unchangeable. If you want to change it anyway, click **Reset model**.

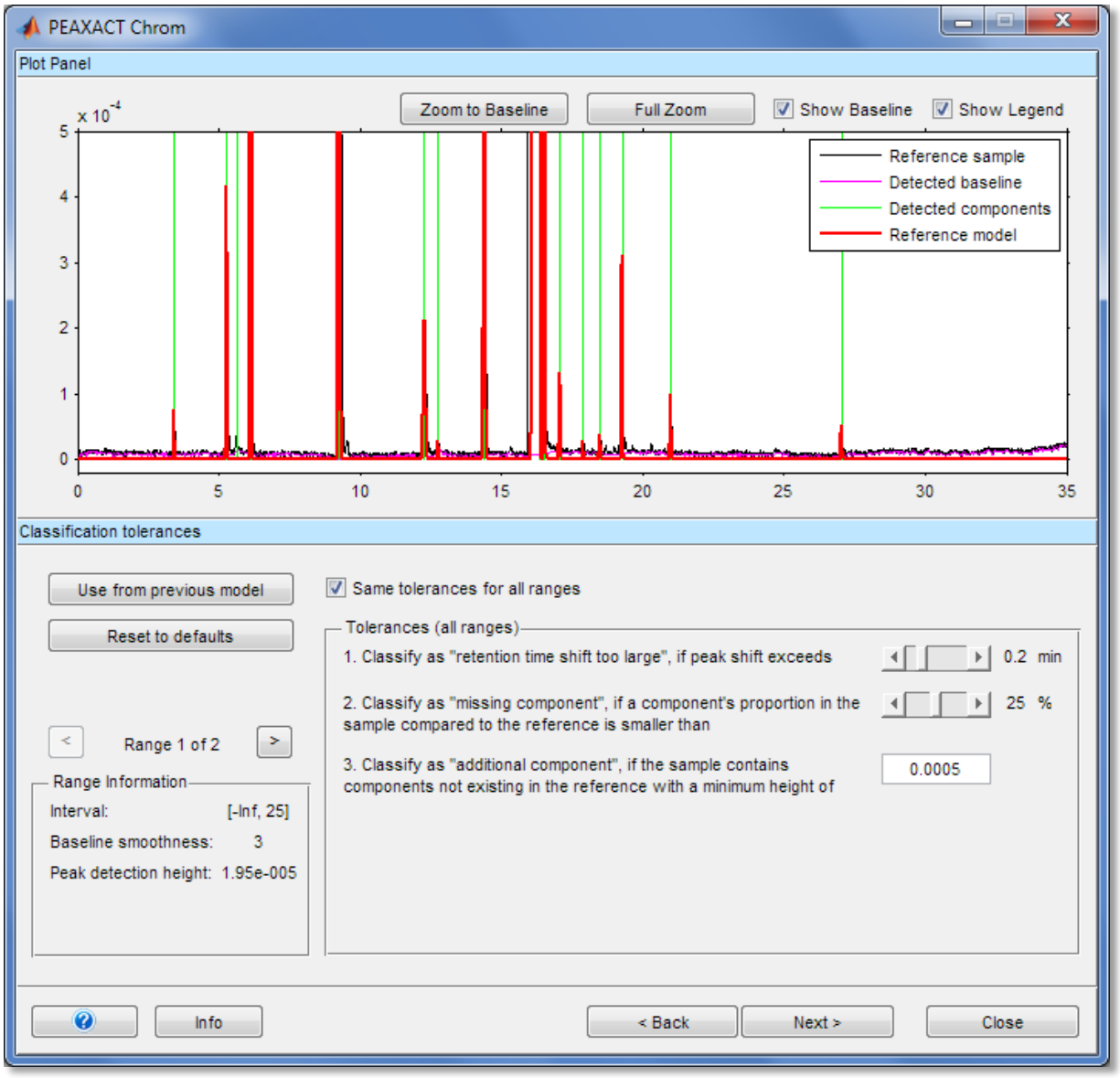
After deciding on a reference model, the table in the Properties Panel reduces to a list of selected components.




**Step 2: After choosing a reference model**

## 4.3.3 Step 3: Setting Classification Tolerances

In the third step you define classification tolerances. Tolerances can either be individual for each range specified in the first step, or alternatively, you can place a checkmark to use **Same tolerances for all ranges**.




**Step 3: Classification tolerances**

There are three tolerances to set:

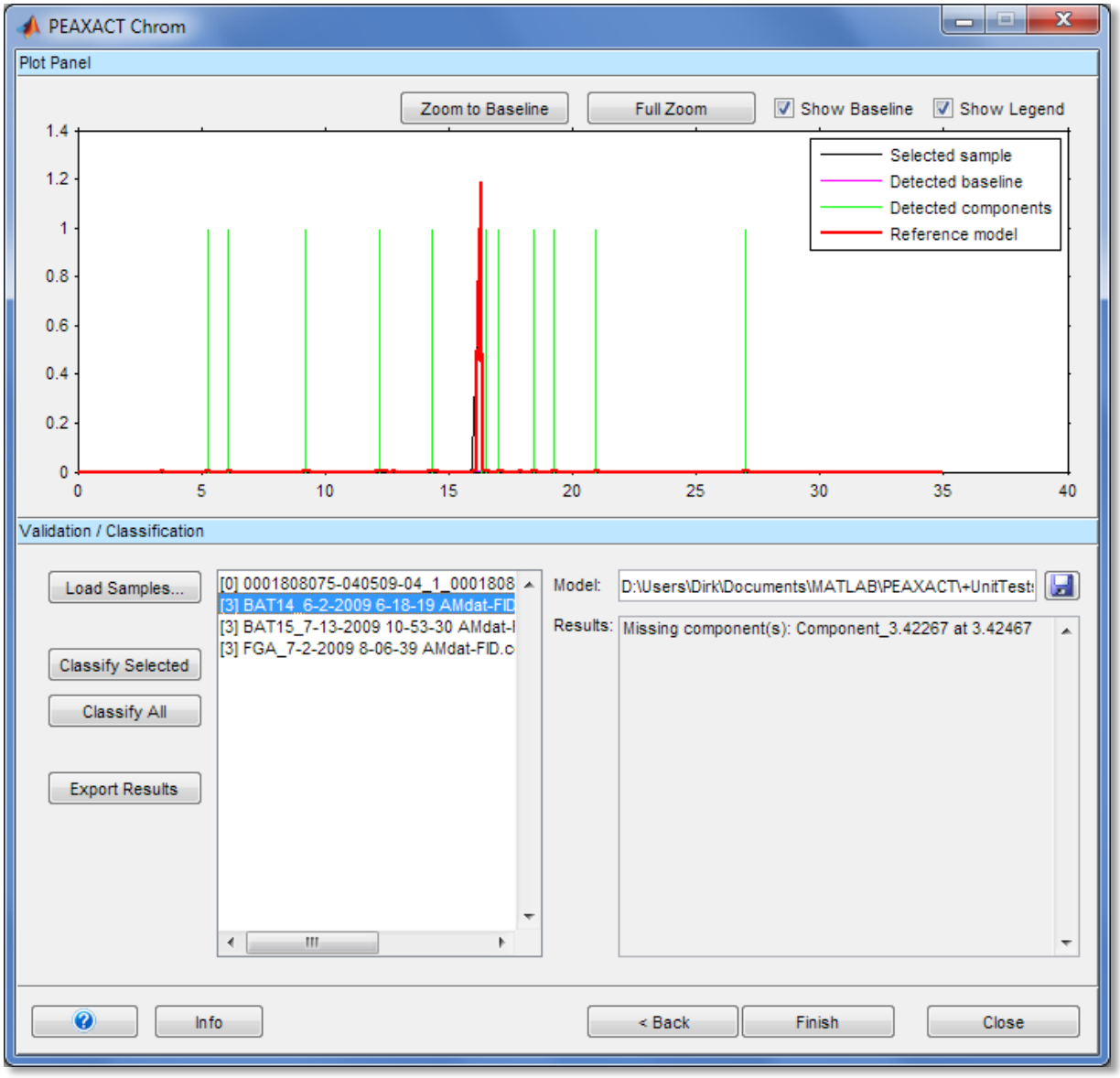
- Maximum tolerated component shift
- Minimum tolerated component fraction compared to reference
- Minimum additional peak height

The descriptions for the parameters are self-explanatory. Please note that the third tolerance for detecting additional components is not the same as the minimum peak detection height that has been set in the first step. This tolerance can be used to only classify those detected peak as additional which exceed a certain height.
At any time, you can reset these tolerances to the values stored in the previous model (**Use from previous model**) or to the default values (**Reset to defaults**).

If you uncheck "Same tolerances for all ranges", tolerances must be specified for each range individually. The **Range information** panel as well as the < > buttons on the left side will help you navigate through the ranges.

## 4.3.4 Step 4: Model Validation

In the fourth step you can validate the model tolerances using chromatograms for which you already know the classification result.




**Step 4: Model validation**

First, **Load Samples** for classification via a file open dialog (you can select multiple chromatogram files at a time). Samples are added to the list in the Properties Panel.

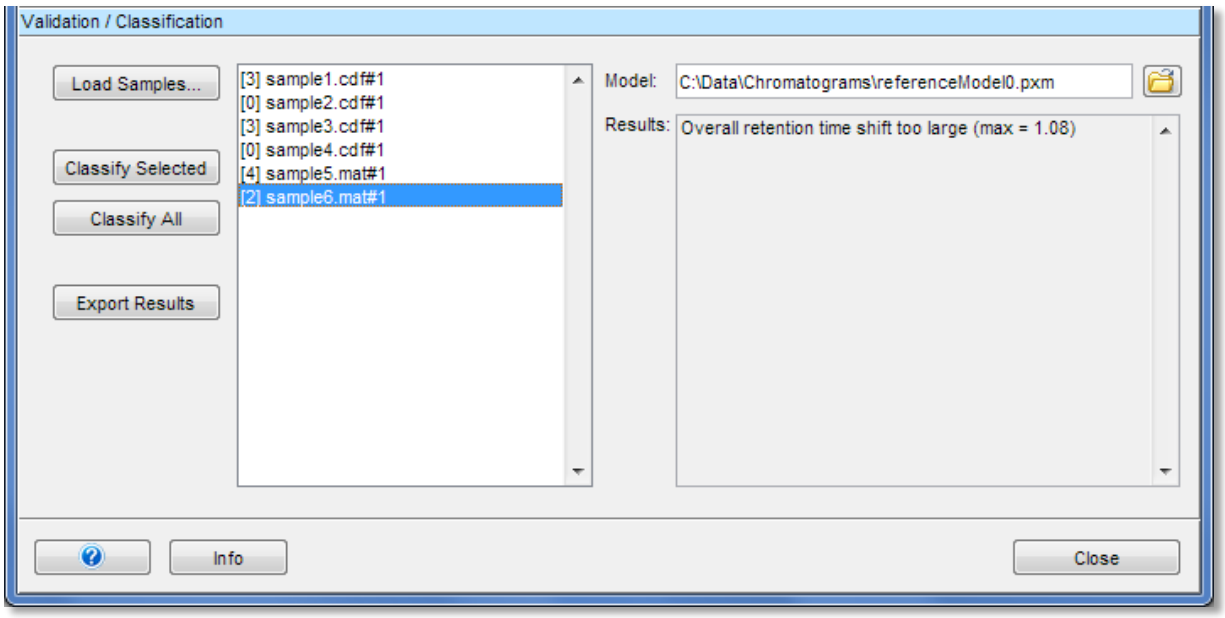
Select a sample from the list and click the **Classify Selected** button to perform the classification based on your tolerances. The resulting classification code is added as prefix to the sample name in the list. You can also classify all samples at once with the **Classify All** button. If the classification results do not match your expectations, go back and forth, change the model tolerances and re-process the data. You may also want to manually check the classification results, e.g. check missing or additional components, by zooming into the Plot Panel.

You can save the model by clicking the button. The **Finish** button closes the assistant after prompting you to save model changes.

# 4.4 Classification

For the classification of samples, the same window is used as for step 4 (validation) of the modeling process. In fact, the classification procedure is identical to model validation except that you typically don't know results in advance.

When you start the PEAXACT Chrom – Analyzer interface from the Start Dialog or from the command line, it directly brings you to step 4. In this case you need to load a model file ( ) to be used for classification. Also, you cannot go back to step 3.




**Classification of samples**

**Load Samples** for classification and click **Classify Selected** or **Classify All** to start the analysis. The resulting classification code is added as prefix to the sample name in the list. Results can be exported to spreadsheet or read-only files using the **Export Results** button. See Section Report File for details.

# 4.5 Known Issues

The success of modeling and model application for classification depends on the quality of your data and is restricted by numerical issues in cases of "bad data".

## Strong fronting/tailing

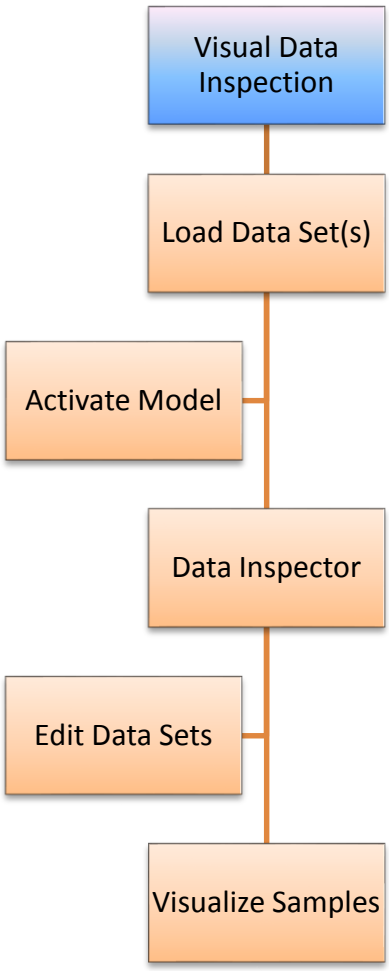
In case of strong fronting or tailing, a small peak overlapping the front or tail of a large peak may not be classified correctly. If the small peak is contained in the model, the component may not be found to be missing even if it is missing in the sample. If the peak is not contained in the model, the component may be found to be additional even if it is not. To avoid wrong classification, the small peak should be contained in the model and should be marked insignificant.

## Large global retention time shifts

In case of a large overall (global) retention time shift – which is smaller than the allowed global shift parameter though – some peaks may be classified as missing because local shifts may become too large.

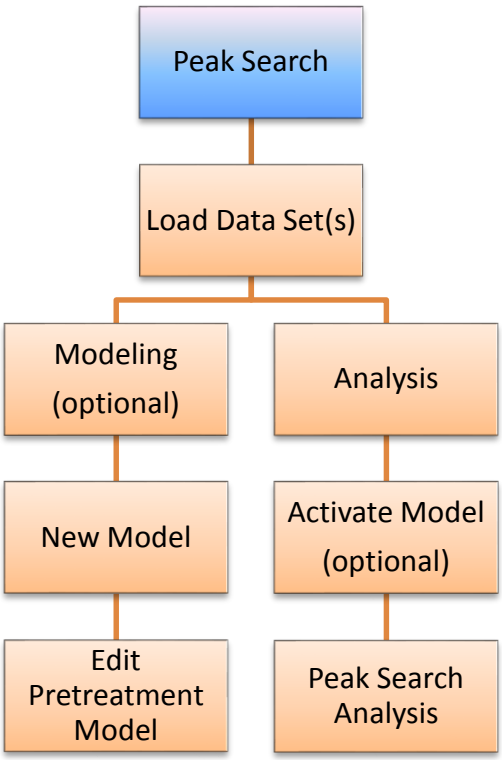
# 5 ANALYSIS WORKFLOW

## 5.1 Visual Data Inspection



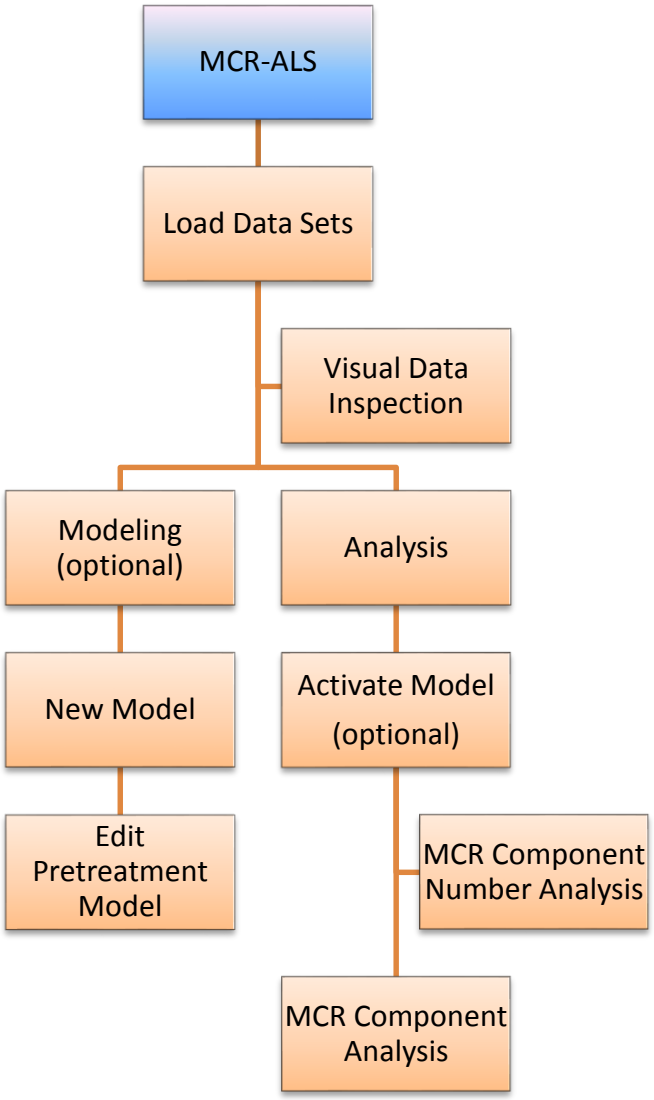

**Data Inspection Workflow**

# 5.2 Peak Search



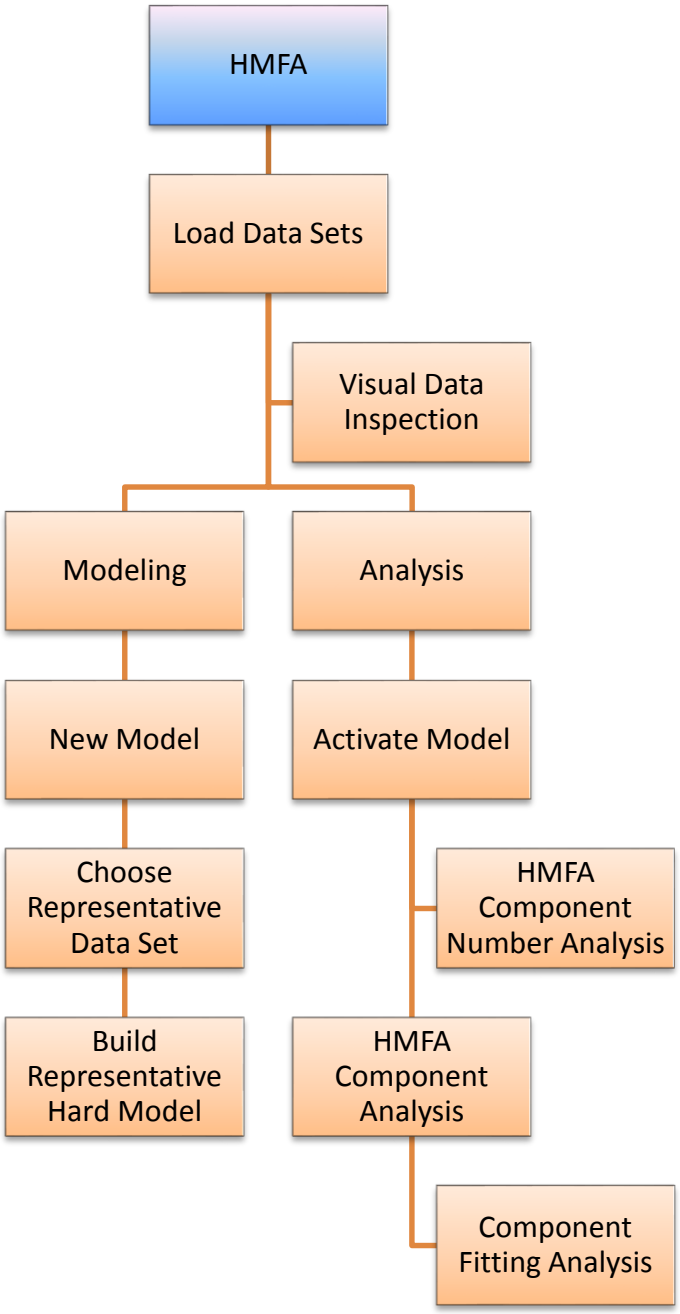

**Peak Search Workflow**

# 5.3 Multivariate Curve Resolution (MCR)



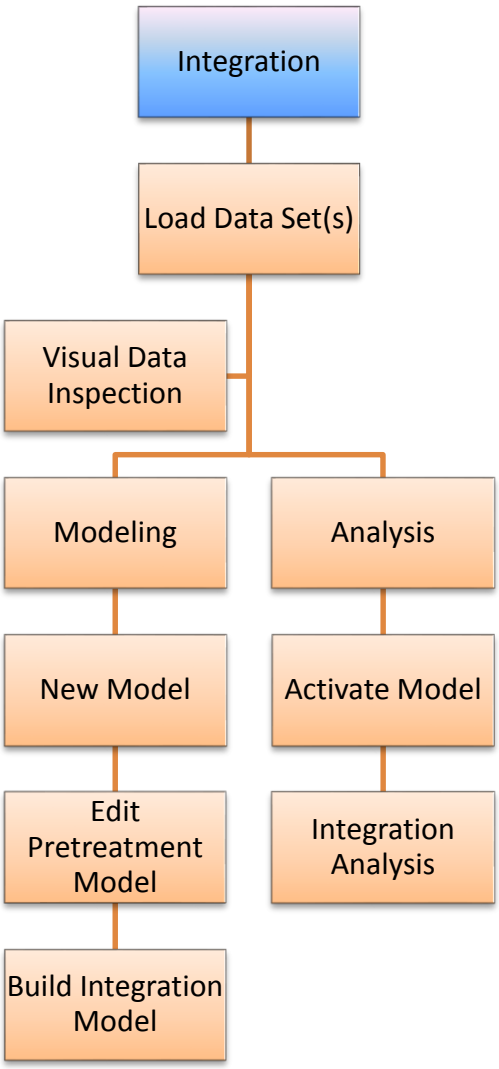

**MCR-ALS Workflow**

# 5.4 Hard Modeling Factor Analysis (HMFA)



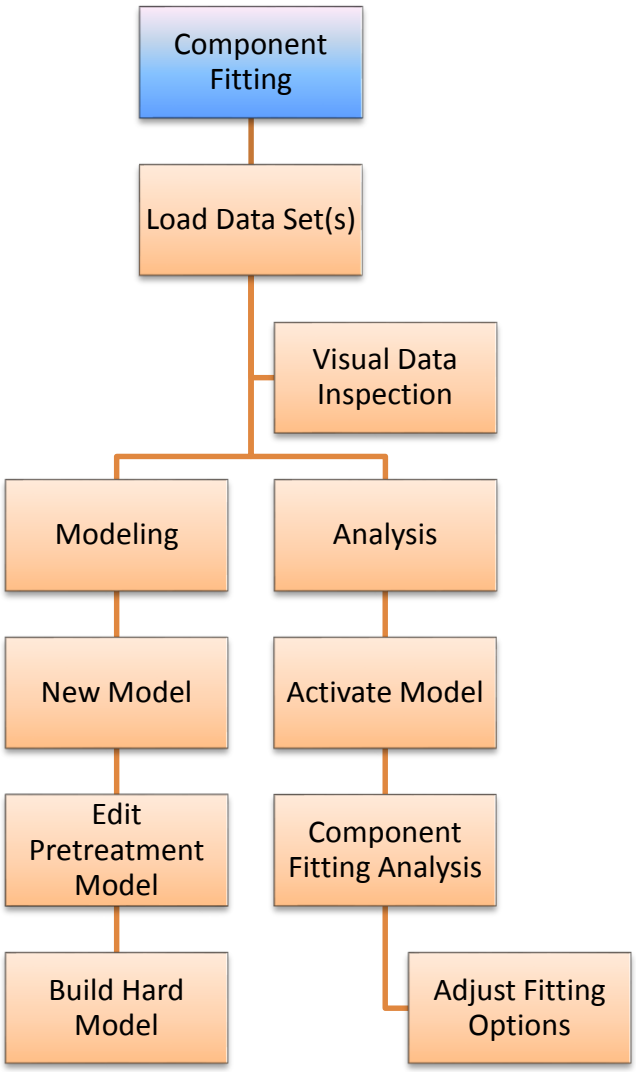

**HMFA Workflow**

# 5.5 Integration Analysis



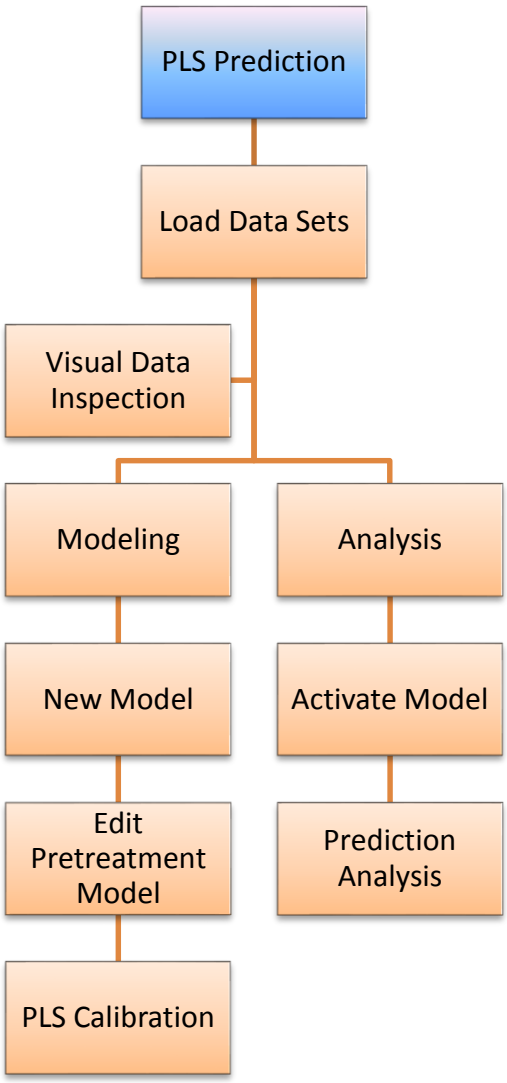

**Integration Analysis Workflow**

# 5.6 Component Fitting Analysis



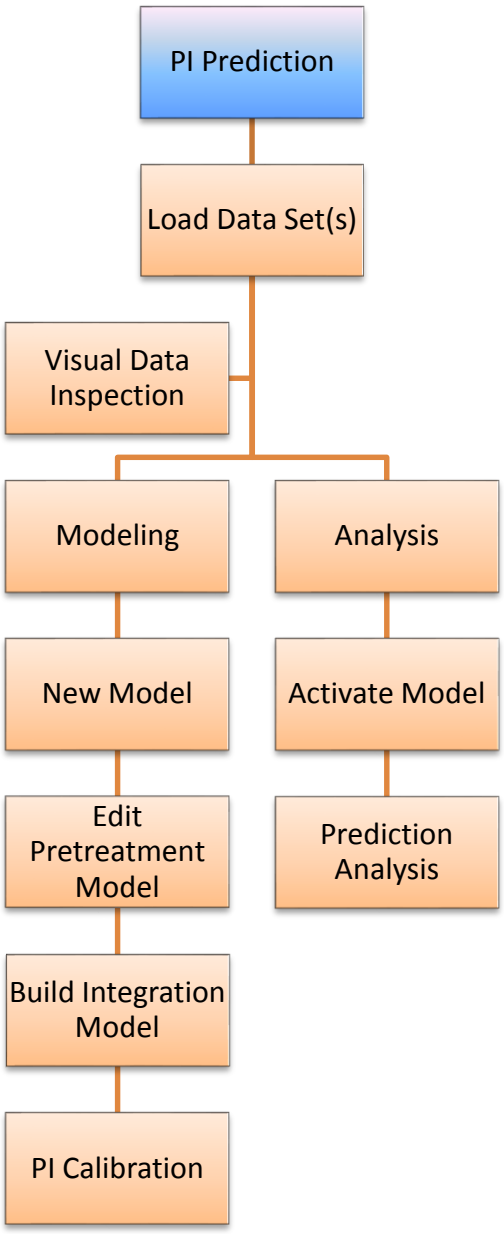

**Component Fitting Analysis Workflow**

# 5.7 Partial Least Squares (PLS) Prediction



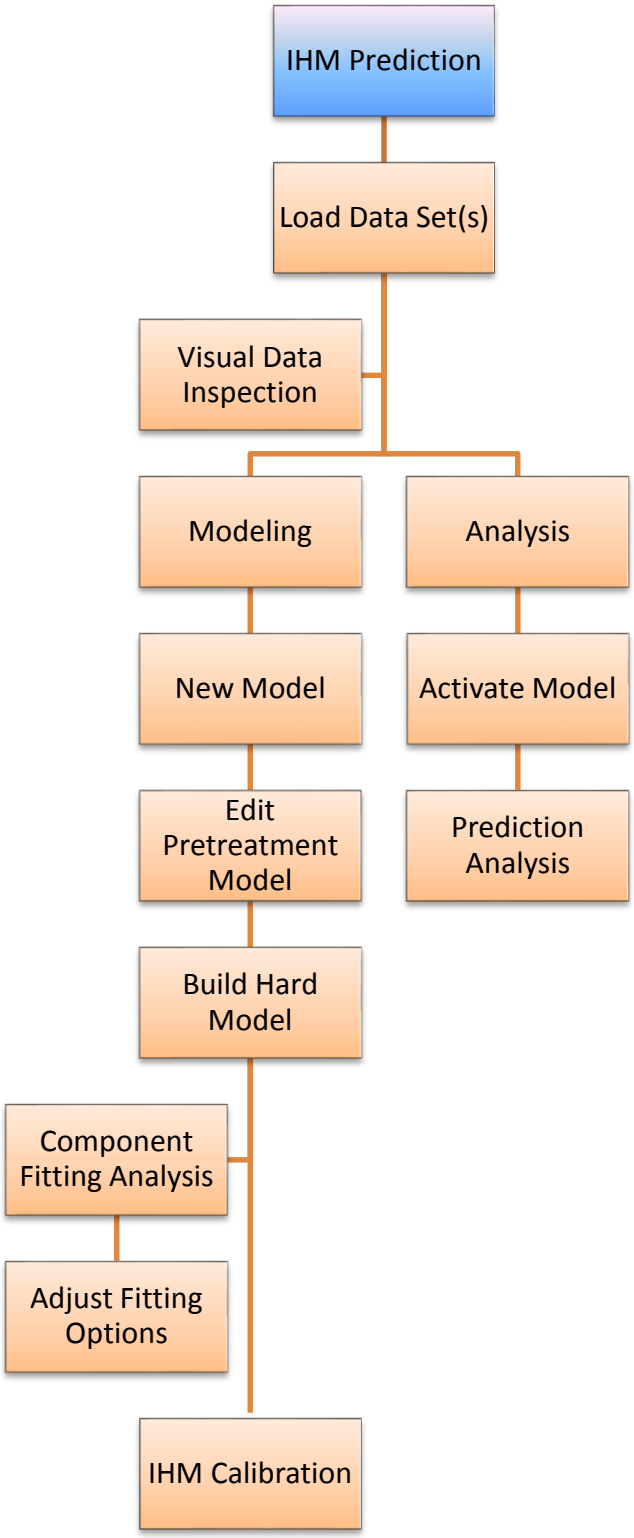

**PLS Prediction Workflow**

# 5.8 Peak Integration (PI) Prediction



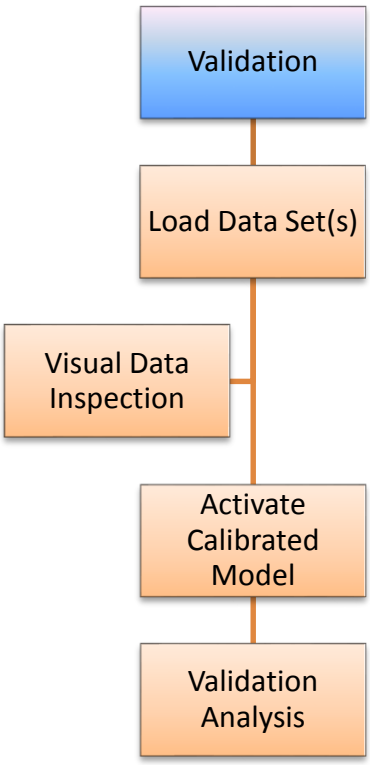

**Peak Integration Prediction Workflow**

# 5.9 Indirect Hard Modeling (IHM) Prediction



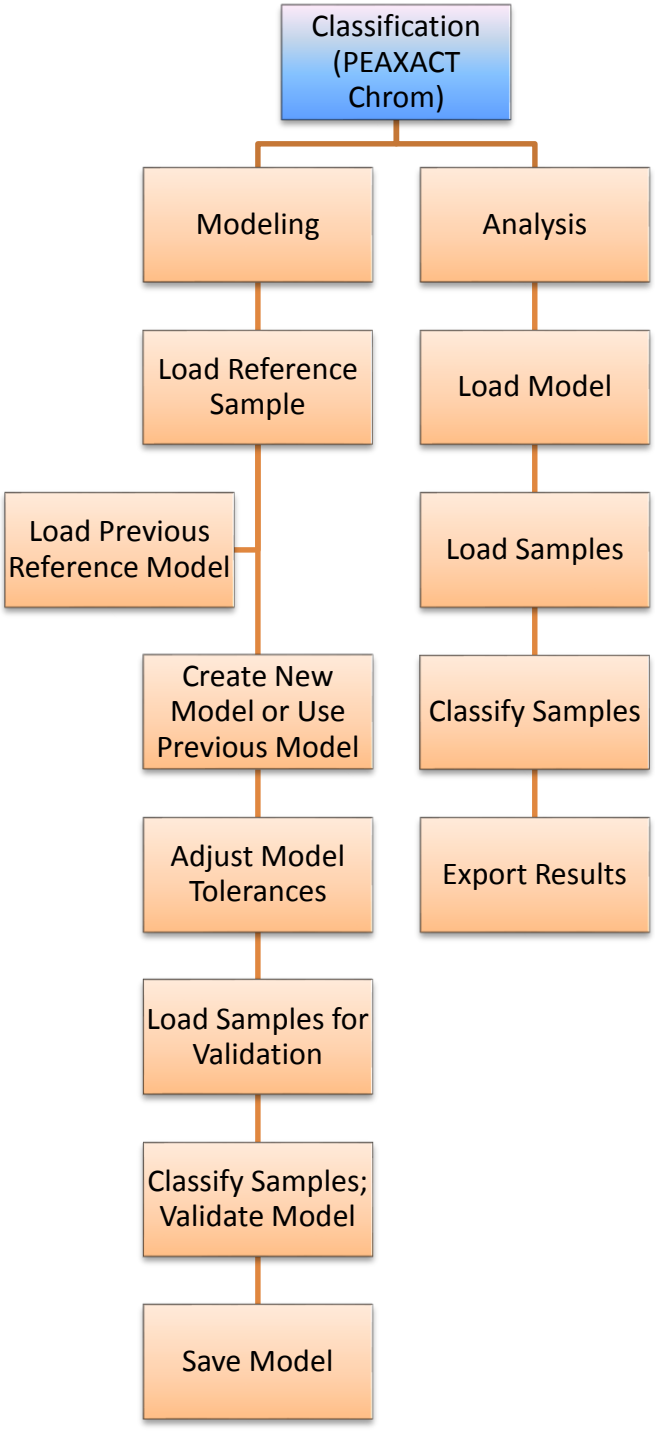

**IHM Prediction Workflow**

# 5.10 Validation




**Validation Workflow**

# 5.11 Classification of Chromatograms




**Classification Workflow**

# 6 TROUBLE SHOOTING

## Graphics problems

### Symptoms

Some PEAXACT windows such as the Desktop or the Data Inspector are not displayed correctly. Those graphics problems e.g. may include upside-down text, duplicated text, and misplaced axes tick labels.

### Cause

Problems are related to hardware-based graphics acceleration performed by some video cards.

### Resolution

Start PEAXACT with command line option `-openglfix`. This forces the usage of software-based graphics acceleration. You should only use this command line option if you encounter graphical problems because software-based acceleration is typically slower than hardware-based acceleration.

- Right-click the Windows Desktop and choose **New > Shortcut**
- Enter
  `"INSTALLDIR\peaxact.exe" -openglfix`
  including quotes; replace `INSTALLDIR` with the PEAXACT installation path on your computer.
- Click **Next** and enter a shortcut name.
- From now on start PEAXACT using the new shortcut.

## PEAXACT immediately shuts down during startup

### Symptoms

When starting PEAXACT from the Windows desktop or start menu, a black command window appears for a short time and closes immediately.

### Cause

This error can have several reasons, e.g. it may result from incomplete installation, missing user privileges, loading a corrupt user preferences profile, and others.

### Resolution

Start PEAXACT from the command prompt in order to get displayed an error message.

- From the Windows start menu, run `cmd` which opens a command window
- At the command prompt, type
  `"c:\Program Files\S-PACT\PEAXACT\peaxact.exe" -logLevel debug`
  You may need to replace the directory with the PEAXACT installation directory on your computer. Don't forget the quotes.
- PEAXACT probably fails to start again, but an error message will be displayed.

| Error Message | See further |
|---|---|
| Error message involves file `mclmcrrt714.dll` | Missing mclmcrrt714.dll |
| Error message involves file `mllapack.dll` | Error loading mllapack.dll |
| Initialize component instance failed<br>Could not access the MCR component cache. | Problems with MCR component cache |
| This application has requested the run time to terminate in an unusual way. | Other Problems |
| Others | Other Problems |

## PEAXACT starts with error regarding missing `mclmcrrt714.dll`

### Symptoms

When starting PEAXACT, a dialog box appears showing a message similar to:
*This application has failed to start because mclmcrrt714.dll was not found.*

### Cause and Resolution

`mclmcrrt714.dll` is a run-time library needed to run PEAXACT. This library is installed during the course of the MATLAB Compiler Runtime (MCR) installation.

This error typically results from one of two situations:

1) The correct version of MATLAB Compiler Runtime (MCR) is not installed.

   PEAXACT depends on a specific version of the run-time library. Please follow the installation instructions in this documentation.

2) A second option is that the system's PATH environment variable was not set correctly.

   During the course of the MATLAB Compiler Runtime installation, the following directory should have been added to your system path:

   `MCRROOT\v714\runtime\win32` or `MCRROOT\v714\runtime\win64`

   To work around this issue, manually update the PATH environment variable on your deployment machine. Alternatively, re-installation of MCR should fix this problem.

   Also note that since these DLLs have different names for each version of the MCR, multiple versions of the MCR can be installed on the same machine. Make sure to have the right version installed for PEAXACT.

## A dialog appears entitled `mllapack.dll` or PEAXACT throws errors related to file `mllapack.dll`

### Symptoms

The application may be started correctly, but accessing certain features invoke a dialog box entitled `mllapack.dll` with the message *The specified procedure could not be found.* The message is language dependent, e.g. the German message is: `Die angegebene prozedur wurde nicht gefunden`.
The PEAXACT Status Window may display additional information on this error.

### Cause

The error comes from a DLL version problem. If an older version of the file `libmmd.dll` is present in the `%WinDir%\System32` directory, the newer version that comes with the MATLAB Compiler Runtime is not used.

### Resolution

Getting rid of older versions of file `libmmd.dll` solves the problem. You may need administrator privileges to make changes to the `%WinDir%\System32` directory.

1) Use the Windows Explorer and browse to `%WinDir%\System32`. The variable `%WinDir%` typically resolves to `c:\Windows`.
2) Rename file `libmmd.dll` to `libmmd.dll_old`

## Problems with MCR component cache

### Symptoms

The following error is displayed during PEAXACT startup:
*Initialize component instance failed.*
*Could not access the MCR component cache.*

### Cause

During startup, PEAXACT extracts files to the MCR component cache directory. The error occurs when these files cannot be written or read. By default, the MCR cache is located in the user's application data directory to which the user normally has full writing permissions. However, depending on the configuration of your Windows user profile, this might not be the case.

### Resolution

There are two possible solutions:

1) Changing your Windows user's privileges. Contact your system administrator and request writing permissions for the following directory: `%LocalAppData%\Temp`
   Because it might conflict with the security policy of your company, this solution is not recommended.
2) Redirecting the MCR cache directory:
   - From the Windows start menu, run `notepad` or any other text editor
   - In the text editor, type
     `set MCR_CACHE_ROOT=c:\MCR`

```
if not exist "%MCR_CACHE_ROOT%\nul" mkdir "%MCR_CACHE_ROOT%"
"c:\Program Files\S-PACT\PEAXACT\peaxact.exe"
```

- You may need to substitute `c:\MCR` with a directory and path for which you have writing permissions. Also, replace `c:\Program Files\S-PACT\PEAXACT` with the PEAXACT installation directory on your computer.
- Save the text file as `peaxact.bat` on your Windows desktop. Note that the file extension must be "bat" and not "txt" (which would be the default).
- Double-click the peaxact.bat file on your Window desktop

## Other Problems

If you encounter a problem not mentioned here, try the following:

- Start PEAXACT with startup option `-logLevel debug`. There is a file shortcut in the PEAXACT installation directory called `PEAXACT 3 Debugging` which starts PEAXACT with this option.
- Check the PEAXACT log file for possible explanations
- Take a screenshot from the command window

If you cannot solve the problem on your own, send the log-file and the screenshot to Technical Support: support@s-pact.de

# 7 LITERATURE


- Alsmeyer, F.; Koß, H. & Marquardt, W.: *Indirect spectral hard modeling for the analysis of reactive and interacting mixtures*. Appl. Spectrosc., 2004, 58, 975-985
- Alsmeyer, F. & Marquardt, W.: *Automatic generation of peak-shaped models*. Appl. Spectrosc., 2004, 58, 986-994
- Kriesten, E.; Alsmeyer, F.; Bardow, A. & Marquardt, W.: *Fully automated indirect hard modeling of mixture spectra*. Chemom. Intell. Lab. Syst., 2008, 91, 181-193
- Kriesten, E.; Mayer, D.; Alsmeyer, F.; Minnich, C. B.; Greiner, L. & Marquardt, W.: *Identification of unknown pure component spectra by indirect hard modeling*. Chemom. Intell. Lab. Syst., 2008, 93, 108-119
- Martens, H. & Næs, T.: *Multivariate Calibration*. John Wiley and Sons, 1989
- Tellinghuisen, J.: *Inverse vs. classical calibration for small data sets*. Fresenius J Anal Chem, 2000, 368, 585–588