This guide provides a step by step walk-through for the script published in Mischel et al. (2016): TERMITE, an R script for fast reduction of LA-ICPMS data and its application to trace element. The reader should carefully work through this manual prior to the usage of the script TERMITE.

General marks:

* A section for Troubleshooting is found at the end of this guide.
* Lines in the script starting with # are commentaries as well as everything in a line following #.
* Variables consisting of text (e.g. “your\_sample\_name”) must be provided with quotation marks.
* TERMITE will never change your raw data files. All files will just be read into R’s internal memory and the calculations will be performed within R.
* The files of the reference materials need to have an assigned, unique name in order for the script to work properly. The name **must** contain the text string provided in the reference material section of the script (Figure 5, name in brackets (“NIST612”, “GSD”, “MACS3”, “KL2-G”, “BAM-B”, “T1-G”, “MACS1”, “StHs”, “ATHO-G” and “NIST610”)). When using one reference material the line RefMat1 must be filled in, the other lines must start with #. If two lines without # are indicated, those two reference materials will be used for calibration (see exemplarily Figure 5). The reference materials must be in chronological order (e.g. Reference Material 1, Reference Material 2), no matter how many reference materials have been analysed during the analytical session. TERMITE calculates one RSFmean from the RSF values for each of the selected reference materials.
* Important to note that even if the files of two reference materials are copied into the ReferenceMaterial\_directory, only the files which are indicated in the reference material section of the script are used.
* Experiments consisting of more than ten spot measurements need one leading zero in the filename (e.g., Spot\_01.asc, Spot\_02.asc). If the experiment consists of more than 100 spots, the user needs to rename the data files using two leading zeroes (e.g., Spot\_001, Spot\_002). This practice is mandatory to guarantee correct sorting of the spots on the sample.
* The output of the script is written into the Results\_directory:
  + A pdf file with the rawCountrate is plotted together with vertical lines indicating the sections used for the background determination and the sample signal as defined in the HelpValue sections of the script (rawCountrate\_your\_sample\_name.pdf)
  + The results are saved as a csv and a pdf file (Results\_your\_sample\_name.csv and Results\_your\_sample\_name.pdf).
  + The Limit of detection (LoD) is saved as a csv and a pdf file (LoD\_ReferenceMaterial\_your\_sample\_name.csv and LoD\_ReferenceMaterial\_your\_sample\_name.pdf
  + The RSF values are saved as a csv and a pdf file (RSFused\_your\_sample\_name.csv and RSFused\_your\_sample\_name.pdf).
* If the user need to update the Reference Material values provided together with this script, the file is found in the directory “TERMITEScriptFolder” named “Standards\_GeoReM.csv”. This file can be opened and edited with Excel, OpenOffice or any text editor. The values must be separated by commas and white space is filled with NA for consistency.

Initial preparations:

1. Download and install R (<https://cran.r-project.org>). If problems arise during the installation, please use the documentation provided on this site.
2. Unzip the file your\_main\_directory.zip. This file contains the script TERMITE and all mandatory files for the script to run together with an example of a line scan and 5 spot measurements.
3. For spot scan analysis the file TERMITE\_spotscan.r is used. For line scan analysis the file TERMITE\_linescan.r is used. Technically these files are identical, but the example ought to cover both types of experiments.
4. Navigate into your main directory.
5. Prepare, if not existing, additional folders in the main directory as shown in Figure 1. The user can decide which names should be assigned to the directories.

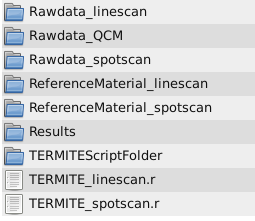
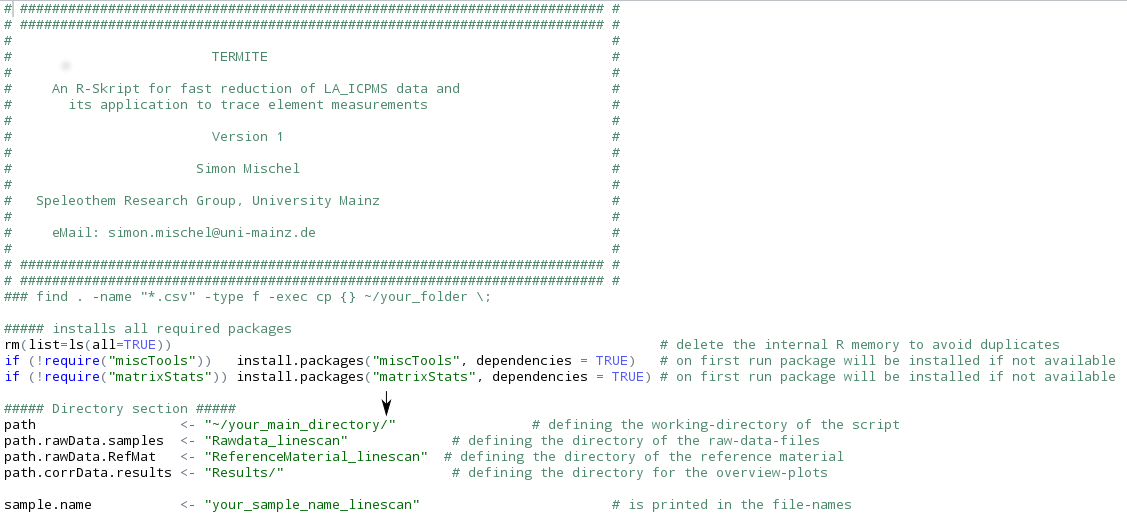
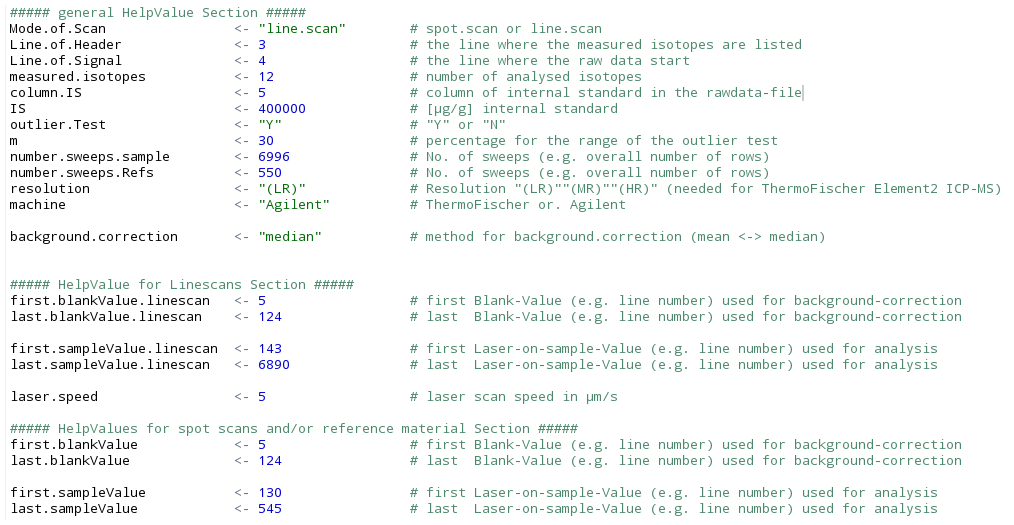


Figure 1: The structure of the directories which must be created by the user if not existing.

1. Copy the raw data sample files of the laser ablation measurement into the corresponding directory (Rawdata\_linescan or Rawdata\_spotscan). If a line scan is evaluated, the Rawdata\_linescan directory should only contain 1 file. Important to note that one dataset must have the same internal standard concentration (e.g. for a stalagmite, Calcium: ~400000 µg/g). If, for example a quality control material (QCM) like MACS-3 is treated as an unknown sample, the internal standard concentration is different (MACS-3, Calcium: 376900 µg/g) and therefore the raw data files of MACS-3 have to be stored in separate folders (e.g., Rawdata\_QCM).
2. Experiments consisting of more than ten samples need one leading zero in the filename (e.g., Sample\_01.asc, Sample\_02.asc). If the experiment consists of more than 100 samples, the user needs to rename the data files using two leading zeroes (e.g., Sample\_001, Sample\_002). This is required to guarantee correct sorting of the samples.
3. Copy the Reference Material data files used for calibration into the folder ReferenceMaterial\_directory.
4. Open the file “TERMITE\_script\_linescan.r” or “TERMITE\_script\_spotscan.r” depending on your experiment and work through the script prior to data evaluation. The steps are explained in the following (Figure 3, 5 and 6).
5. The packages “miscTools” and “matrixStats” will be installed automatically if the corresponding lines are run in R. (the 3 lines after the part “installs all required packages”, Figure 2). If problems arise try to choose a different download mirror or try using “http” instead of “https”. In case of Errors these packages can be installed manually by executing the commands install.packages(“matrixStats”) and install.packages(“miscTools”) inside R.
6. As a first step the names of the folders must be passed into the script in the directory part if not already there (Figure 2). Important to note is that the names must be passed exactly to the script according to the names assigned in step 5. Please note that R is case sensitive and distinguishes between lower and upper case letters. Users must be cautious to provide the directories using slash (/) instead of backslash (\).

Figure 2: The directory section of the script. The slash (/) at the end of path is mandatory (black arrow).

1. The slash (/) in Figure 2 at the end in the object path is mandatory.
2. The name of the sample (sample.name, e.g. “your\_sample\_name”) should be assigned. This helps the user to identify the sample after data reduction as this text string is printed in all file names.
3. Prior to data reduction the user should work through the section “HelpValues” and provide all important parameters for data reduction in the script (Figure 3).

Figure 3: The section of the script where all important parameters regarding the structure of the raw data is filled in. Please change only the blue numbers or the green text strings with quotation marks.

1. In the following the HelpValue section is explained step by step

* Mode.of.Scan: insert “spot.scan” or “line.scan” depending on your experimental setup.
* Line.of.Header is the line in the original raw data file where the names of the isotopes are found in the data files (Figure 4).
* Line.of.Signal is the line where the signal of the ablation run start, normally starting with the background signal (Figure 4)
* Measured.isotopes is the number of all measured isotopes including the Internal Standard (IS).
* Column.IS is the number of the column in the raw data file where the internal standard isotope (IS) is located (Figure 4).
* IS is the concentration of the internal standard in the sample (e.g. Ca → 400000 µg/g)
* The outlier test (for spot scan measurements and for reference materials) can be completely switched on and off. Please note that in case of a linescan the outlier test will only be evaluated during the calculation of the reference materials.
* m is the percentage of the range for the outlier test.
* number.sweeps.sample is the number of total sweeps recorded in the raw data file of the sample (Figure 4).
* number.sweeps.Refs is the number of total sweeps recorded in the raw data file of the reference material. Typically spot scans are measured with the same experimental setup (e.g. background and ablation time etc.) as the reference materials. In this case both values (number.sweeps.sample and number.sweeps.Refs) contain the same value.
* The string machine let the user decide which ICPMS instrument was used (provide “ThermoFischer” or “Agilent”)
* Using the ThermoFisher Element2 mass spectrometer the mode of resolution during ablation can be set to low (“(LR)”), mid (“MR”) and high (“HR”). This text string must be provided to enable TERMITE to delete this text string out of the header to obtain the names of the isotopes measured during ablation.
* The user can decide which background.correction for the background will be applied (“mean” or “median”)
* The section “HelpValues for Linescans” needs to be completed if line scans are evaluated. In the case of spot measurements the values in this section will not be needed during the calculation and can be left unchanged.
* The values in the section HelpValues for spot scans and the reference material must be always provided according to the experiment.

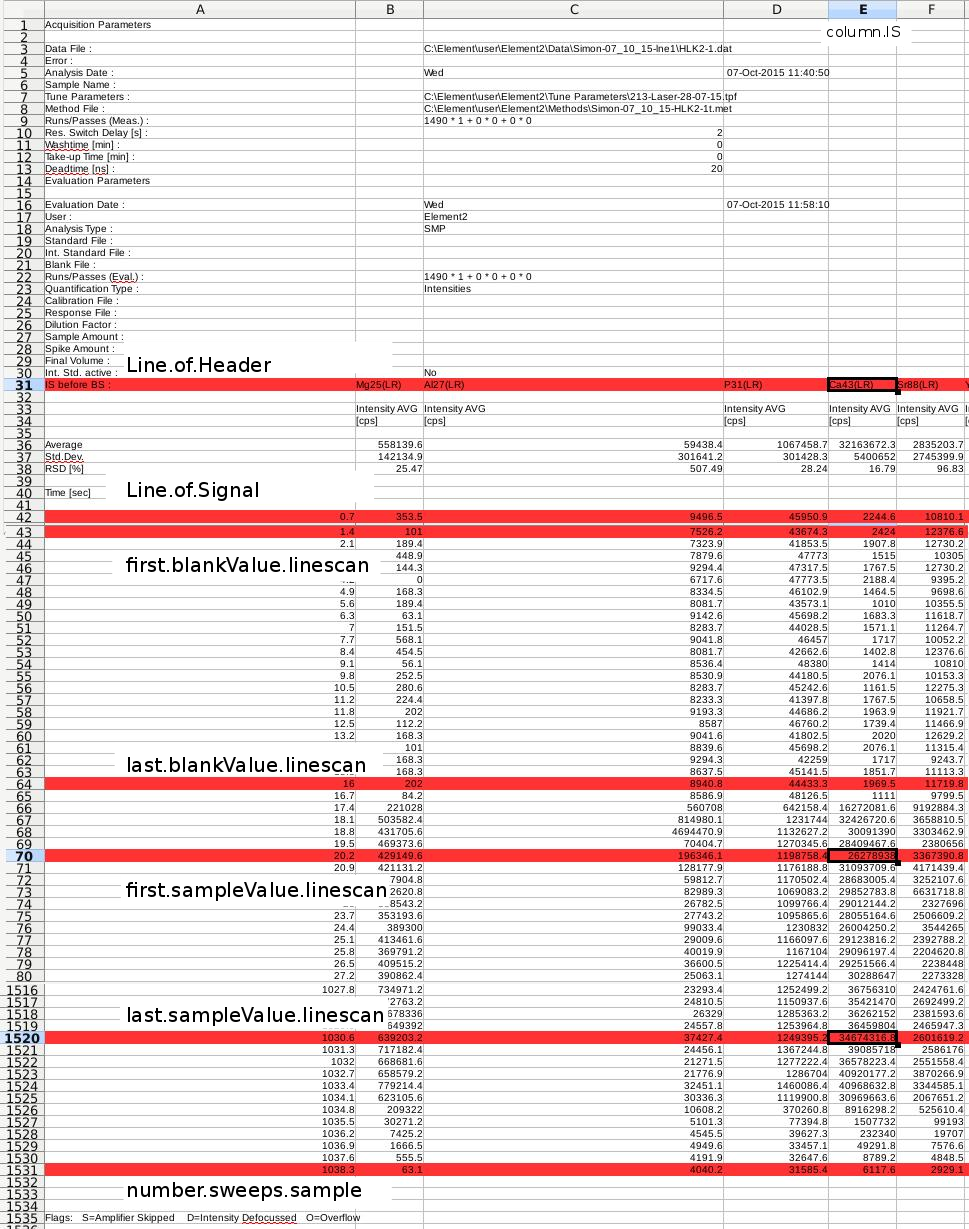
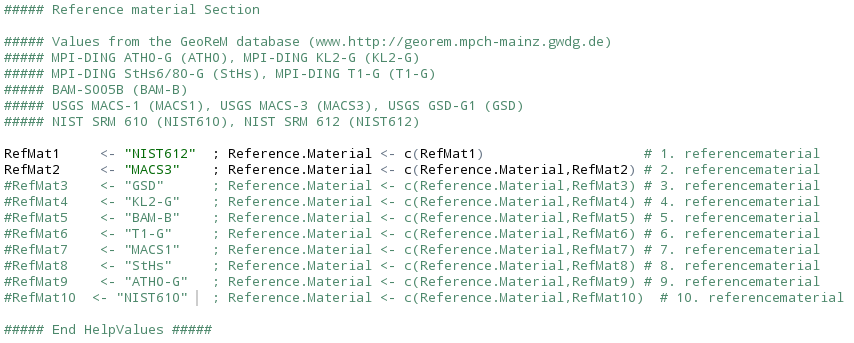


Figure 4: Example of a raw data file obtained using the Thermo Scientific Element2 ICPMS opened in Microsoft Excel or OpenOffice. Indicated are the lines of the header where the names of the measured elements are found (“Line.of.Header”) and the beginning of the recorded signal during one laser ablation session (“Line.of.Signal”), which needs to be specified in the script. Also highlighted is the column containing the data of the isotope used as the internal standard (“column.IS”, 43Ca in this example). The lines of the first and last blank value used for the calculations are highlighted (first.blankValue.linescan & last.blankValue.linescan) as well as the first and the last value used for the sample (first.sampleValue.linescan & last.sampleValue.linescan). The total number of lines recorded (number.sweeps.sample) is also indicated.

* The first and last value (e.g. the line number) used for the blank calculation needs to be filled in as well as the first and last value (e.g. the line number) which will be used for the calculation of the sample (Figure 4). Typically, the first and the last one or two values are truncated to ensure consistency of the data. This is done in the section for the line scans as well as in the spot scan section.
* The first and the last values (e.g. the line number) used during the calculation of the sample concentration need to be provided in the section for the line scans as well as in the spot scan section.
* The scan speed of the laser needs to be provided in the line scan section as this value allows the calculation of the distance of the measurements using the time section of the raw data file.
* In the “Reference Material section” the user can decide which and how many reference materials are used for calibration. The script will only work properly if at least **two** reference material files from one material are present in the directory. One ablation of the reference materials should be performed before a set of spots (e.g., 30-45 spots) and one after the set. This enables to correct for the drift of the machine during the analytical session. For each reference material a Relative Sensitivity Factor (see Manuscript for details) is calculated. Finally, TERMITE applies the mean of the single RSF values to correct the measured element concentrations of the unknown sample.
* Figure 5 shows an example where two reference materials are used for calibration (e.g. NIST SRM 612 and USGS MACS-3). If only one reference material (for example only NIST612) is used, RefMat line number 2 to 10 should start with a #. All reference materials should be in directly successive lines, i.e., if 3 reference materials are evaluated (e.g. NIST612, MACS3 and NIST610), the # before RefMat3 must be deleted and “KL2-G” must be replaced with “NIST610”.
* If no reference material file containing the text string (e.g. “NIST610”) is placed in the “ReferenceMaterial\_directory”, an error will be produced (“Error in file… No such file or directory”). If the correct files are in the directory but the wrong line is used, an error is produced (e.g. “Error in eval: object “RefMat3” not found”). If these errors occur, the user should check the content of the directory as well as the Reference material section.

Figure 5: The reference material section with the reference materials provided with this script.

1. As a last step the user should check the complete script again if all values are inserted correctly and then run the script either by executing the whole sequence or every single line.

Troubleshooting:

General considerations: Warnings will not stop the script from execution and the script should finish. Any error that occurs during data reduction will terminate the script and troubleshooting should start. The user is advised to try to understand the error messages provided by R either by reading the provided troubleshooting section or by using a search engine on the internet.

* If the raw data contain any E-values which are errors introduced during the data saving process, these values will be set to NA prior to data reduction. In R, this will produce warnings, which can be displayed using the command warnings(). The message will be (in short): Warning message: In FUN(x[[i]], …): NAs introduced by coercion.
* If you got an “Error in file: cannot open the connection” together with an addition: Warning message: cannot open file, the user should first check if the spelling in the directory section is correct (Figure 2).
* If an “Error in file: cannot open the connection with addition: Warning message: In file: cannot open file ‘NA’: No such file or directory” occurs, the most likely case is a wrong spelling in the Directory section (Figure 2). Please check also the spelling of the directory on the hard-drive of your PC. Another reason for this error could be the spelling of the reference material files.
* An “error in pdf cannot open file” is most likely the result of a misspelled directory name. Therefore, the pdf cannot be written.
* If other errors occur, please check again all values in the sections for correctness.
* If in rare cases R will not stop the calculation for a long time, please consider the amount of data which is evaluated. 24000 spot measurement files will result in the script to work for about 3 h. This will produce a raw countrate file with some GB in size.