elementR package

elementR[1] is an R package facilitating the reduction of elemental microchemistry data from solid-phase LA-ICPMS analysis (laser ablation inductive coupled plasma mass spectrometry). The **elementR** package provides a reactive and user friendly interface for conducting all steps needed for an optimal data reduction while leaving maximum control for user.

Documentation Summary

1. How to start ?
2. What is the structure of the elementR application ?
3. What are the data to provide ?
4. What contain the different steps of the elementR reduction procedure ?
    - Step 1: Building the project -
    - Step 2: Filtration of the standard replicates -
    - Step 3: Verification of the machine drift -
    - Step 4: Filtration of the sample replicates -
    - Step 5: Averaging of the sample replicates -
    - Additional step: Optional settings -
5. What is exactly happening when elementR performs the machine drift verification and correction ?
6. What are the exported data from elementR ?
7. More questions ?

1 To cite this package:

Charlotte Sirot, Francois Guilhaumon, Franck Ferraton, Audrey Darnaude, Jacques Panfili and Amber Child (2016). elementR: A shiny application for reducing elemental LA-ICPMS data from solid structures. R package version 1.0. https://CRAN.R-project.org/package=elementR

1. How to start ?

To install and load the **elementR** package, open your usual R environment (R, Rstudio...) and run the following instructions:

```
install.packages("elementR", dependencies = T)
library(elementR)
```

The only function required to perform the data reduction is the `runElementR()` function. To launch the data reduction procedure, just run:

```
runElementR()
```

At this point, the web browser opens a new page with the operational application based on the **elementR** code.

IMPORTANT: To be installed and run properly, the **elementR** package needs an up-to-date version of the R software (at least ≥ 3.2.3, to update your version click here)

2. What is the structure of the elementR application ?

The data reduction proceeded by the **elementR** package is organized in a session framework including five main steps (indicate on the left side bar of the opened page of the web browser):

→ Step 1. The setting of the main parameters of the procedure (see step1)
→ Step 2. The filtration of the standards (see step2)
→ Step 3. The machine drift verification & correction (see step3)
→ Step 4. The filtration of the samples (see step4)
→ Step 5. The sample replicate averaging (see step5)

Note that these steps have to be carried out in the order mentioned above, the third step, for instance, being not available until the first and second one are validated.

The left side bar allows to navigate between these steps and to know exactly which part of the procedure is currently running, validates or remains to complete.

The two last tabs "Configuration"and "Source code for app" are not part of the data reduction procedure but bring additional information detailed later in this document (see additional settings).

3. What are the data to provide ?

Running **elementR** code requires to provide appropriate data in specified format and organized in a session framework.

- What data ? -

→ a set of data from standard analysis (at least one). **elementR** is compatible with all standard types (NIST 612, NIST 610, MACS...). However, standards of a single session must have all the same type (that is the reason we call them **standard replicates** and that they are stored **in a same folder**).

→ a calibration file providing for each investigated chemical element:

$$\text{Calib}_X = \frac{[X]_{calibType}}{[Internal\ standard]_{calibType}}$$

where:

- Calib_X is the value to include in the calibration file for the chemical element X
- $[X]_{calibType}$ is the concentration of the chemical element X contained in the calibration material
- $[Internal\ standard]_{calibType}$ is the concentration of the internal standard element contained in the calibration material

→ a set of data from sample analysis (at least one, obviously :D). Sample may have one or more replicates, these replicates being averaged in the last step.

- What data format ? -

→ The format of data compatible with the **elementR** application are those from worksheet from Excel (.xls, .xlsx), LibreOffice (.ods) or text (.csv).

- For the text format, the separator used by default is the semi-colon. However, you can change this parameter before uploading your project in the configuration tab.
- For any format (except for text format), the decimal has to be indicated by a point (for text format, you can however customize the decimal in the configuration tab before uploading your project)
- For Excel format, **elementR** reads the first worksheet of the file. For OpenOffice format, you indicate the sheet to import by calling it "data"

→ Sample and standard data must be all organized in the same way (fig. 1):

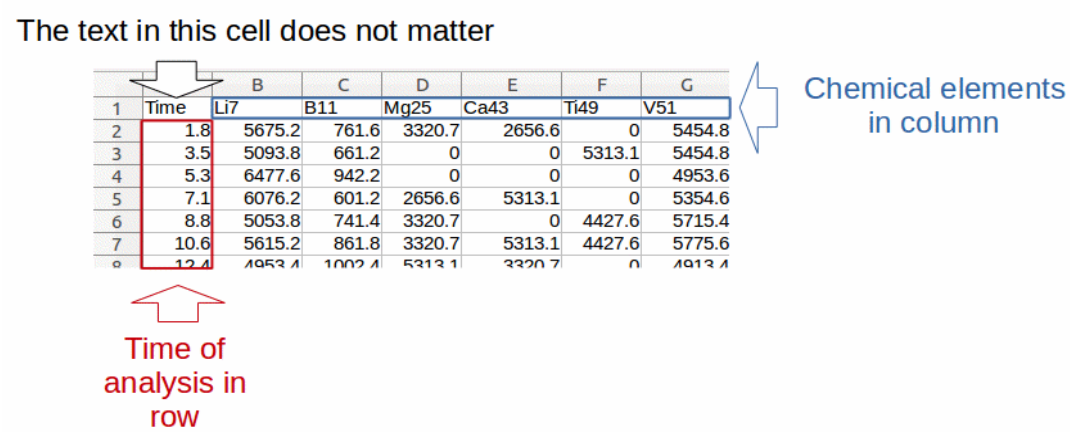
- In column 1, the time of successive analysis
- In the following columns, the chemical elements with element names as column heads

- All data must contain the same chemical elements in the same order
- Data must contain only numerical characters (except for the names of the columns)

Note that:

In any case: when you upload your data, **elementR** checks the validity of these points and indicates the problem if any.

Some compatible format of standard and sample data are provided with the **elementR** package, take a look ! (in the folder where R packages are installed[2], open the **elementR** folder and look at the “Example_Session”, you will find sample and standard examples of data format).



		B	C	D	E	F	G
1	Time	Li7	B11	Mg25	Ca43	Ti49	V51
2	1.8	5675.2	761.6	3320.7	2656.6	0	5454.8
3	3.5	5093.8	661.2	0	0	5313.1	5454.8
4	5.3	6477.6	942.2	0	0	0	4953.6
5	7.1	6076.2	601.2	2656.6	5313.1	0	5354.6
6	8.8	5053.8	741.4	3320.7	0	4427.6	5715.4
7	10.6	5615.2	861.8	3320.7	5313.1	4427.6	5775.6
8	12.4	4953.4	1002.4	5313.1	3320.7	0	4913.4

Figure 1: example of data format compatible with **elementR** application for sample and standard replicate data: chemical elements in column (in blue) with their name at the top of each column, time of analysis in row (in red). The text included in the cell at the top of the analysis time does not matter.

→ The file comprising the calibration data must have its own organization (fig. 2):

- In columns, the chemical elements with their name at the head of each column
- The text in the cells of the first column does not matter

- The data of the calibration file must contain the same chemical elements in the same order than in the standard and sample files
- Data must contain only numerical characters (except for the first column and the names of the chemical elements)

Note that:

In any case: when you upload your data, **elementR** checks the validity of these points and indicates the problem if any.

Some compatible format of calibration data are provided with the **elementR** package, take a look ! (in the folder where R packages are installed[2], reach the **elementR** folder and look at the “Example_Session”, you will find the file called “Calibration_File_NIST612.csv”).

2 If you do not manage to find this folder on your computer, you can find its path by running:

```
install.packages("devtools", dependencies = T)
library(devtools)
filePath <- system.file("", package="elementR")
```

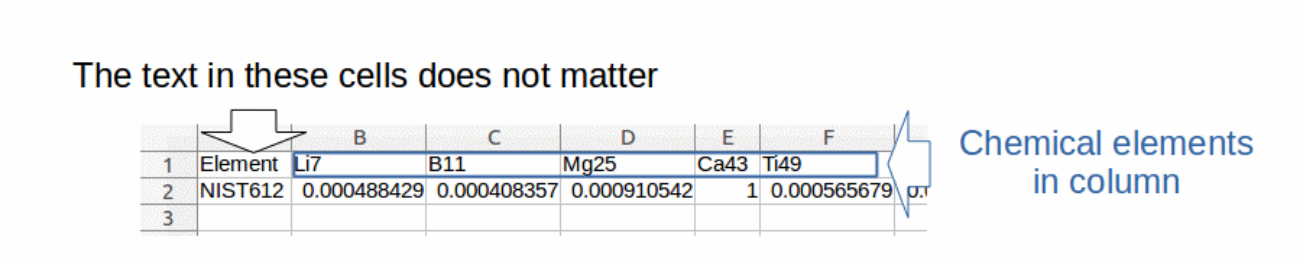


Figure 2: example of data structure compatible with **elementR** application for calibration data: chemical elements in column (in blue) with their name at the top of each column. The text included in the cells of the first column does not matter.

- How organize data ? -

→ The data must be organized in a session framework (Fig 3):

- All file corresponding to standard replicates must be included in the same folder called “standards” (= batch of standard replicates)
- All samples must be included in the same folder called “samples”
- All sample replicates of the same sample (even if there is only a single replicate) must be included in a folder with the name of the considered sample (= a batch of sample replicates)
- The “standards” and the “samples” sub-folders have to be included in a folder called with the name of the project

Note that:

In any case: when you upload your data, **elementR** checks the validity of these points and indicates the problem if any.

A compatible format of session is provided with the **elementR** package, have a look ! (in the folder where R packages are installed, reach the **elementR** folder[3] and look at the “Example_Session” organization).

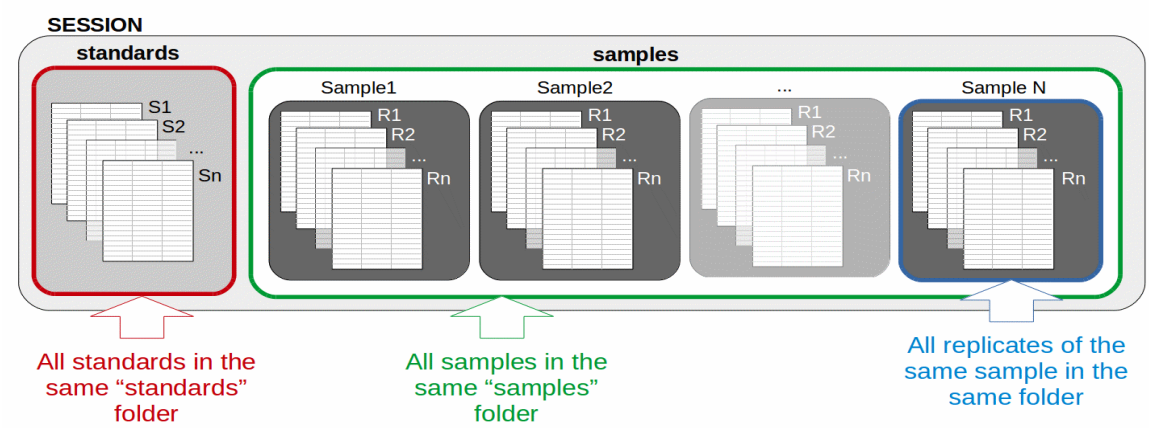


Figure 3: example of session organization compatible with **elementR** application: all standards (S1, S2, Sn) are in a “standards” folder (in red), all samples are in a “samples” folder (in green). Each sample replicate R1, R2, Rn (even if there is only a single replicate) must be included in a sub-folder with the name of the sample (in blue). The name of the sample sub-folders and of the standard and sample replicates do not matter.

3 if you do not manage to find this folder on your computer, see here

4. What contain the different steps of a elementR reduction procedure ?

- Step 1: Building the project -

In this first step, you build the project and set its parameters (Fig. 4).

Here, you have the choice to upload a new session (choose "New project") or to load an existing project for checking, finishing or editing results (choose "Load Project").

At any step of the session procedure, you can change the running project by going back on this first step and by clicking on the "Starting another project" button.

→ In the “New project” mode
By choosing this mode, you can select the directory containing the data to be filtered (see their organization in the 3.) and set the main information regarding the running session:

- the chemical element considered as internal standard (by default, Ca)
- the calibration data corresponding to the type of standard
- the rank of each standard and sample in ICPMS analysis (the rank is an open choice. However, be careful to complete all fields with non-redundant numeric figures)
- When all completed, you can valid your choice and access to the next step of the data reduction procedure.

Note that:

The **elementR** package provides an example of session to run. To this aim, launch the `runElementR()` function and press the “Run Example” button.

For filling easily the rank, you can click on the corresponding box and write directly the rank of your replicate. You can then go to the next replicate by pressing “tab” of your keyboard.

→ In the “Load project” mode
By choosing this mode, you have to select the .RData file corresponding to the project to open and all data of this session will be automatically uploaded. You can valid your choice and access to the next step of the data reduction procedure.

Note that:

The **elementR** package provides an example of session to load. To this aim, run the `runElementR()` function and press the “Load Example” button .

When this first step is over, an “Export project” button is appeared at the bottom of the left side bar. It allows to export the final data and save the project to save user's results (see 6.). Moreover, a brief summary of the session running is then displayed.

VERY IMPORTANT: the project and associated data are exported only when you have pressed the "Export Project" button. There is no automatic saving.

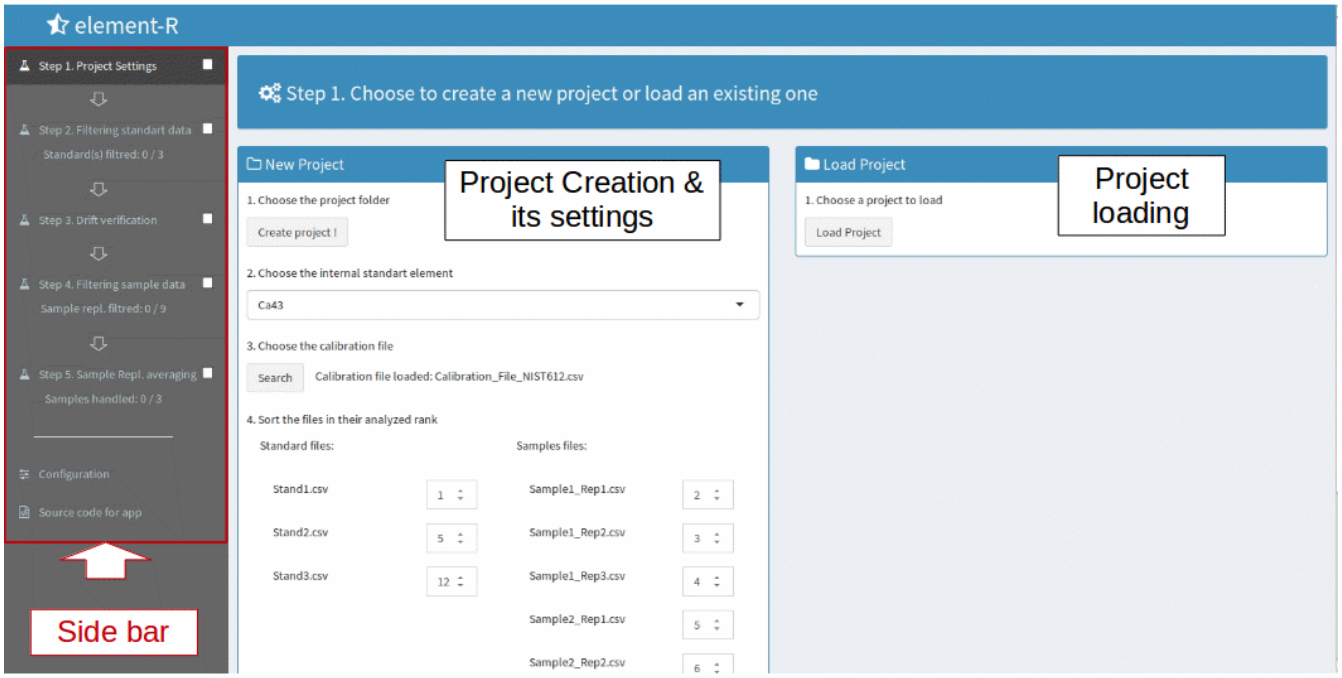


Figure 4: Standard filtration tab functionalities

- Step 2: Filtration of the standard replicates -

In this second step, you proceed to standard replicates filtration.

To this aim, you must first select the standard replicate to filter (Fig. 5).
Once the replicate chosen, raw and reduced data of the considered replicate appear in two boxes (Fig. 5). The left box aims to select background and plateau thresholds thanks to the selectors below the plot. The right box plots calculated data in order to check the validity of your selected thresholds (i.e. reduced and intermediate data, see below).

elementR allows to check all the intermediate and final data of the filtration procedure. To reach them, go to the right box and selected the wanted data (see Appendix 1 for more details about render data):

- “Blank” for the blank values
- “Raw” for raw data
- “Plateau” for plateau data
- “Blank removed” for `data_SuppBlank`.
- “>LOD” for `data_SupLOD`

- “Normalized” for `data_Norm`
- “Outlier free” for the `data_OutlierFree`

Once all thresholds approved, all information can be saved by clicking on the "Save" button. Moreover, you can delete this validation at any step of the procedure by going back on this tab and by clicking on the "Delete" button. You can repeat the procedure as many times as wanted until obtaining satisfying results.

When step 2 is running, a "graphic export" box (collapsed by default, enlarged by clicking on the white cross) allows you to export graphic displayed on this page with the export format sets in the “Configuration” tab.

VERY IMPORTANT: the project and associated data are exported only when you have pressed the "Export Project" button. There is no automatic saving.

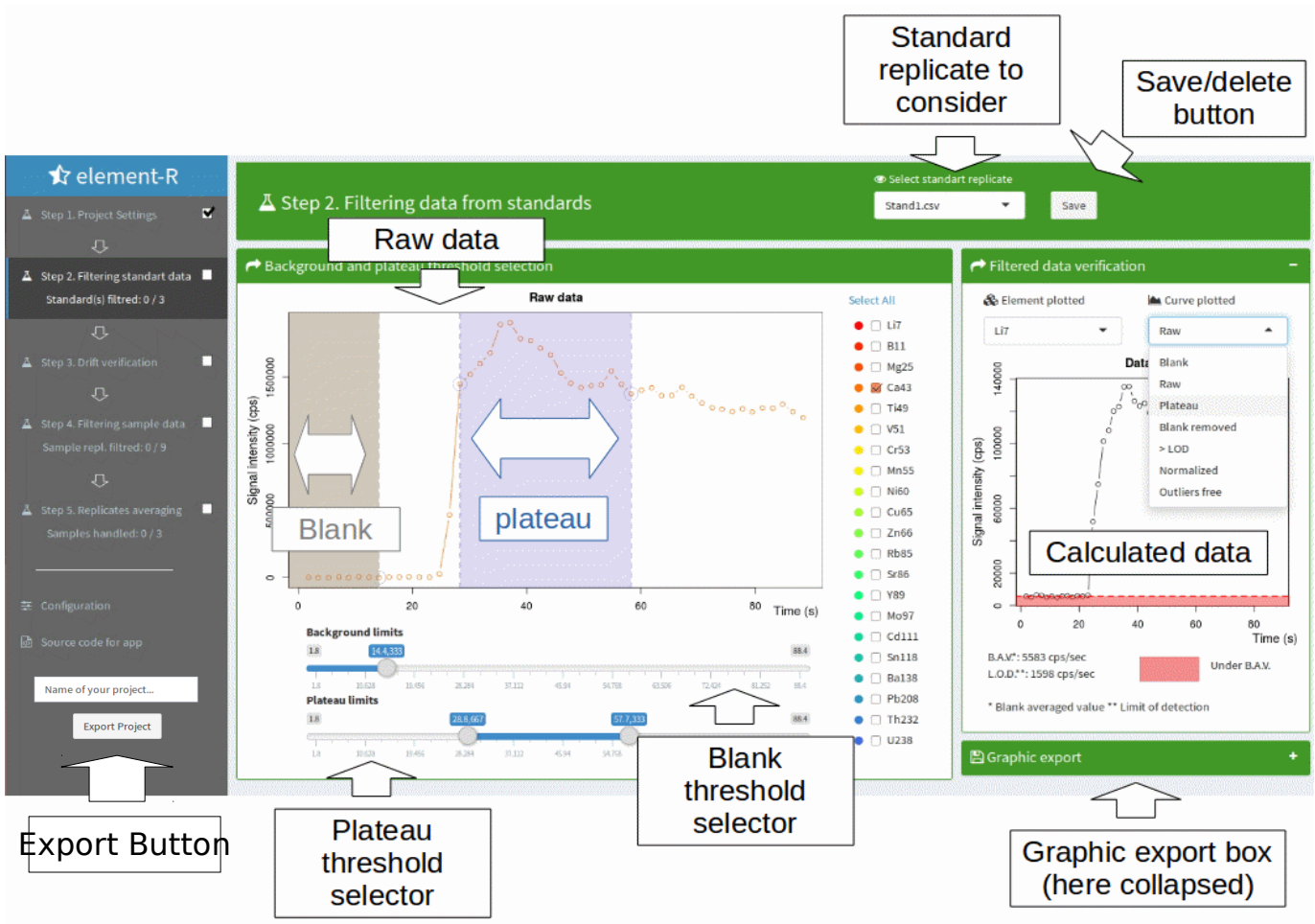


Figure 5: Standard filtration tab functionalities

- Step 3: Verification of the machine drift -

In this third step, you check the machine drift occurring during the ICPMS session.

This machine drift is evaluated thanks to linear regressions (for more details about the regression, its parameters and the correction of the machine drift performed by **elementR**, see 5.). If a machine drift occurs, the chemical element(s) affected by this drift and the whole parameters of the linear regression performed on their values are displayed. If no machine drift occurs, six elements are displayed randomly. However, you can display more elements by adding them in the "element to plot" widget (Fig. 6).

At this step, you have the choice to correct the machine drift for some elements (by clicking on the “correction” check box of these considered elements) or for all of them (by clicking on the “Correct all” check box beside the names of the elements to display). Note that the chemical elements with a non significant slope can not be corrected.

Once approved, the information regarding machine drift correction can be saved by clicking on the “Save machine drift" button. Moreover, user can delete the validation at any step of the procedure by returning to this tab and by clicking on the "Change machine drift" button. You can repeat the procedure as many times as wanted until obtaining satisfying results.

When step 3 is running, "graphic export" box (collapsed by default, enlarged by clicking on the white cross) allows user to export graphic displayed in this page with the export format set in the “Configuration" tab.

VERY IMPORTANT: the project and associated data are exported only when user have pressed the "Export Project" button. There is no automatic saving.

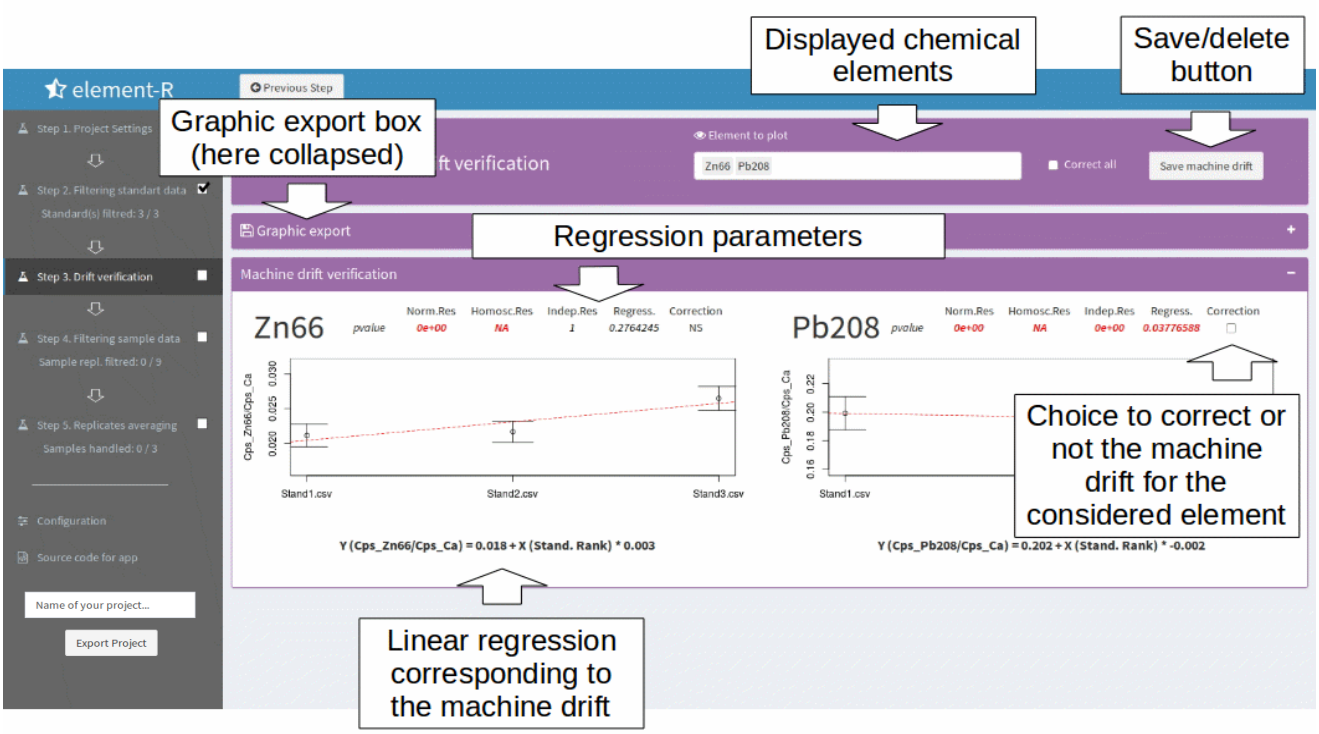


Figure 6: Drift verification tab functionalities

- Step 4: Filtration of the sample replicates -

In this forth step, you proceed to sample replicates filtration.

To this aim, you must first select the sample and the replicate to filter (Fig. 7).
Once the replicates chosen, raw and reduced data of the considered replicate appear in two boxes (Fig. 7). The left box aims to select background and plateau thresholds thanks to the selectors below the plot. The right box plots calculated data in order to check the validity of your selected thresholds (i.e. reduced and intermediate data, see below).

elementR allows to check all the intermediate and final data of the filtration procedure. To reach them, go to the right box and selected the wanted data:

- “Blank”, “Raw”, “Plateau”, “Blank removed”, “>LOD” and “Normalized” have the same results than for standard filtration (see here)
- “Outlier free” is not performed for sample filtration as all sample values mattered
- “Concentration” for `data_Norm` convert in concentration (`data_Conc`)
- “Conc. corrected” for `data_Norm` convert in concentration by taking in account the machine drift for the chemical element(s) chosen to be correct (`data_ConcCorr`)

Once all thresholds approved, all information can be saved by clicking on the "Save" button. Moreover, user can delete this validation at any step of the procedure by returning on this tab and by clicking on the "Delete" button. You can repeat the procedure as many times as wanted until obtaining satisfying results.

When the step 4 is running, a "graphic export" box (collapsed by default, enlarged by clicking on the white cross) allows user to export graphic displayed in this page with the export format set in the “Configuration” tab.

VERY IMPORTANT: the project and associated data are exported only when user have pressed the "Export Project" button. There is no automatic saving.

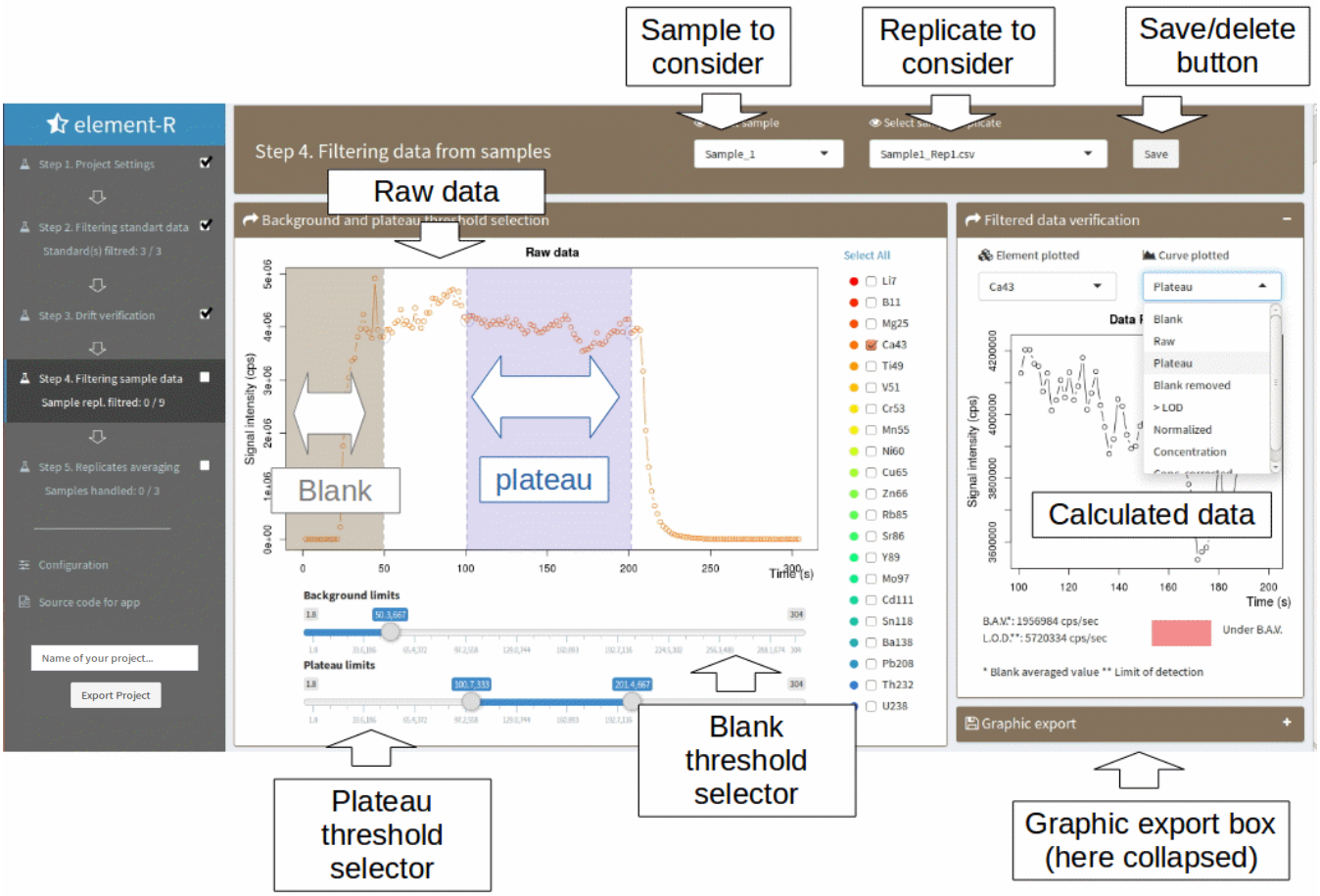


Figure 7: Sample filtration tab functionalities

- Step 5: Averaging of the sample replicates -

In this last step, you proceed to sample replicates averaging according to two modes: "spot" and "raster".

To this aim, you have first to select the sample to average and indicate its mode of averaging (Fig. 8).

→ the "Spot" mode (Fig. 8):

A reactive table displays the average and standard deviation (SD) per chemical element for each replicate of the considered sample and for all replicates (total mean and total SD). You can modify which replicates to keep for final calculation, the total average and SD being automatically recalculated.

Once approved, all information can be saved by clicking on the "Save averaging" button. Moreover, you can delete the validation by going back on this tab and by clicking on the "Delete averaging" button. You can repeat the procedure as many times as wanted until obtaining satisfying results.

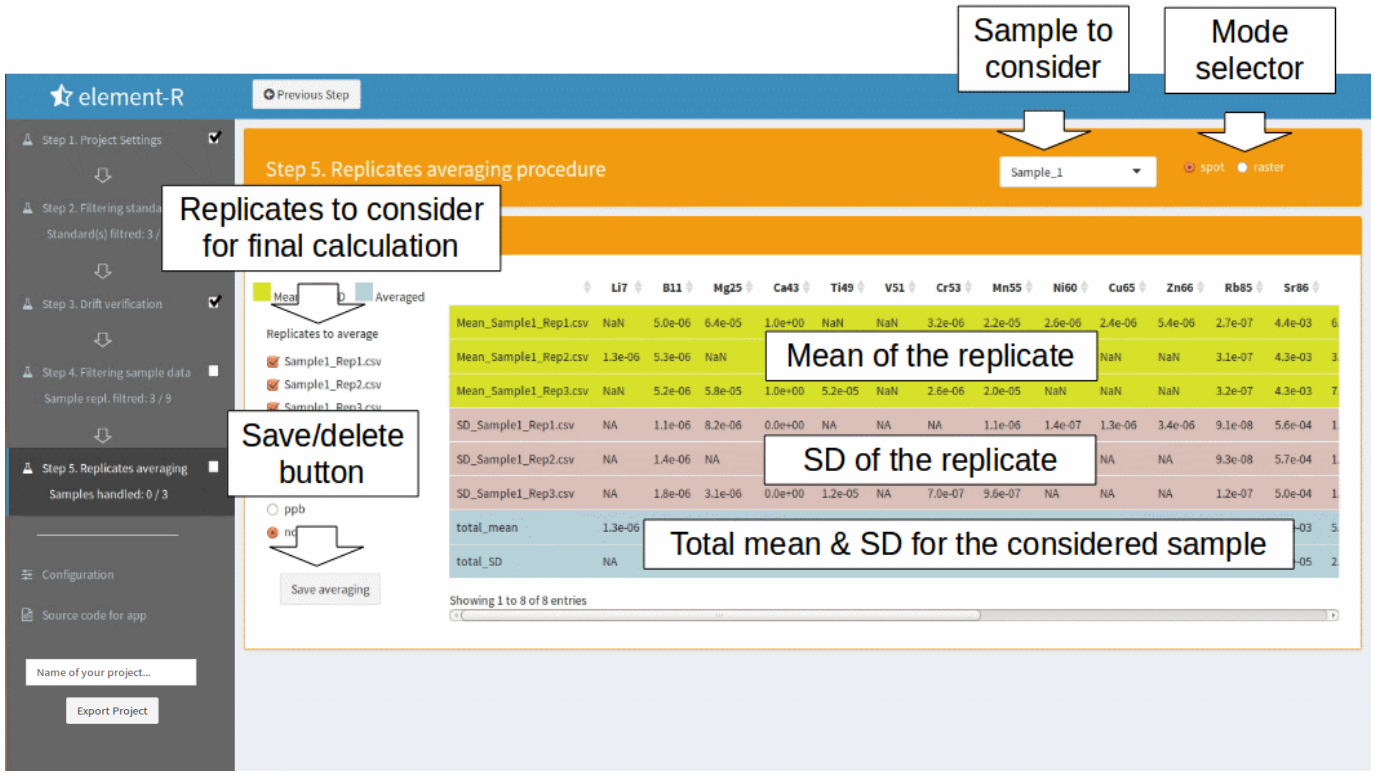


Figure 8: Sample replicate averaging tab functionalities in spot mode

→ The "Raster" mode (fig. 9 & 10):

A reactive plot displays the reduced data (i.e. in concentration, corrected or not from the machine drift according to your instructions in step 3) for each replicates of the considered sample. You can choose to select which replicates are part of the final calculation and realign visually their data by moving back and forth the replicates curves (thanks to "realign" numerical inputs). Once all curves are realigned, you can proceed to the data averaging by clicking on the "Mean" button. At this point, a black curve appears corresponding to the final averaged value of the considered sample (see fig.10).

Averaged data can be saved by clicking on the "Save averaging" button or delete if the realignment is not optimal by clicking on the "Delete averaging" or "Delete Realignment" button. You can repeat the procedure as many times as wanted until obtaining satisfying results.

When step 5 is running in raster mode, a "graphic export" box (collapsed by default, enlarged by clicking on the white cross) allows user to export graphic displayed in this page with the export format set in the “Configuration" tab.

VERY IMPORTANT: the project and associated data are exported only when user have pressed the "Export Project" button. There is no automatic saving.

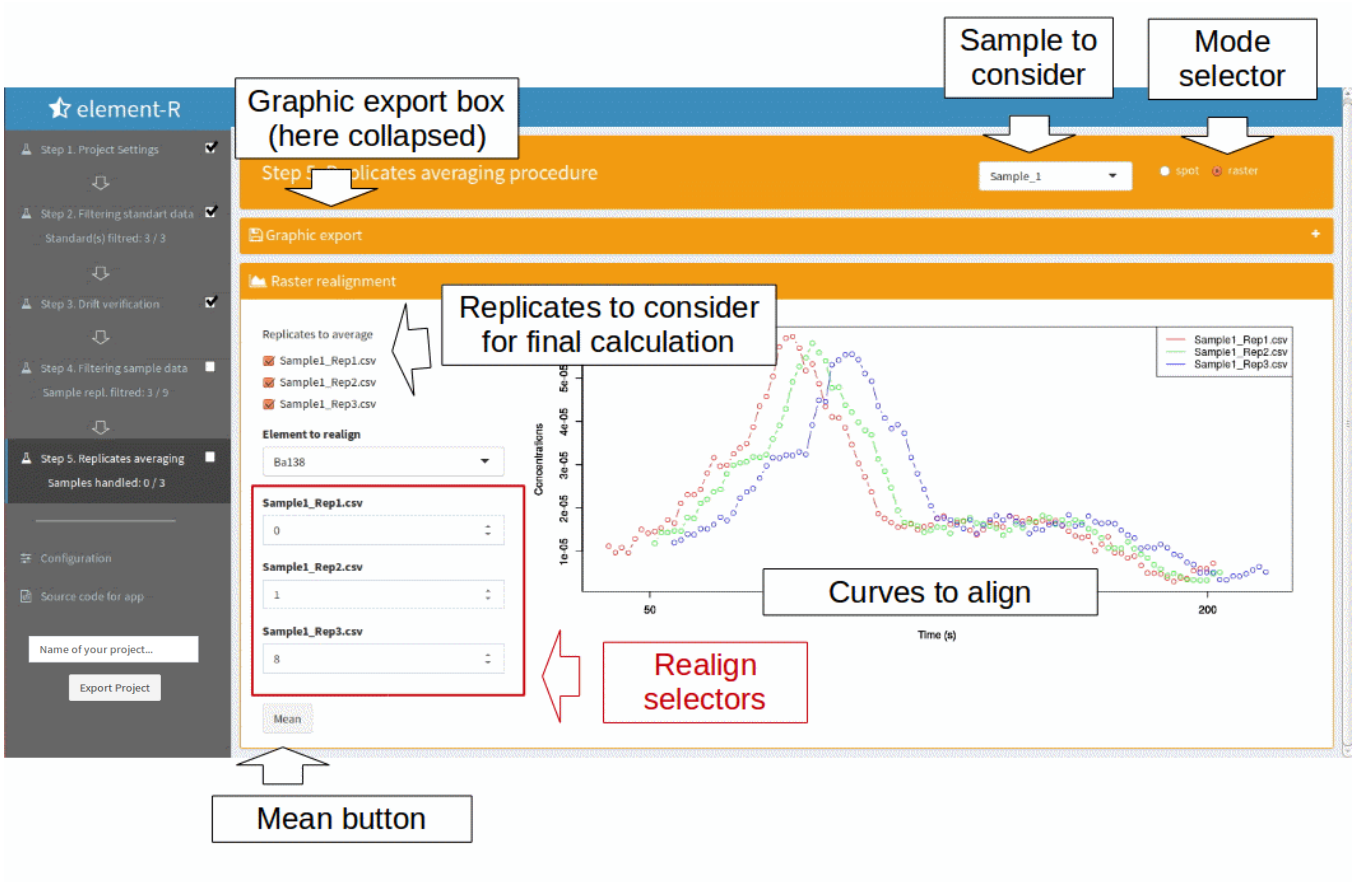


Figure 9: Sample replicate averaging tab functionalities in raster mode

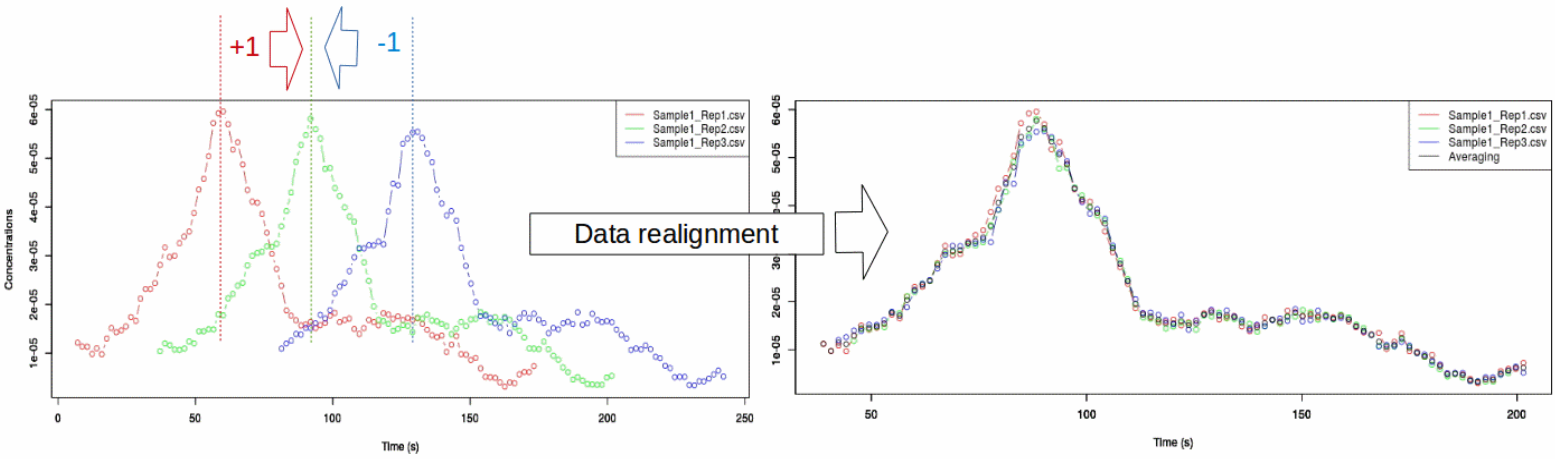


Figure 10: Sample replicate realignment and averaging

- Additional step: Optional settings -

This tab displays optional settings:

- the value to put when data are below the limit of detection (step 2 & 4)
- the color of the plotted chemical element (as some colors are not properly seen, you can change the color of particular chemical elements or all colors by clicking on the colored rectangle of the considered element). This option only available after the validation of the running project.
- Graphic and data export parameters

Note that:

This parameter settings is an optional step and its validation is optional.

You can set these parameters at any step of the procedure.

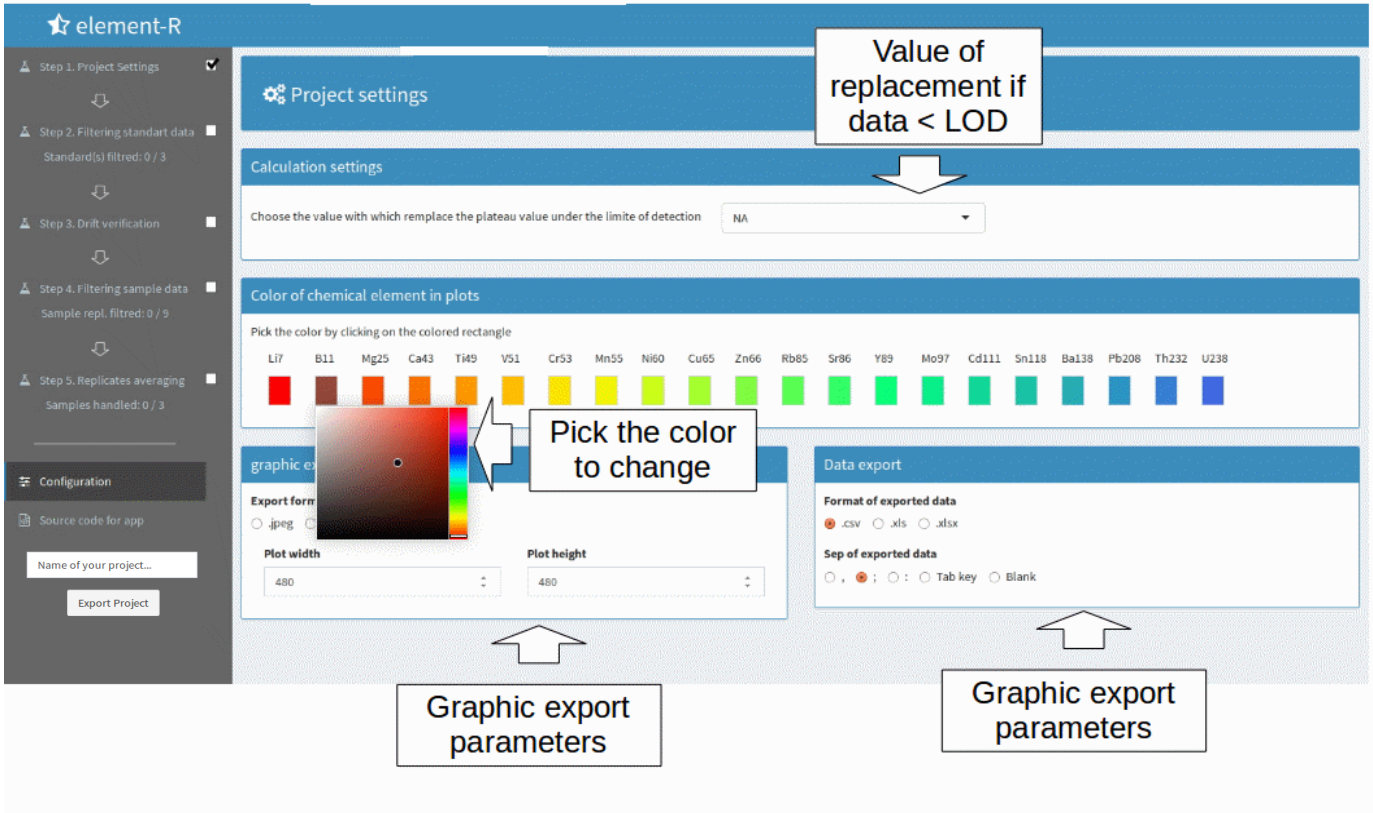


Figure 11: Configuration tab functionalities

5. What is exactly happening when elementR performs the machine drift verification and correction ?

The verification of the ICPMS drift performed by the **elementR** package is conducted by evaluating the evolution of the standard signal of each chemical element throughout the session.

→ Machine drift verification (Fig. 11)

In the case of having more than 2 standard replicates, a linear regression of the standard signal as a function of their rank in ICPMS analysis is processed (R function *lm* of the *stats* package). All parameters of the regression are statistically tested and displayed (normality, homoscedasticity and independence of the residuals with respectively the *shapiro.wilks* function of the *stats* package, *hmctest* and *dwtest* functions of the *lmtest* package (Hothorn *et al.* 2015), *p.value* of the regression) and coefficients (slope and intercept).

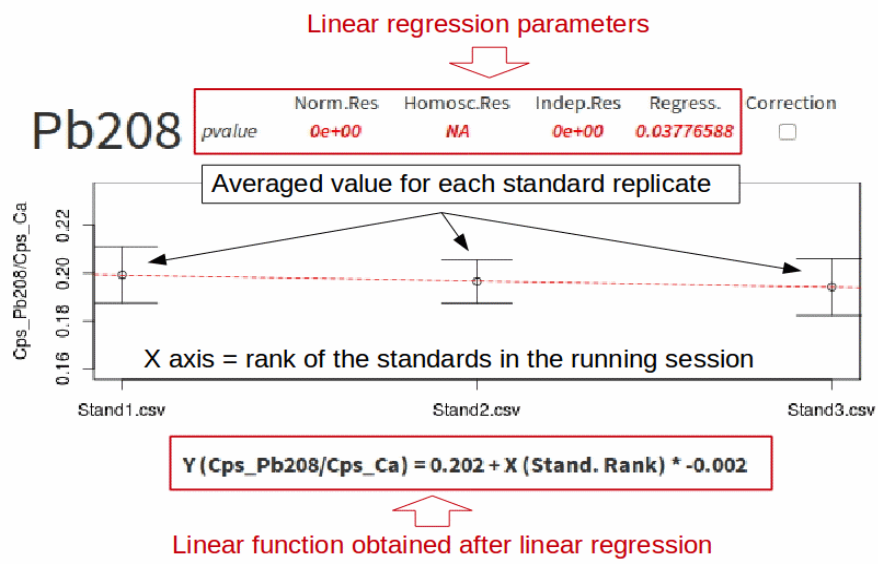


Figure 11: Machine drift verification performed by **elementR**

In the case of only two standards available, the linear regression is therefore performed without testing its significance. The intercept and slope are then displayed and you will be able to correct the drift for this(ese) chemical element(s).

In the case of only has one standard available, the value of the single standard is used for converting signal intensity in concentration. No correction is possible.

If no standard available (for some element, for instance), the data to convert in concentration are replace by NA. No correction is possible.

→ Machine drift correction (Fig. 12)

Note that in any case, you have the choice to correct or not the temporal drift of the machine if you consider this drift relevant or not.

If chemical element(s) are picked to be corrected, **elementR** proceeds to an additional step after the concentration conversion as followed:

Example of Pb208 in the “Example_Session”:

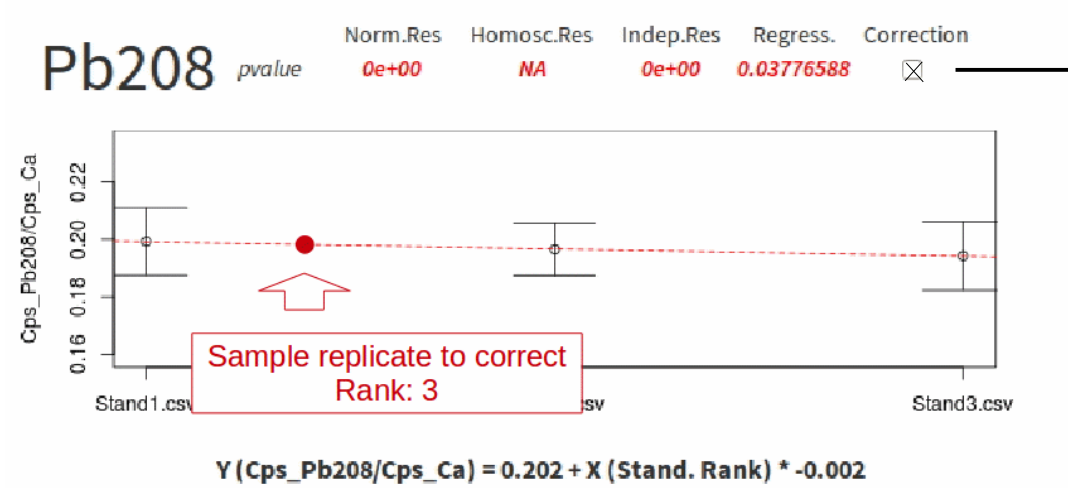


Figure 12: Machine drift correction performed by **elementR** for sample replicate at rank 3

Without drift correction vs. **With drift correction**

Conversion of the intensity signal in concentration for Pb208:

$$[Conc.] = \frac{Calib_{Pb208} \times IntenSignal_{sample.Pb208}}{IntenSignal_{standard.Pb208}}$$

Where:
→ *[Conc.]* is the final concentration
→ $Calib_{Pb208}$ is the Pb208 value of the considered calibration type (here, NIST 612)
→ $IntenSignal_{sample.Pb208}$ is the Pb208 intensity for the considered sample replicate
→ $IntenSignal_{standard.Pb208}$ is the average intensity for all standard replicates within the running session

Conversion of the intensity signal in concentration for Pb208:

1. Calculation of the theoretical value (*TheoretStand*) of the standard intensity at the rank 3 (according to the equation above):

$$TheoretStand_{Rank\,3.Pb208} = -0.02 \times 3 + 0.202$$

2. Conversion of the signal intensity in concentration by taking in account the machine drift correction

$$[Conc.] = \frac{TheoStand_{Rank\,3.Pb208} \times IntenSignal_{sample.Pb208}}{IntenSignal_{standard.Pb208}}$$

Where:
→ *[Conc.]* is the final concentration
→ $TheoretStand_{Rank3.pb208}$ is the theoretical Pb208 value at rank 3 of the calibration type (here, NIST 612)
→ $IntenSignal_{sample.Pb208}$ is the Pb208 intensity for the considered sample replicate
→ $IntenSignal_{standard.Pb208}$ is the average intensity for all standard replicates within the running session

6. How export data from elementR ?

As soon as data are loaded, an “export project” button appears at the bottom of the left

sidebar. You can name your project and press it at any step of the data reduction procedure to export:

- the entire project and its settings in an .Rdata file. This file is the one to load for re-opening an existing project (see step 1 and on Fig.5)

- all the intermediate and final data in worksheet format. Note that only validated data are exported.

All these files are exported in the sub-folder called “Done” located in the running project folder.

Note that:

The **elementR** package provides an example of exported data in sub-folder called “Done” from the **elementR** folder[4]

By default, **elementR** export data in .csv. If you want to change this parameter, go to the “Configuration” tab.

- Exports for replicates -

For each replicate filtered (here, named X), **elementR** exports (name of data in Annexe 1):

- “data_Blank_X” corresponding to `data_Blank`
- “data_plateau_X” corresponding to `data_Plateau`
- “data_SuppBlank_X” corresponding to `data_SuppBlank`
- “data_SupLOD_X” corresponding to `data_SupLOD`
- “data_Norm_X” corresponding to `data_Norm`
- “data_OutlierFree_X” corresponds to `data_OutlierFree`
- “data_Conc_X” corresponds to `data_Conc`
- “data_ConcCorr_X” corresponds to `data_ConcCorr`

- Exports for sample batch -

When the last step (step 5) has been performed on at least one sample (here, called Y), **elementR** exports:

→ “finalReplicates_X” corresponding to the `data_ConcCorr` taking in account the realignment (export only in raster mode)

→ “final_Y” corresponding to the final data of the sample (i.e. the averaged value for all chosen replicates).

- Overview tables -

Some data summarizing session settings or recapitulating the main results of steps are

4 if you do not manage to find this folder on your computer, see here

also exported in order to provide as many details as possible on the running procedure:

- “SummarySettings” summarizes the main settings for each filtered replicate (rank in ICPMS analysis, blank and plateau thresholds, blank averaged value & limit of detection for each chemical element)

- “regression_parameters” indicates the parameters of the linear regression corresponding to the machine drift (residuals normality, homoscedasticity, independence, *pvalue* of the slope, intersect and slope)

- “SummaryStandard” displays the average and the standard deviation per chemical element for each standard replicate and for all standards of the session

VERY IMPORTANT:

You can export as many times as wanted at any step of the data reduction procedure. Be careful, the overwritten data will be lost and unrecoverable.

The project and associated data are exported only when user have pressed the "Export Project" button. There is no automatic saving.

7. More questions ?

If, despite the care devoted to program this package and write this documentation, you have difficulties to install or run **elementR**, if you have questions about the procedures or calculation, or if you want to report bugs, do not hesitate to consult the official **elementR** documentation provided by the CRAN, to ask question on GitHub (click on the “Source code” tab) or to send us an email to charlott.sirot@gmail.com

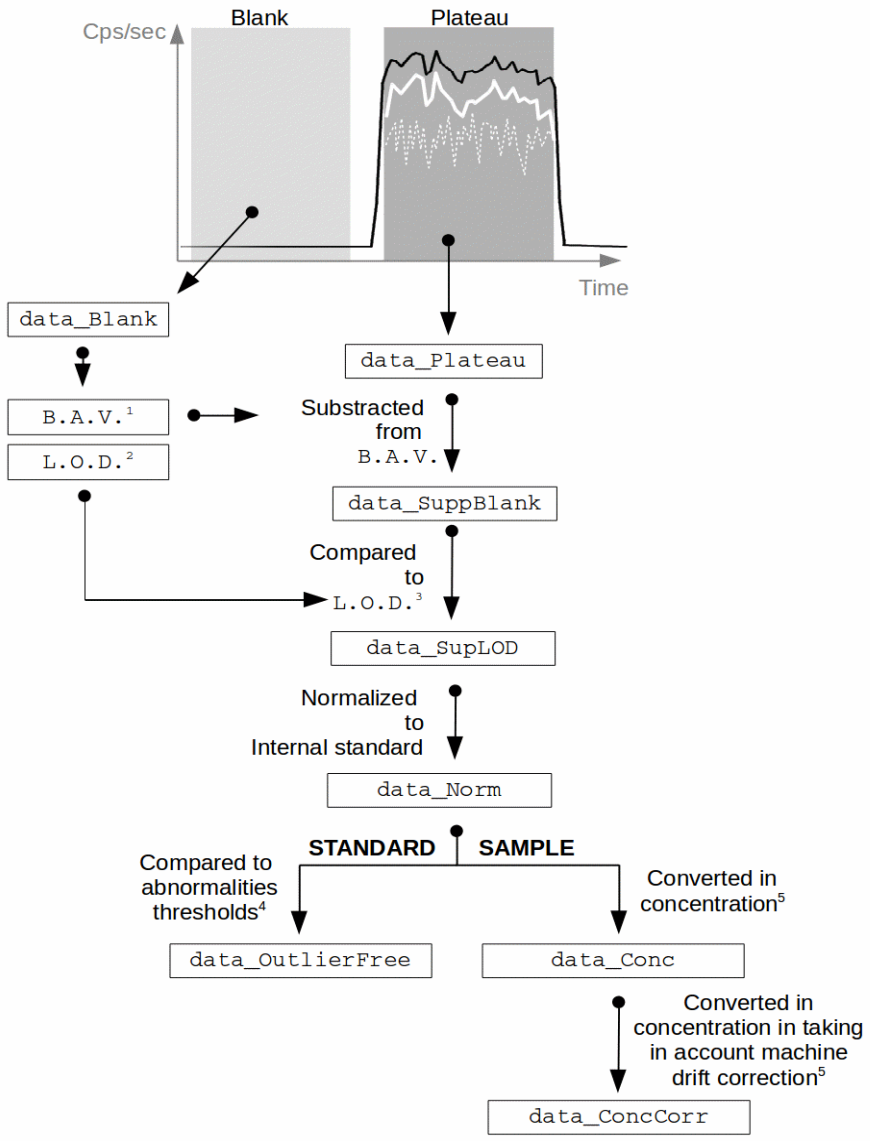
8. References

Elsdon & Gillanders. Interactive effects of temperature and salinity on otolith chemistry: challenges for determining environmental histories of fish. Can. J. Fish. Aquat. Sci. Vol. 59, 2002.

Fowler et al. Experimental assessment of the effect of temperature and salinity on elemental composition of otolith using laser ablation ICPMS. J. Fish. Aquat. Sci. Vol. 52, 1995.

Milton & Chenery. The effect of otolith storage methods on the concentrations of elements detected by laser-ablation ICPMS. J. of Fish Biology, Vol. 53, 1998.

Thorrold et al. 1998. Accurate classification of juvenile weakfish Cynoscion regalis to estuarine nursery areas based on chemical signatures in otoliths. Marine Ecology Press Series, Vol. 173, 1998.



Appendix

1. Calculations included in step 2 and step 4 of the data reduction procedure performed by **elementR**:

1. Blank averaged value (B.A.V.): averaged values for each chemical element of `data_Blank`

2. Limit of detection (L.O.D.): 3 times standard_deviation of `data_Blank`

3. Comparison of `data_SuppBlank` to L.O.D.: `data_SuppBlank` values below the L.O.D. are replacing by NA, or 0 or B.A.V. (according to user choice)

4. Comparison of `data_Norm` to abnormalities thresholds: `data_Norm` values below the LimitMin or up to LimitMax are replacing by NA where:

LimitMin = mean(`data_Norm`) – 2 x standard_deviation(`data_Norm`)
LimitMax = mean(`data_Norm`) + 2 x standard_deviation(`data_Norm`)

5. See 5.