RootProf

TUTORIAL 3

Supervised quantitative analysis

Contents

Chapter 1

The data set

Unidimensional patterns from X-ray diffraction measurements on polycrystalline mixtures form out dataset. Experimental samples have been produced by crystallization processes aiming at obtaining co-crystals formed by an active pharmaceutical ingredient (API) and a co-former. In our case polymorph III of carbamazepine constitutes the API, saccharine the co-former. The experimentally determined weight fractions of carbamazepine (CBZ III), saccharine (SAC) and co-crystal (CBZ-SAC) are reported in Table 1. The corresponding files are included as demo files. They are formed by two columns, the first containing the 2θ values, the second the corresponding values of diffracted intensity.

Table 1: Weight fractions of prepared mixtures. Samples 6-8 (shadowed) are composed by pure phases.

Sample n.	CBZ III	SAC	CBZ-SAC	File name
0	0	0.565	0.435	Rocco_S3_mac.txt
1	0.500	0.500	0	Rocco_S5_mac.txt
2	0.500	0	0.500	Rocco_S7_Como.txt
3	0.347	0.334	0.319	Rocco_S11_mac.txt
4	0.263	0.482	0.255	Rocco_S21_mac.txt
5	0.238	0.364	0.399	Rocco_S22_mac.txt
6	1	0	0	Rocco_CBZ_III_nomac.txt
7	0	1	0	Rocco_SAC_pura_nomac.txt
8	0	0	1	Rocco_CBZSAC_90511_n.txt

Chapter 2

Calibration of pre-processing

Motivation

Find the pre-processing procedure which produces the best quantitative results, by using a subset of patterns with known weight fractions.

The command file

The list of commands is the following.

```
whichanalysis 4
calib 1
figpaper 1
dataType 2
range 10 50
skipdata 2
file  Rocco_S3_mac.txt
referw 0 0.565 0.435
file  Rocco_S5_mac.txt
referw 0.5 0.5 0
file  Rocco_S7_Como.txt
referw 0.5 0 0.5
file  Rocco_S11_mac.txt
referw 0.347 0.334 0.319
test
file  Rocco_S21_mac.txt
referw 0.263 0.482 0.255
test
file  Rocco_S22_mac.txt
referw 0.238 0.364 0.399
test
file  Rocco_CBZ_III_nomac.txt
referw 1 0 0
purephase
file  Rocco_SAC_pura_nomac.txt
referw 0 1 0
purephase
file  Rocco_CBZSAC_90511_n.txt
purephase
referw 0 0 1
```

The commands have been included in the demo file named *fileInputCalibratio*nPreProcessing. See the user guide for an explanation of their meaning. Input files tagged by the *test* command are those used for calibrating the pre-processing. The command *skipdata* has been added to speed up calculations.

Running RootProf

Start ROOT by clicking on his icon, or by typing "root" on a terminal window. Then write the root command:

```
Root> .x RootProf_v15.C("fileInputCalibrationPreProcessing")
```

After some seconds, graphic windows will start appearing on your screen, while text output will appear on the text window of ROOT.

If you want to store the text output on an external file, just replace the previous command with the following ones:

```
Root> .> outputCalibrationPreProcessing
.x RootProf_v15.C("fileInputCalibrationPreProcessing")
.>
```

Then the text output will be being redirected in the file named *outputCalibrationPreProcessing*. When the run ends, the root prompt will appear again on the ROOT terminal, and you are able to edit each single graphic window and read the output file by your text editor.

The graphic output

The pre-processing methods considered for automatic calibration are summarized in Table 1. All the combinations of these methods are explored during automatic calibration, for each of them the agreement factor R between calculated and known weight fraction is calculated, by considering only the mixtures tagged by the *test* command. Lower the value of R, better the pre-processing performances. Optimal pre-processing conditions are those which minimize R, and are reported in the output file, at the end of the calibration procedure. As graphic output, projections of the hypersurphace of sampled R values are produced, separately for the 4 pre-processing levels (Figs. 1-4), and for their 2D projections (Figs. 5-11)). From these figures one can monitor the efficiency of each processing method on the current dataset.

Table 1: Pre-processing values considered for automatic calibration

Pre-Processing type	0	1	2	3	4	5
Level 1: Modifications	no-modification	Smoothing	Deconvolution	Log10	Powering by 0.8	Powering by 1.2
Level 2: Rescaling	No-rescaling	Mean centering		Standard Normal Variate		
Level 3: Bakground subtraction	No-background subtraction	Clipping window=16	Clipping window=20	Clipping window=40	Clipping window=60	Clipping window= 100
Level 4: Filtering	No-filtering	Multiplicative Scatter Correction	Multiplicative Scatter Correction, all profiles	Principal component filtering	Principal component filtering, all profiles	

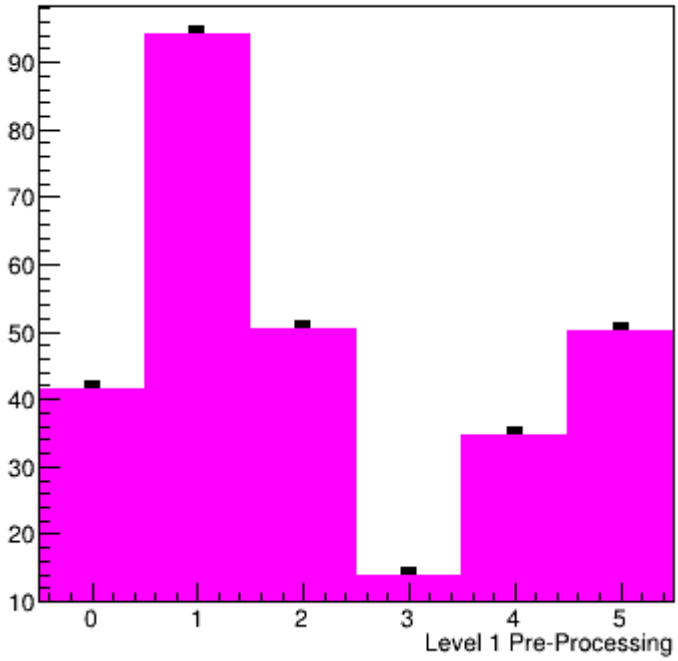

Fig. 1 Calibration of Level 1 Pre-Processing

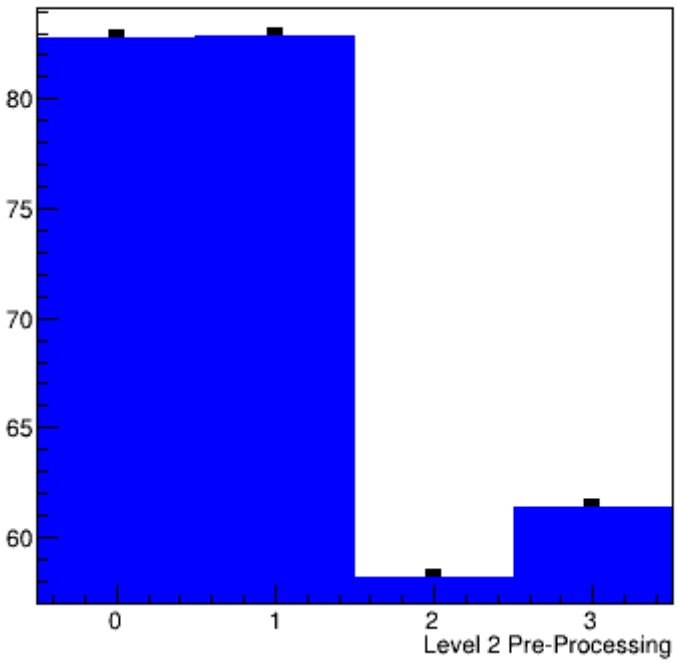

Fig. 2 Calibration of Level 2 Pre-Processing

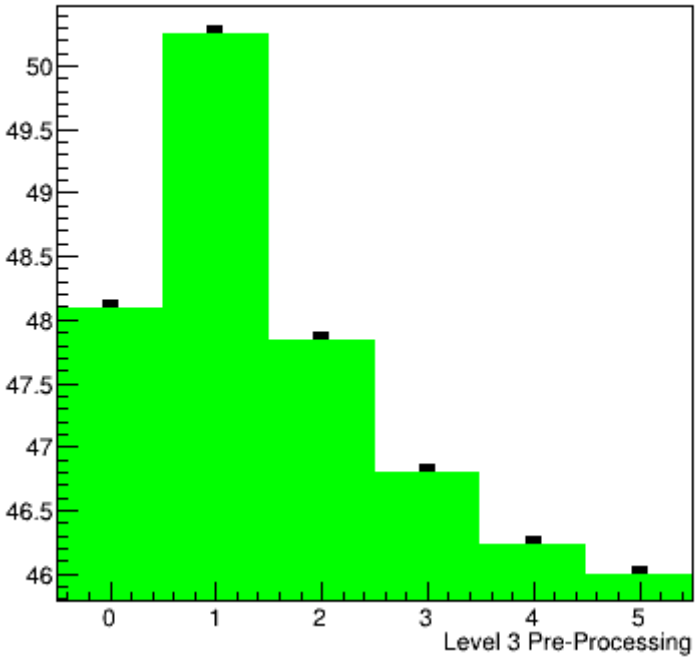

Fig. 3 Calibration of Level 3 Pre-Processing

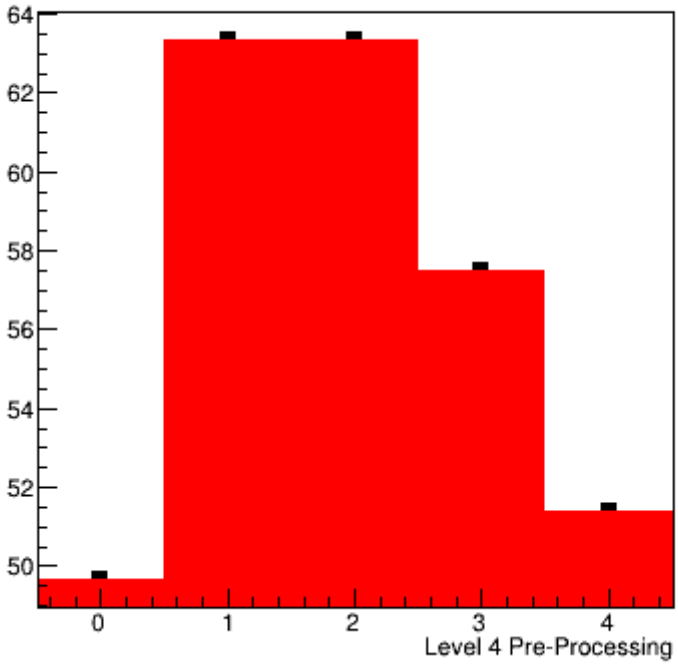

Fig. 4 Calibration of Level 4 Pre-Processing

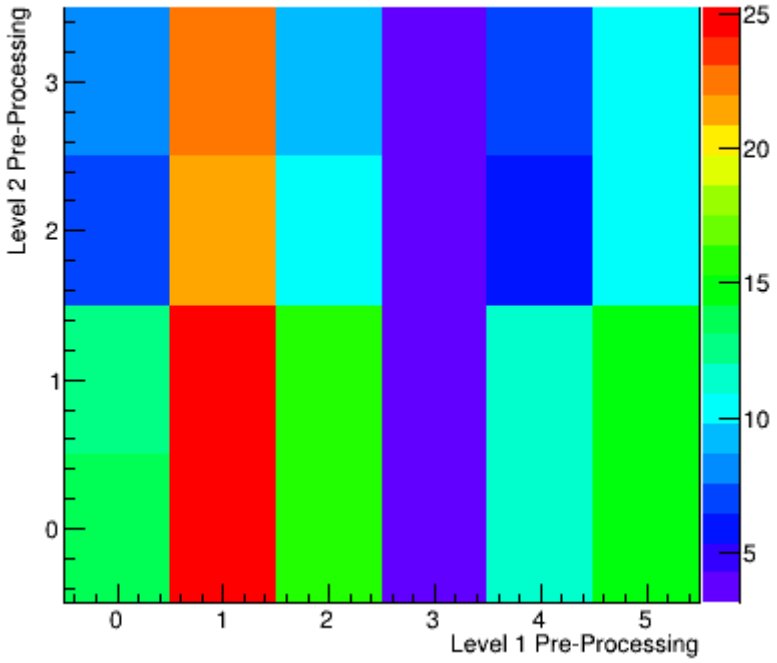

Fig. 5 Calibration of Level 1 and Level 2 Pre-Processing

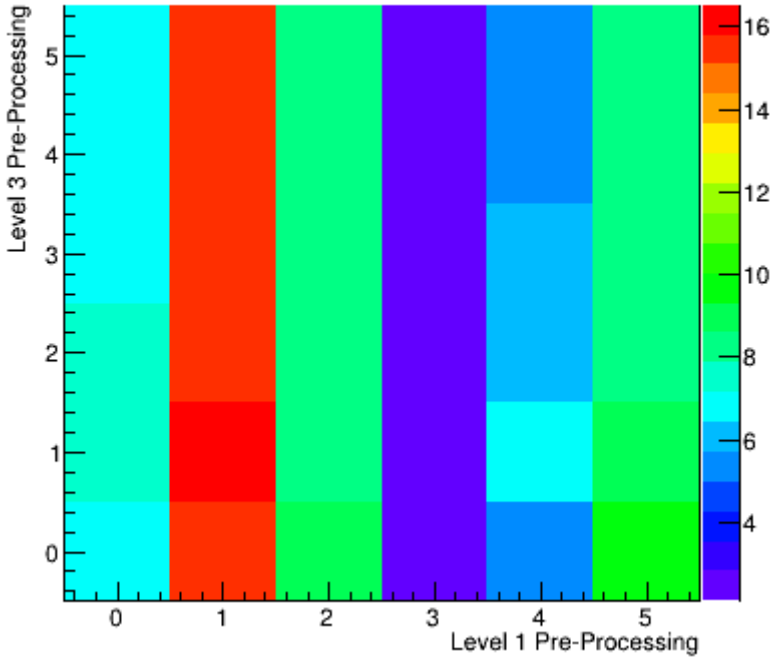

Fig. 6 Calibration of Level 1 and Level 3 Pre-Processing

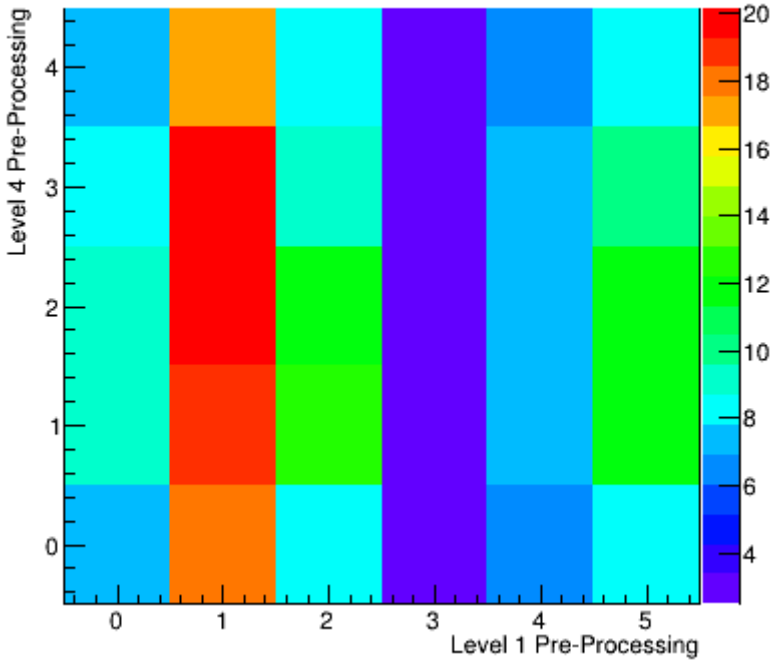

Fig. 7 Calibration of Level 1 and Level 4 Pre-Processing

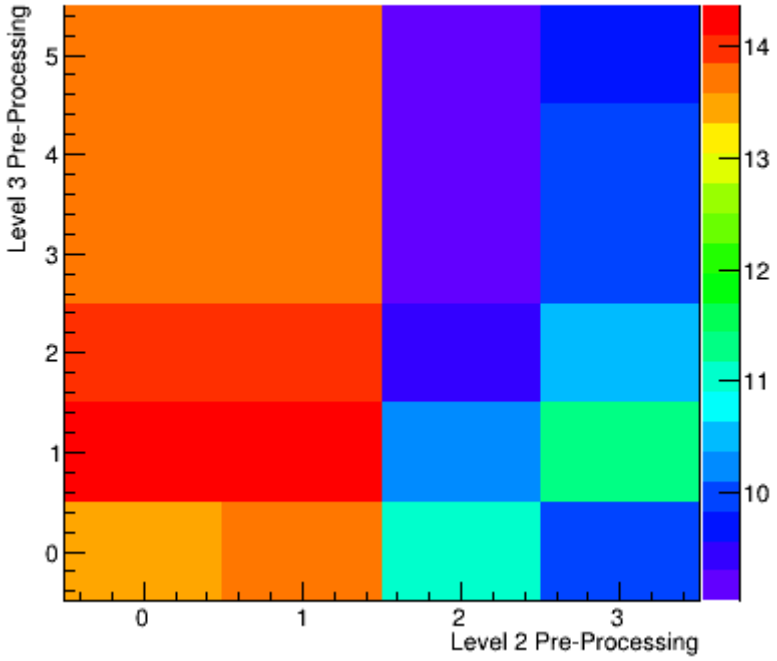

Fig. 8 Calibration of Level 2 and Level 3 Pre-Processing

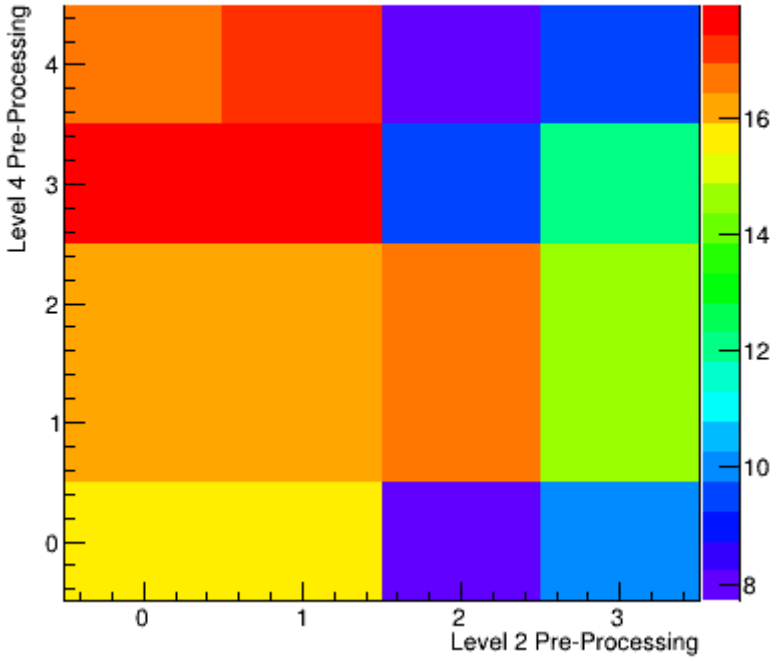

Fig. 9 Calibration of Level 2 and Level 4 Pre-Processing

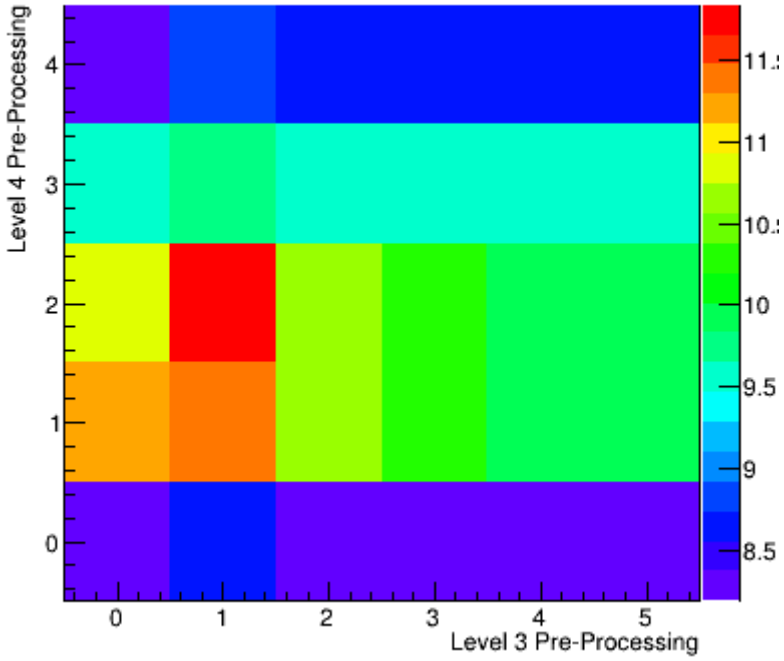

Fig. 10 Calibration of Level 3 and Level 4 Pre-Processing

After having chosen the best pre-processing strategy, which in this case is *preprocess* 3 1 48 4, the program execute a quantitative analysis on profiles processed with this strategy. In Fig.11 is reported an example of the MultiFit procedure applied to Sample 0. Note that the spectrum is

significantly different from that in Fig. 1 of chapter 2 of Quantitative tutorial, were *preprocess* 0 2 100 was applied. This is mainly due to the Log10 function applied at level 1 pre-processing.

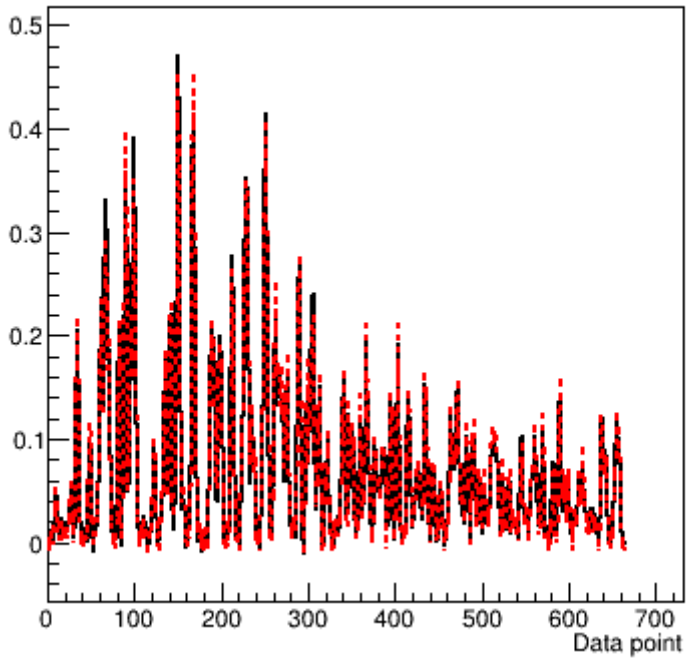

Fig. 11 MultiFit on Sample 0

The summary of the fitting procedure are shown in Fig.12. It is very similar to the True concentrations (Fig.13), more than Fig.2, chapter 2 of Quantitative tutorial, where *preprocess* 0 2 100 was applied.

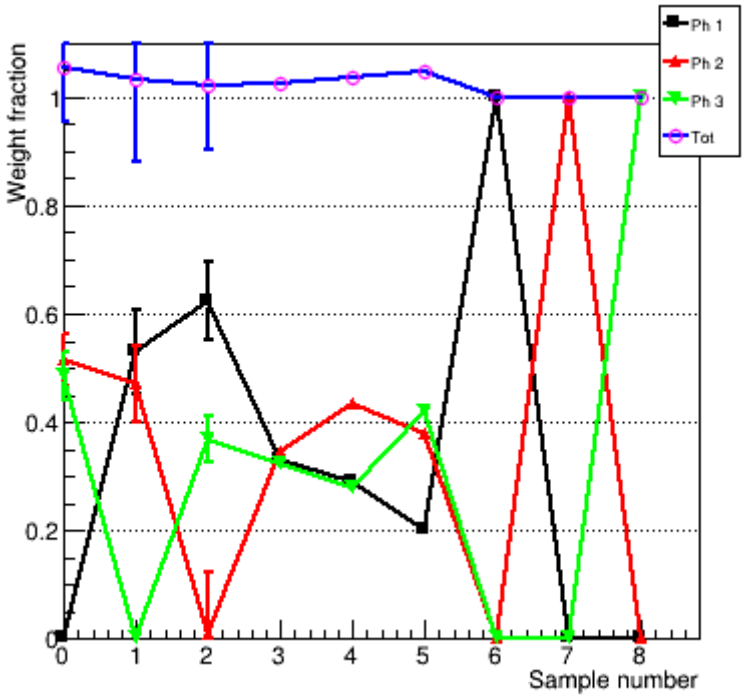

Fig. 12 Quantitative Fit graph

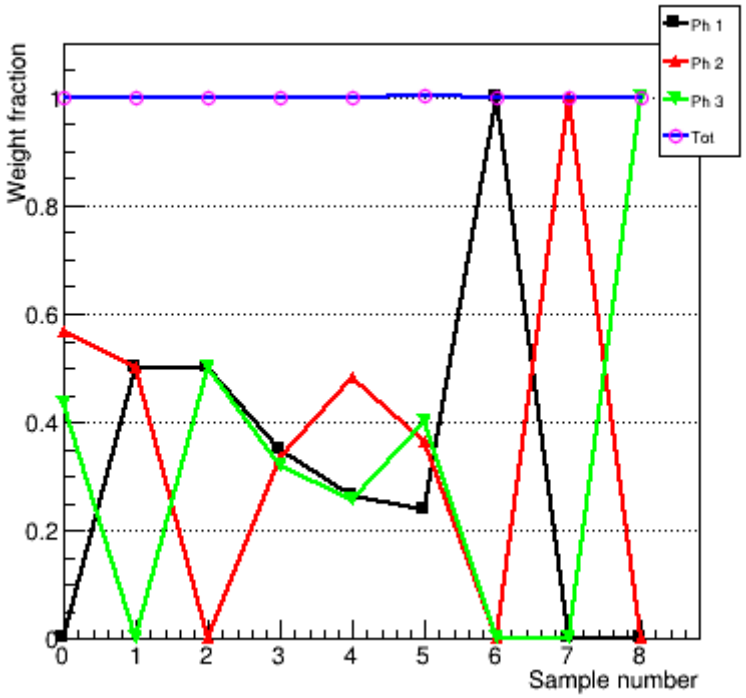

Fig. 13 Quantitative Fit graph (reference weight fractions)

The output file

The content of the output file named *outputCalibrationPreProcessing* is very long. Only the last part is reported below, which summarizes the results of the calibration procedure.

```
----------------------------
Best Pre-Processing: L1=3 L2=2 L3=16 L4=3  minimum FOM=0.074

----------------------------
```

The above section reports the results of the automatic calibration procedure on pre-processing. The last line show the pre-processing methods for which the minimum R factor is found (FOM=0.074). Thus the optimal pre-preprocessing command would be *preprocess* 3 2 16 3.

```
 Reading input files:
 ---------------------
Sample 0 -> file Rocco_S3_mac.txt
         Found 666  points
Sample 1 -> file Rocco_S5_mac.txt
         Found 666  points
Sample 2 -> file Rocco_S7_Como.txt
         Found 666  points
Sample 3 -> file Rocco_S11_mac.txt
         Found 666  points
Sample 4 -> file Rocco_S21_mac.txt
         Found 666  points
Sample 5 -> file Rocco_S22_mac.txt
         Found 666  points
Sample 6 -> file Rocco_CBZ_III_nomac.txt
         Found 666  points
Sample 7 -> file Rocco_SAC_pura_nomac.txt
         Found 666  points
Sample 8 -> file Rocco_CBZSAC_90511_n.txt
         Found 666  points
```

The above section reports the number of data points read within each input file.

```
 FIT RESULTS:
 ------------------------
 Spectrum 0: Rocco_S3_mac.txt
 Chi-Square=1.02e-01, Reduced Chi-Square=1.54e-04, NDF=662
 Weight fraction Phase 1  0.000 +- 0.097
 Weight fraction Phase 2  0.543 +- 0.004
 Weight fraction Phase 3  0.511 +- 0.005
 Total weight fraction 1.053 +- 0.097
 Rescaled weight fraction Phase 1  0.000 +- 0.000
 Rescaled weight fraction Phase 2  0.515 +- 0.048
 Rescaled weight fraction Phase 3  0.485 +- 0.045
 ------------------------
 Spectrum 1: Rocco_S5_mac.txt
 Chi-Square=1.93e-01, Reduced Chi-Square=2.92e-04, NDF=662
 Weight fraction Phase 1  0.545 +- 0.007
 Weight fraction Phase 2  0.487 +- 0.006
 Weight fraction Phase 3  0.000 +- 0.153
 Total weight fraction 1.032 +- 0.153
 Rescaled weight fraction Phase 1  0.529 +- 0.078
 Rescaled weight fraction Phase 2  0.471 +- 0.070
 Rescaled weight fraction Phase 3  0.000 +- 0.000
 ------------------------
 Spectrum 2: Rocco_S7_Como.txt
```

```
Chi-Square=2.09e-01, Reduced Chi-Square=3.16e-04, NDF=662
Weight fraction Phase 1  0.636 +- 0.008
Weight fraction Phase 2  0.010 +- 0.118
Weight fraction Phase 3  0.375 +- 0.007
Total weight fraction 1.022 +- 0.118
Rescaled weight fraction Phase 1  0.623 +- 0.072
Rescaled weight fraction Phase 2  0.010 +- 0.114
Rescaled weight fraction Phase 3  0.367 +- 0.043
------------------------
Spectrum 3: Rocco_S11_mac.txt
Chi-Square=4.65e-02, Reduced Chi-Square=7.02e-05, NDF=662
Weight fraction Phase 1  0.340 +- 0.004
Weight fraction Phase 2  0.355 +- 0.003
Weight fraction Phase 3  0.330 +- 0.003
Total weight fraction 1.024 +- 0.006
Rescaled weight fraction Phase 1  0.332 +- 0.003
Rescaled weight fraction Phase 2  0.346 +- 0.003
Rescaled weight fraction Phase 3  0.322 +- 0.003
------------------------
Spectrum 4: Rocco_S21_mac.txt
Chi-Square=5.80e-02, Reduced Chi-Square=8.76e-05, NDF=662
Weight fraction Phase 1  0.300 +- 0.004
Weight fraction Phase 2  0.450 +- 0.003
Weight fraction Phase 3  0.287 +- 0.004
Total weight fraction 1.037 +- 0.006
Rescaled weight fraction Phase 1  0.289 +- 0.003
Rescaled weight fraction Phase 2  0.434 +- 0.003
Rescaled weight fraction Phase 3  0.277 +- 0.003
------------------------
Spectrum 5: Rocco_S22_mac.txt
Chi-Square=5.64e-02, Reduced Chi-Square=8.51e-05, NDF=662
Weight fraction Phase 1  0.212 +- 0.004
Weight fraction Phase 2  0.396 +- 0.003
Weight fraction Phase 3  0.439 +- 0.004
Total weight fraction 1.048 +- 0.006
Rescaled weight fraction Phase 1  0.202 +- 0.003
Rescaled weight fraction Phase 2  0.378 +- 0.003
Rescaled weight fraction Phase 3  0.419 +- 0.003
------------------------
Spectrum 6: Rocco_CBZ_III_nomac.txt
Chi-Square=1.34e-06, Reduced Chi-Square=2.03e-09, NDF=662
Weight fraction Phase 1  1.000 +- 0.000
Weight fraction Phase 2  0.000 +- 0.000
Weight fraction Phase 3  0.000 +- 0.000
Total weight fraction 1.000 +- 0.001
Rescaled weight fraction Phase 1  1.000 +- 0.001
Rescaled weight fraction Phase 2  0.000 +- 0.000
Rescaled weight fraction Phase 3  0.000 +- 0.000
------------------------
Spectrum 7: Rocco_SAC_pura_nomac.txt
Chi-Square=5.15e-07, Reduced Chi-Square=7.78e-10, NDF=662
Weight fraction Phase 1  0.000 +- 0.000
Weight fraction Phase 2  1.000 +- 0.000
Weight fraction Phase 3  0.000 +- 0.000
Total weight fraction 1.000 +- 0.000
Rescaled weight fraction Phase 1  0.000 +- 0.000
Rescaled weight fraction Phase 2  1.000 +- 0.000
```

```
 Rescaled weight fraction Phase 3  0.000 +- 0.000
 ------------------------
 Spectrum 8: Rocco_CBZSAC_90511_n.txt
 Chi-Square=4.97e-07, Reduced Chi-Square=7.51e-10, NDF=662
 Weight fraction Phase 1  0.000 +- 0.000
 Weight fraction Phase 2  0.000 +- 0.000
 Weight fraction Phase 3  1.000 +- 0.000
 Total weight fraction 1.000 +- 0.000
 Rescaled weight fraction Phase 1  0.000 +- 0.000
 Rescaled weight fraction Phase 2  0.000 +- 0.000
 Rescaled weight fraction Phase 3  1.000 +- 0.000
  ----------------------------------
 Phase 1: RMSE=0.045 RMSE'=0.055 R2=0.983 R2(NPP)=0.959 AKLD=0.036
      Calibration fit: Chi-Square=2.40e+02 intercept=0.000+-0.000
slope=0.999+-0.001
      Limits estimates: LOD=0.04  LOQ=0.11
 NPP: Calibration fit: Chi-Square=2.16e+02 intercept=-0.008+-0.010
slope=1.000+-0.035
 NPP: Limits estimates: LOD=3.40  LOQ=10.29
  ----------------------------------
 Phase 2: RMSE=0.026 RMSE'=0.032 R2=0.995 R2(NPP)=0.986 AKLD=0.026
      Calibration fit: Chi-Square=3.24e+02 intercept=-0.003+-0.003
slope=1.002+-0.003
      Limits estimates: LOD=0.84  LOQ=2.54
 NPP: Calibration fit: Chi-Square=1.98e+01 intercept=0.162+-0.010
slope=0.569+-0.025
 NPP: Limits estimates: LOD=5.67  LOQ=17.19
  ----------------------------------
 Phase 3: RMSE=0.048 RMSE'=0.059 R2=0.975 R2(NPP)=0.869 AKLD=0.041
      Calibration fit: Chi-Square=1.19e+02 intercept=0.000+-0.000
slope=1.000+-0.000
      Limits estimates: LOD=0.05  LOQ=0.14
 NPP: Calibration fit: Chi-Square=4.28e+01 intercept=0.017+-0.009
slope=0.991+-0.028
 NPP: Limits estimates: LOD=3.04  LOQ=9.22
  ----------------------------------
 Average total weight fraction: 1.036 +- 0.036
 Overall agreement: FOM=0.119 AKLD=0.104
```

The section above shows the results of the quantitative analysis performed on all mixtures, by using the optimal pre-processing. The value of the Overall agreement is 0.119 is, lower than that found without PreProcessing calibration (0.123).

Notes

- In our experience, the best performances with X-ray powder diffraction patterns are obtained by using normalization rescaling and background subtraction at high values of clipping window (100). For IR spectra, instead, the best performances are obtained by the standard normal variate rescaling followed by background subtraction with low values of the clipping window (10).

- The calibration procedure applies to MultiFit or Unfolding techniques, depending on the settings of the *unfold* command.
- The more promising pre-processing method can be checked by using the First Sight analysis (see Qualitative analysis tutorial) with the proper *preprocessing* command.

Chapter 3

Optimal scaling of pure phase profiles

Motivation

Calibrating the MultiFit procedure by using known weight fraction for a subset of mixtures. The procedure can be seen as supervised multivariate analysis, with the subset of profiles used as calibration set.

The command file

The list of commands is the following.

```
whichanalysis 4
calib 2
figpaper 1
dataType 2
range 10 50
preprocess 0 2 100
file  Rocco_S3_mac.txt
referw 0 0.565 0.435
file  Rocco_S5_mac.txt
referw 0.5 0.5 0
file  Rocco_S7_Como.txt
referw 0.5 0 0.5
file  Rocco_S11_mac.txt
referw 0.347 0.334 0.319
test
file  Rocco_S21_mac.txt
referw 0.263 0.482 0.255
test
file  Rocco_S22_mac.txt
referw 0.238 0.364 0.399
test
file  Rocco_CBZ_III_nomac.txt
referw 1 0 0
purephase
file  Rocco_SAC_pura_nomac.txt
referw 0 1 0
purephase
file  Rocco_CBZSAC_90511_n.txt
purephase
referw 0 0 1
```

The commands have been included in the demo file named *fileInputCalibrationLSQ*. See the user guide for an explanation of their meaning. Input files tagged by the *test* command are those included in the calibration set.

Running RootProf

Start ROOT by clicking on his icon, or by typing “root” on a terminal window. Then write the root command:

```
Root> .x RootProf_v15.C(“fileInputCalibrationLSQ”)
```

or

```
Root> .> outputCalibrationLSQ
.x RootProf_v15.C(“fileInputCalibrationLSQ”)
.>
```

After some seconds, graphic windows will start appearing on your screen, while text output will appear on the terminal window, or redirected in the file named *outputCalibrationLSQ*. When the run ends, the root prompt will appear again on the ROOT terminal, and you are able to edit each single graphic window and read the output file by your text editor.

The graphic output

The graphic window in Fig.1 shows the curve which is used to obtain the estimates of the calibration parameters. The known weight fractions of profiles belonging to the calibration set (black line) are fitted by the weight fractions calculated by quantitative analysis, performed by MultiFit or Unfolding techniques, according to the value of the *unfold* command (red, dashed line). In the fitting model the pure phases are multiplied by constants, which represent the free parameters of the fitting. The best fit is thus performed on NxP points, where N is the number of profiles in the calibration set, and P is the number of pure phases. In this case N=3, P=3, and there are 9 points.

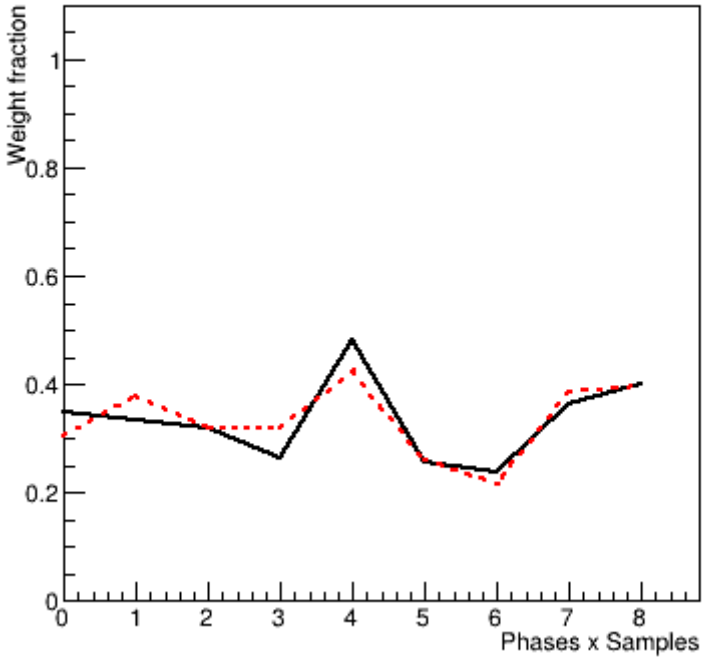

Fig. 1 Calibration by LSQ

The output file

The content of the output file named *outputCalibrationLSQ* is reported below, with comments added.

```
Input from file: fileInputCalibrationLSQ
-------------------------------
  whichanalysis 4

  calib 2

  figpaper 1

  dataType 2

  range 10 50

  preprocess 0 2 100

  file  Rocco_S3_mac.txt

  referw 0 0.565 0.435

  file  Rocco_S5_mac.txt

  referw 0.5 0.5 0

  file  Rocco_S7_Como.txt
```

```
  referw 0.5 0 0.5

  file  Rocco_S11_mac.txt

  referw 0.347 0.334 0.319

  test

  file  Rocco_S21_mac.txt

  referw 0.263 0.482 0.255

  test

  file  Rocco_S22_mac.txt

  referw 0.238 0.364 0.399

  test

  file  Rocco_CBZ_III_nomac.txt

  referw 1 0 0

  purephase

  file  Rocco_SAC_pura_nomac.txt

  referw 0 1 0

  purephase

  file  Rocco_CBZSAC_90511_n.txt

  purephase

  referw 0 0 1
```

The section above shows the commands read from the command file. It should be checked to ensure that they are interpreted correctly.

```
Starting pure phase scale calibration


 Reading input files:
 ----------------------
Sample 0 -> file Rocco_S3_mac.txt
         Found 1999  points
Sample 1 -> file Rocco_S5_mac.txt
         Found 1999  points
Sample 2 -> file Rocco_S7_Como.txt
         Found 1999  points
Sample 3 -> file Rocco_S11_mac.txt
         Found 1999  points
Sample 4 -> file Rocco_S21_mac.txt
         Found 1999  points
Sample 5 -> file Rocco_S22_mac.txt
```

```
         Found 1999  points
Sample 6 -> file Rocco_CBZ_III_nomac.txt
         Found 1999  points
Sample 7 -> file Rocco_SAC_pura_nomac.txt
         Found 1999  points
Sample 8 -> file Rocco_CBZSAC_90511_n.txt
         Found 1999  points
```

The section above reports the number of data points read within each input file.

```
------------------------
 Spectrum 3: Rocco_S11_mac.txt
 Chi-Square=3.94e+02, Reduced Chi-Square=1.98e-01, NDF=1995
 Weight fraction Phase 1  0.220 +- 0.006
 Weight fraction Phase 2  0.281 +- 0.005
 Weight fraction Phase 3  0.270 +- 0.008
 Total weight fraction 0.771 +- 0.011
 Rescaled weight fraction Phase 1  0.286 +- 0.007
 Rescaled weight fraction Phase 2  0.364 +- 0.006
 Rescaled weight fraction Phase 3  0.350 +- 0.007
 ------------------------
 Spectrum 4: Rocco_S21_mac.txt
 Chi-Square=6.86e+02, Reduced Chi-Square=3.44e-01, NDF=1995
 Weight fraction Phase 1  0.257 +- 0.008
 Weight fraction Phase 2  0.347 +- 0.006
 Weight fraction Phase 3  0.243 +- 0.010
 Total weight fraction 0.846 +- 0.014
 Rescaled weight fraction Phase 1  0.303 +- 0.008
 Rescaled weight fraction Phase 2  0.410 +- 0.008
 Rescaled weight fraction Phase 3  0.287 +- 0.009
 ------------------------
 Spectrum 5: Rocco_S22_mac.txt
 Chi-Square=3.83e+02, Reduced Chi-Square=1.92e-01, NDF=1995
 Weight fraction Phase 1  0.187 +- 0.006
 Weight fraction Phase 2  0.340 +- 0.005
 Weight fraction Phase 3  0.398 +- 0.007
 Total weight fraction 0.925 +- 0.011
 Rescaled weight fraction Phase 1  0.202 +- 0.006
 Rescaled weight fraction Phase 2  0.368 +- 0.005
 Rescaled weight fraction Phase 3  0.430 +- 0.006
 ------------------------
```

The section above reports the results of a first run of quantitative analysis, performed without special rescaling pure phase profiles on the profiles belonging to the calibration set. The obtained weight fraction represents the starting point for the calibration procedure.

```
Calibration fit results:
 ------------------------------
 Chi-Square=0.012  Reduced Chi-Square=0.0019 NDF=6
 Phase 0: Best fit coefficient= 1.00 +- 58.48
 Phase 1: Best fit coefficient= 1.02 +- 59.26
 Phase 2: Best fit coefficient= 1.16 +- 64.43
```

The section above report the result of the best fit applied on the true weight fractions of the profiles belonging to the calibration set (see Fig.1). The coefficients so obtained are the scaling factors to be applied to pure phase profiles.

```
FIT RESULTS:
 ------------------------
 Spectrum 0: Rocco_S3_mac.txt
 Chi-Square=4.23e+02, Reduced Chi-Square=2.12e-01, NDF=1995
 Weight fraction Phase 1  0.043 +- 0.006
 Weight fraction Phase 2  0.382 +- 0.005
 Weight fraction Phase 3  0.354 +- 0.007
 Total weight fraction 0.779 +- 0.010
 Rescaled weight fraction Phase 1  0.055 +- 0.008
 Rescaled weight fraction Phase 2  0.491 +- 0.007
 Rescaled weight fraction Phase 3  0.454 +- 0.007
 ------------------------
 Spectrum 1: Rocco_S5_mac.txt
 Chi-Square=1.28e+03, Reduced Chi-Square=6.42e-01, NDF=1995
 Weight fraction Phase 1  0.342 +- 0.011
 Weight fraction Phase 2  0.339 +- 0.008
 Weight fraction Phase 3  0.019 +- 0.012
 Total weight fraction 0.700 +- 0.018
 Rescaled weight fraction Phase 1  0.488 +- 0.013
 Rescaled weight fraction Phase 2  0.485 +- 0.013
 Rescaled weight fraction Phase 3  0.027 +- 0.016
 ------------------------
 Spectrum 2: Rocco_S7_Como.txt
 Chi-Square=6.49e+02, Reduced Chi-Square=3.25e-01, NDF=1995
 Weight fraction Phase 1  0.374 +- 0.008
 Weight fraction Phase 2  0.053 +- 0.006
 Weight fraction Phase 3  0.259 +- 0.008
 Total weight fraction 0.686 +- 0.013
 Rescaled weight fraction Phase 1  0.546 +- 0.010
 Rescaled weight fraction Phase 2  0.078 +- 0.008
 Rescaled weight fraction Phase 3  0.377 +- 0.009
 ------------------------
 Spectrum 3: Rocco_S11_mac.txt
 Chi-Square=3.94e+02, Reduced Chi-Square=1.98e-01, NDF=1995
 Weight fraction Phase 1  0.220 +- 0.006
 Weight fraction Phase 2  0.275 +- 0.005
 Weight fraction Phase 3  0.233 +- 0.006
 Total weight fraction 0.728 +- 0.010
 Rescaled weight fraction Phase 1  0.303 +- 0.007
 Rescaled weight fraction Phase 2  0.378 +- 0.006
 Rescaled weight fraction Phase 3  0.320 +- 0.007
 ------------------------
 Spectrum 4: Rocco_S21_mac.txt
 Chi-Square=6.86e+02, Reduced Chi-Square=3.44e-01, NDF=1995
 Weight fraction Phase 1  0.257 +- 0.008
 Weight fraction Phase 2  0.339 +- 0.006
 Weight fraction Phase 3  0.209 +- 0.009
 Total weight fraction 0.805 +- 0.013
 Rescaled weight fraction Phase 1  0.319 +- 0.008
 Rescaled weight fraction Phase 2  0.422 +- 0.008
 Rescaled weight fraction Phase 3  0.260 +- 0.009
```

```
 ------------------------
 Spectrum 5: Rocco_S22_mac.txt
 Chi-Square=3.83e+02, Reduced Chi-Square=1.92e-01, NDF=1995
 Weight fraction Phase 1  0.187 +- 0.006
 Weight fraction Phase 2  0.333 +- 0.005
 Weight fraction Phase 3  0.343 +- 0.006
 Total weight fraction 0.862 +- 0.010
 Rescaled weight fraction Phase 1  0.216 +- 0.006
 Rescaled weight fraction Phase 2  0.386 +- 0.005
 Rescaled weight fraction Phase 3  0.397 +- 0.006
 ------------------------
 Spectrum 6: Rocco_CBZ_III_nomac.txt
 Chi-Square=1.16e-06, Reduced Chi-Square=5.81e-10, NDF=1995
 Weight fraction Phase 1  1.000 +- 0.000
 Weight fraction Phase 2  0.000 +- 0.000
 Weight fraction Phase 3  0.000 +- 0.000
 Total weight fraction 1.000 +- 0.000
 Rescaled weight fraction Phase 1  1.000 +- 0.000
 Rescaled weight fraction Phase 2  0.000 +- 0.000
 Rescaled weight fraction Phase 3  0.000 +- 0.000
 ------------------------
 Spectrum 7: Rocco_SAC_pura_nomac.txt
 Chi-Square=1.51e-06, Reduced Chi-Square=7.55e-10, NDF=1995
 Weight fraction Phase 1  0.000 +- 0.000
 Weight fraction Phase 2  1.000 +- 0.000
 Weight fraction Phase 3  0.000 +- 0.000
 Total weight fraction 1.000 +- 0.000
 Rescaled weight fraction Phase 1  0.000 +- 0.000
 Rescaled weight fraction Phase 2  1.000 +- 0.000
 Rescaled weight fraction Phase 3  0.000 +- 0.000
 ------------------------
 Spectrum 8: Rocco_CBZSAC_90511_n.txt
 Chi-Square=3.58e-06, Reduced Chi-Square=1.80e-09, NDF=1995
 Weight fraction Phase 1  0.000 +- 0.000
 Weight fraction Phase 2  0.000 +- 0.000
 Weight fraction Phase 3  1.000 +- 0.000
 Total weight fraction 1.000 +- 0.000
 Rescaled weight fraction Phase 1  0.000 +- 0.000
 Rescaled weight fraction Phase 2  0.000 +- 0.000
 Rescaled weight fraction Phase 3  1.000 +- 0.000
```

The section above reports the results of a second run of quantitative analysis, performed by rescaling pure phase profiles with the previously determined coefficient.

```
   ------------------------------------
 Phase 1: RMSE=0.035 RMSE'=0.036 R2=0.988 R2(NPP)=0.945 AKLD=0.029
      Calibration fit: Chi-Square=1.74e+02 intercept=0.012+-0.004
slope=0.988+-0.004
      Limits estimates: LOD=1.40  LOQ=4.26
 NPP: Calibration fit: Chi-Square=1.57e+02 intercept=0.031+-0.006
slope=0.906+-0.021
 NPP: Limits estimates: LOD=2.35  LOQ=7.11
  ------------------------------------
 Phase 2: RMSE=0.045 RMSE'=0.044 R2=0.982 R2(NPP)=0.967 AKLD=0.037
```

```
      Calibration fit: Chi-Square=3.31e+02 intercept=0.022+-0.004
slope=0.977+-0.004
      Limits estimates: LOD=1.45  LOQ=4.40
 NPP: Calibration fit: Chi-Square=8.30e+01 intercept=0.109+-0.007
slope=0.715+-0.017
 NPP: Limits estimates: LOD=3.23  LOQ=9.79
  -----------------------------------
 Phase 3: RMSE=0.043 RMSE'=0.051 R2=0.981 R2(NPP)=0.913 AKLD=0.028
      Calibration fit: Chi-Square=1.85e+02 intercept=-0.008+-0.005
slope=1.008+-0.005
      Limits estimates: LOD=1.59  LOQ=4.81
 NPP: Calibration fit: Chi-Square=1.41e+02 intercept=0.062+-0.012
slope=0.809+-0.031
 NPP: Limits estimates: LOD=4.77  LOQ=14.44
  -----------------------------------
 Average total weight fraction: 0.760 +- 0.005
 Overall agreement: FOM=0.122 AKLD=0.094
```

The above section includes the comparison between the phase abundances estimated by the calibrated RootProf and the reference ones. The quantities reported are the same as those described in Chapter 2 of the Quantitative tutorial. As expected. the overall FOM and AKLD are lower than that obtained by the unsupervised MultiFit (0.123 and 0.108, respectively).

Notes

- By considering results reported on Table 1, it results that mixtures with non-zero phase abundances (3,4,5) are particularly suitable for calibration purposes. Binary mixtures (0,1,2) have lower performances.
- Pure phase profiles (6,7,8) give very poor calibration performances. A calibration set constituted by pure phases only should not be used. It does not add any new information with respect to the unsupervised case.
- The calibration procedure applies to MultiFit or Unfolding techniques, depending on the settings of the *unfold* command.
- The calibration parameters can be view as a multiplicative rescaling of the pure phase profiles. In case of powder diffraction profiles, however, they can also be considered as a multiplicative correction of the mass absorption coefficients (inserted via the *rhoin* command).

Table 1: Results of the calibration procedure by using different compositions of the calibration set.

Samples in Calibration Set	**Calibration Parameters**	**Overall AKLD**

--	1.00, 1.00, 1.00	0.108
0,1,2	1.00, 0.88, 0.98	0.104
3,4,5	1.00, 1.02, 1.16	0.094
0,1,2,3,4,5	1.00, 0.93, 1.05	0.097
6,7,8	1.00, 1.00, 1.00	0.108
0,1,2,6,7,8	1.00, 0.87, 0.97	0.104
3,4,5,6,7,8	1.00, 1.02, 1.16	0.094
0,1,2,3,4,5,6,7,8	1.00, 0.93, 1.05	0.097

Chapter 4

Supervised quantitative analysis

Motivation

Perform a complete run of supervised quantitative analysis, by executing an automatic calibration of pre-processing and a best fit determination of pure phase rescaling on a subset of mixtures with known weight fractions (calibration set).

The command file

The list of commands is the following.

```
Whichanalysis 4
calib 0
figpaper 1
dataType 2
range 10 50
preprocess 0 2 100
file  Rocco_S3_mac.txt
referw 0 0.565 0.435
file  Rocco_S5_mac.txt
referw 0.5 0.5 0
file  Rocco_S7_Como.txt
referw 0.5 0 0.5
file  Rocco_S11_mac.txt
referw 0.347 0.334 0.319
test
file  Rocco_S21_mac.txt
referw 0.263 0.482 0.255
test
file  Rocco_S22_mac.txt
referw 0.238 0.364 0.399
test
file  Rocco_CBZ_III_nomac.txt
referw 1 0 0
purephase
file  Rocco_SAC_pura_nomac.txt
referw 0 1 0
purephase
file  Rocco_CBZSAC_90511_n.txt
purephase
referw 0 0 1
```

The commands have been included in the demo file named *fileInputSupervised*. See the user guide for an explanation of their meaning. Input files tagged by the *test* command are those included in the calibration set.

Running RootProf

Start ROOT by clicking on his icon, or by typing "root" on a terminal window. Then write the root command:

Root> .x RootProf_v15.C("fileInputSupervised")

or

Root> .>outputSupervised

.x RootProf_v15.C("fileInputSupervised")

.>

After some seconds, graphic windows will start appearing on your screen, while text output will appear on the terminal window, or redirected in the file named *outputSupervised*. When the run ends, the root prompt will appear again on the ROOT terminal, and you are able to edit each single graphic window and read the output file by your text editor.

The graphic output

The graphic output is composed by those seen in chapters 2 and 3, plus the standard output for quantitative analysis (see Quantitative analysis tutorial).

The output file

The output file is very long. It consists of the output file listed in chapter 2, followed by:

```
Calibration fit results:
 ------------------------------
 Chi-Square=0.0048  Reduced Chi-Square=0.0008 NDF=6
 Phase 0: Best fit coefficient= 1.00 +-  0.05
 Phase 1: Best fit coefficient= 0.99 +-  0.07
 Phase 2: Best fit coefficient= 1.06 +-  0.06
```

The section above shows the results of the best fit performed on known weight fractions, to determine the best estimates of the pure phase rescaling constant (the best fit coefficients).

```
FIT RESULTS:
 ------------------------
 Spectrum 0: Rocco_S3_mac.txt
 Chi-Square=1.02e-01, Reduced Chi-Square=1.54e-04, NDF=662
 Weight fraction Phase 1  0.000 +- 0.097
 Weight fraction Phase 2  0.545 +- 0.004
```

```
Weight fraction Phase 3  0.480 +- 0.004
Total weight fraction 1.025 +- 0.097
Rescaled weight fraction Phase 1  0.000 +- 0.000
Rescaled weight fraction Phase 2  0.532 +- 0.050
Rescaled weight fraction Phase 3  0.468 +- 0.044
------------------------
Spectrum 1: Rocco_S5_mac.txt
Chi-Square=1.93e-01, Reduced Chi-Square=2.92e-04, NDF=662
Weight fraction Phase 1  0.545 +- 0.007
Weight fraction Phase 2  0.489 +- 0.006
Weight fraction Phase 3  0.000 +- 0.149
Total weight fraction 1.034 +- 0.150
Rescaled weight fraction Phase 1  0.527 +- 0.076
Rescaled weight fraction Phase 2  0.473 +- 0.068
Rescaled weight fraction Phase 3  0.000 +- 0.000
------------------------
Spectrum 2: Rocco_S7_Como.txt
Chi-Square=2.09e-01, Reduced Chi-Square=3.16e-04, NDF=662
Weight fraction Phase 1  0.637 +- 0.008
Weight fraction Phase 2  0.010 +- 0.113
Weight fraction Phase 3  0.352 +- 0.006
Total weight fraction 0.999 +- 0.113
Rescaled weight fraction Phase 1  0.637 +- 0.072
Rescaled weight fraction Phase 2  0.010 +- 0.112
Rescaled weight fraction Phase 3  0.353 +- 0.040
------------------------
Spectrum 3: Rocco_S11_mac.txt
Chi-Square=4.65e-02, Reduced Chi-Square=7.02e-05, NDF=662
Weight fraction Phase 1  0.340 +- 0.004
Weight fraction Phase 2  0.356 +- 0.003
Weight fraction Phase 3  0.310 +- 0.003
Total weight fraction 1.006 +- 0.006
Rescaled weight fraction Phase 1  0.338 +- 0.003
Rescaled weight fraction Phase 2  0.354 +- 0.003
Rescaled weight fraction Phase 3  0.308 +- 0.003
------------------------
Spectrum 4: Rocco_S21_mac.txt
Chi-Square=5.80e-02, Reduced Chi-Square=8.76e-05, NDF=662
Weight fraction Phase 1  0.300 +- 0.004
Weight fraction Phase 2  0.452 +- 0.003
Weight fraction Phase 3  0.270 +- 0.003
Total weight fraction 1.022 +- 0.006
Rescaled weight fraction Phase 1  0.293 +- 0.003
Rescaled weight fraction Phase 2  0.443 +- 0.003
Rescaled weight fraction Phase 3  0.264 +- 0.003
------------------------
Spectrum 5: Rocco_S22_mac.txt
Chi-Square=5.64e-02, Reduced Chi-Square=8.51e-05, NDF=662
Weight fraction Phase 1  0.212 +- 0.004
Weight fraction Phase 2  0.398 +- 0.003
Weight fraction Phase 3  0.413 +- 0.003
Total weight fraction 1.023 +- 0.006
Rescaled weight fraction Phase 1  0.207 +- 0.003
Rescaled weight fraction Phase 2  0.389 +- 0.003
Rescaled weight fraction Phase 3  0.404 +- 0.003
------------------------
Spectrum 6: Rocco_CBZ_III_nomac.txt
```

```
 Chi-Square=8.77e-07, Reduced Chi-Square=1.32e-09, NDF=662
 Weight fraction Phase 1  1.000 +- 0.000
 Weight fraction Phase 2  0.000 +- 0.000
 Weight fraction Phase 3  0.000 +- 0.000
 Total weight fraction 1.000 +- 0.000
 Rescaled weight fraction Phase 1  1.000 +- 0.000
 Rescaled weight fraction Phase 2  0.000 +- 0.000
 Rescaled weight fraction Phase 3  0.000 +- 0.000
 ------------------------
 Spectrum 7: Rocco_SAC_pura_nomac.txt
 Chi-Square=9.84e-07, Reduced Chi-Square=1.49e-09, NDF=662
 Weight fraction Phase 1  0.000 +- 0.000
 Weight fraction Phase 2  1.000 +- 0.000
 Weight fraction Phase 3  0.000 +- 0.000
 Total weight fraction 1.001 +- 0.000
 Rescaled weight fraction Phase 1  0.000 +- 0.000
 Rescaled weight fraction Phase 2  1.000 +- 0.000
 Rescaled weight fraction Phase 3  0.000 +- 0.000
 ------------------------
 Spectrum 8: Rocco_CBZSAC_90511_n.txt
 Chi-Square=2.00e-07, Reduced Chi-Square=3.03e-10, NDF=662
 Weight fraction Phase 1  0.000 +- 0.000
 Weight fraction Phase 2  0.000 +- 0.000
 Weight fraction Phase 3  1.000 +- 0.000
 Total weight fraction 1.000 +- 0.000
 Rescaled weight fraction Phase 1  0.000 +- 0.000
 Rescaled weight fraction Phase 2  0.000 +- 0.000
 Rescaled weight fraction Phase 3  1.000 +- 0.000
  ------------------------------------
 Phase 1: RMSE=0.049 RMSE'=0.060 R2=0.981 R2(NPP)=0.955 AKLD=0.037
      Calibration fit: Chi-Square=2.09e+02 intercept=0.000+-0.000
slope=0.999+-0.000
      Limits estimates: LOD=0.04  LOQ=0.13
 NPP: Calibration fit: Chi-Square=2.05e+02 intercept=-0.007+-0.010
slope=1.012+-0.036
 NPP: Limits estimates: LOD=3.39  LOQ=10.29
  ------------------------------------
 Phase 2: RMSE=0.022 RMSE'=0.027 R2=0.995 R2(NPP)=0.986 AKLD=0.025
      Calibration fit: Chi-Square=3.05e+02 intercept=0.012+-0.003
slope=0.988+-0.003
      Limits estimates: LOD=0.86  LOQ=2.62
 NPP: Calibration fit: Chi-Square=2.70e+01 intercept=0.171+-0.010
slope=0.570+-0.025
 NPP: Limits estimates: LOD=5.72  LOQ=17.34
  ------------------------------------
 Phase 3: RMSE=0.051 RMSE'=0.062 R2=0.974 R2(NPP)=0.868 AKLD=0.038
      Calibration fit: Chi-Square=4.61e+01 intercept=0.000+-0.000
slope=1.000+-0.000
      Limits estimates: LOD=0.06  LOQ=0.19
 NPP: Calibration fit: Chi-Square=4.47e+01 intercept=0.010+-0.009
slope=0.970+-0.027
 NPP: Limits estimates: LOD=3.04  LOQ=9.21
  ------------------------------------
 Average total weight fraction: 1.018 +- 0.035
 Overall agreement: FOM=0.122 AKLD=0.100
```

This section report the final result of the quantitative analysis, performed on all profiles by using the optimal pre-processing and pure phase rescaling. Note that the overall agreement (0.122) is lower than that obtained by unsupervised analysis (0.123, see chapter 3 of the tutorial on quantitative analysis).

Notes

- Supervised quantitative analysis could be very time consuming. You can speed up calculations by using the Unfolding approach (command *unfold 1*), or by skipping some data points (command *skipdata* 10).
- By using the commands *skipdata* or *unfold*, calculations are much faster, and probably a different pre-processing scheme and best fit coefficients could be obtained. The final overall results will be however only slightly different.
- Given the long computations, it would be advisable to run RootProf in batch mode by using the linux command

  root.exe -l -b -q 'RootProf_v15.C("fileInput")' > outputFile

  when using the command *calib* 1. In this case the command *savefig* 1 or 2 should be added to the command file, to save figures for later inspection.