**Pygilent**

**Fundamentals**

**Overview**

The Pygilent programs are designed to standardise and automate the data processing of trace element (TE) ratio measurements made by the Agilent 8900. Pygilent\_Ca should be used to determine the concentration of one element (usually Ca) within a ‘batch’ (also known as a ‘sequence’ or a ‘run’). This is typically performed as a first step after the samples have been prepared so that the samples can be matrix-matched (diluted to the same Ca concentration) for the main TE ratio run. Pygilent\_TE should be used to determine the TE ratios of a batch. We will refer to each of these batches as the Ca checks and the TE batch, respectively.

**Batch setup**

A batch should be setup such that there are blocks of samples with an inner layer of bracketing by bracketing standards, then an outer layer of bracketing by blanks (see diagram below). The optimal number of samples that should be analysed between bracketing standards will vary between instruments and methods. We typically run no more than seven samples between bracketing standards for the TE batch, but this is less strict for the Ca checks.

|  |
| --- |
| BLK |
| Bracking stnd |
| Sample |
| Sample |
| Sample |
| Sample |
| Sample |
| Sample |
| Sample |
| Bracketing stnd |
| BLK |

Samples and standards should be matrix-matched for the TE batch. Blanks should be made using the same stock of diluting acid (typically 0.5 M HNO3) as the samples and standards.

If you wish to utilise a calibration curve approach (discussed later), then include a block of at least three different calibration standards at the beginning and end of the batch, with further blocks inserted into the middle of longer batches.

Batch data from the Agilent must be exported as .csv files, with the option to export replicate data turned on. This can be found in the data handling section of the Mass Hunter software settings.

**Ca checks with Pygilent\_Ca**

**Overview**

Use the Pygilent\_Ca.exe to begin processing a batch folder to find the concentration of one element within the samples. This can be applied to any element, not just calcium. After importing the data Pygilent\_Ca performs the following sequence of processes automatically:

1. Blank correction on all samples and standards. Changes in the blank concentration throughout the run are corrected for by linear interpolation between pairs of blanks.
2. Standard-sample bracketing: ratioing every sample by a linearly interpolated value of the blank-corrected bracketing standard counts.
3. Calculate the pipetting volumes based on user data input.
4. Export the data output to an excel file.

**Step-by-step guide**

1. Select the batch folder you want processed.
2. Select all the blanks within the batch you want included in the blank correction. Press Submit.
3. This next window shows the blank counts throughout the batch. Use the drop-down menu to view select which isotope to view. Click on data points to select them for removal (used to remove outliers). Click submit to continue.
4. Select all the bracketing standards within the batch. Click submit.
5. The same as step (3) but with the bracketing standards. Select those you want omitted from the processing. Submit to continue.
6. Choose the isotope(s) you want the concentration data for. You can select multiple isotopes, which may be useful for comparisons (say Ca43, Ca46 and Ca48). However, only one concentration value is submitted, so all isotopes must be the same element and/or all selected isotopes must have the same concentration within the bracketing standard.
7. Enter the following, then click submit:
   1. The concentration of the selected isotope(s) within the bracketing standard.
   2. The volume of sample added to each sample dilution for this batch.
   3. The volume of acid used to dilute each sample for this batch.
   4. The target volume for each sample for the TE batch.
   5. The target concentration for each sample of the TE batch.
8. (Optional) Enter a name for the output file.

**Output information**

Once completed, the data will be exported to the Pygilent\_out folder (created if it non-existant) with the suffix \_Ca\_*timestamp*.xlsx. The file will contain seven sheets plus one sheet for each isotope selected in step 6 above. Below is a description of all the sheets, but users will only need to refer to the sheet with the relevant isotope name for a summary of the data.

* Reps, CPS and PA contain the raw, unprocessed data (counts for all replicates, mean counts, and detector mode, respectively).
* CaConc (mM) is the concentrations of all isotopes in all samples & standards within the batch assuming a concentration within the bracketing standard given in step 7 above.
* Undiluted CaConc (mM) is the same concentration data corrected to account for the dilution data entered in step 7 above.
* Vol smpl (ul) the volume of sample required to matrix-match.
* Vol acid (ul) the volume of diluting acid required to matrix-match.
* *Isotope name* a summary of the above for the selected isotope with the blanks removed and data sorted to separate the bracketing standards from the samples for readability and printing.

**A note on PA**

The Agilent 8900 has two detector modes, pulse (P) and analog (A). In the acquisition method of the batch setup, it is possible to select a detector mode option for each isotope (Analog or Auto). In Auto, the detector mode switches between the modes when required. Below approximately 106 CPS the response of A detectors is non-linear therefore P is used. Above 106 CPS the counts overwhelm P detectors, so A are used. Each replicate is assigned a single detector mode, therefore, when analysing multiple replicates, it is possible for both detector modes to be used for one sample. When this occurred the PA label is given as M (for mixed). When setting the detector mode option to Analog, the detector never switched to P mode even at low counts. This is useful for when it is known that an element will be analysed exclusively in A mode and therefore a more accurate blank correction can be performed.

It is very important that samples and standards are analysed using the same detector mode for any given isotope (and preferably never M). It is also important that samples are not analysed in A mode below 106 CPS. However, because the concentration of the isotope is unknown in samples, we need to prepare for samples to have a large range of counts. To account for this issue, include multiple isotopes of the same element with differing natural abundances. For the calcium setup, we include Ca43, Ca46 and Ca48. Ca48 is forced to remain in Analog mode, whilst the others are set to Auto. Ca48 is used by default, but the expected PA column is compared to the PA column in the output. If any of the samples have an expected PA of ‘P’ then either Ca43 or Ca46 should be used, whichever yields the same PA between samples and bracketing standards.

As a further fall-back in case all isotopes are unsuitable (i.e. the unlikely event where all isotopes have different PA between standards and samples), we include several measurements a standard that is 10% the concentration of the bracketing standard within the batch. Then, the batch can be re-processed with this weaker standard as the bracketing standard.

**TE batch with Pygilent\_TE**

**Overview**

The Pygilent\_TE.exe is designed to process Agilent TE data to produce calibrated TE ratios as well as helpful statistics and figures that can be used to assess the data quality. As with above, the default ratio isotope is Ca48, but in theory any element/isotope can be used if the correct standard values are applied (more on this later). In brief, the Pygilent\_TE performs the following:

1. Associates each isotope to an element in the stndvals.csv.
2. If using a calibration curve approach, after the user selects the calibration standards, the detector mode of each standard for each isotope is checked to see if it matches the bracketing standard, if not, then that isotope within that calibration standard is omitted from the curve.
3. Each sample and standard is blank corrected by linear interpolation then ratioed with the ratio isotope (typically Ca48) according to:

Where and are the mean CPS of the analyte and ratio element, respectively, is the time the sample was analysed relative to the two bracketing blanks (bI and bII).

1. Each value of R is then corrected for signal drift between bracketing standards (sI and sII) using the following equation:

Where is the time the sample was analysed relative to the two bracketing standards (sI and sII).

1. In single-point calibration mode every B value is then multiplied by the true standard value for that analyte isotope within the bracketing standard. With the calibration curve approach, each of the B values for the calibrants are regressed against their respective true values. All samples are then corrected using the following equation.

Where is the slope of the curve and is the intercept.

1. The standard error in B is calculated via a Gaussian approach:

Where and are the variances in the CPS between replicates of the analyte and ratio isotope, respectively, and is the covariance of the CPS between replicates of the analyte and ratio isotope.

For the calibration curve approach, further errors deriving from the error in the fit are added in quadrature to the above.

1. The limit of detection (LoD) is calculated by determining the blank equivalent concentration in a similar manner to the Ca checks (blank correcting the bracketing standards then using single-point calibration). Outlier blanks are automatically removed (> q75+1.5\*IQR or < q25-1.5\*IQR). Then the LoD is calculated as the mean blank concentration plus three standard deviations of the blanks throughout the run.
2. Long-term precision data is acquired from the data archive (AgilentArchive.csv). Standards from the stndvals.csv can be selected to find their long-term precision, performed by searching their sample names in the archive. Outliers are automatically removed.
3. Data are exported in two excel files, a simple and a full version.
4. Data is reformatted and written to the AgilentArchive.csv.
5. (Optional) figures are produced.

**Step-by-step guide**

1. Select the batch folder you want processed. Wait for data import to complete.
2. Select all the blanks within the batch you want included in the blank correction. Press Submit.
3. This window shows the blank counts throughout the batch. Use the drop-down menu to view select which isotope to view. Click on data points to select them for removal (used to remove outliers). Click submit to continue.
4. Select all the bracketing standards within the batch. Click submit.
5. The same as step (3) but with the bracketing standards. Select those you want omitted from the processing. Submit to continue.
6. Choose the ratio isotope for each gas mode.
7. Choose the calibration approach (calibration curve is recommended if enough calibration standards have been included in the batch. Note, that if calibration curve is selected then both the single-point and calibration curve data will be available in the final full output.
8. (Calibration curve only) Select the names of the calibrants.
9. (Calibration curve only) Make sure all calibrants are selected in the batch order. Note, by default the program will try to find adjacent calibrants but may miss some out if they are not adjacent.
10. (Calibration curve only) Assign each standard to a named standard from stndvals.csv.
11. (Single-point only) assign the bracketing standard to a named standard from stndvals.csv.
12. (Calibration curve only) This window plots the calibration curve for each isotope, which can be changed using the drop-down menu at the top. Select data points to omit them from the curve (outliers). Note, unlike the other data removal figures, here the removal is isotope-specific.
13. Select named standards from stndvals.csv to include as long-term precision data in the output file. For our analysis we typically use 8301f / 8301c.
14. (Optional) Enter a name for the output files.
15. Confirm whether you would like figures of long-term trends to be produced.
16. (If ‘Yes’ to figures) Choose either Blanks CPS or a named standard from stndvals.csv to plot.
    1. Blanks CPS gives the raw blank count variability from AgilentArchive.csv with outliers removed and highlights the data from this batch.
    2. The other named standards provide the long-term records of the respective standards from AgilentArchive.csv along with the expected values from stndvals.csv. Data are coloured to highlight different bracketing standards used.

Select each isotope using the drop-down menu. Click ‘Save’ to save the current figure as a png. ‘Save all’ exports figures from all isotopes to file as png. Exit moves onto the next standard.

1. (If the AgilentArchive.csv already contains a batch with the same name) Choose whether to overwrite the original data with the new data or not. If not, then the data will not be added to the archive.
2. Complete!

**Output information**

Once completed, the data will be exported as two excel files to the Pygilent\_out folder (created if it non-existant) with the suffixes “\_full\_*timestamp*.xlsx” and “\_short\_*timestamp*.xlsx”. Any figures saved will be exported to the figures subfolder within Pygilent\_out. The file “\_short\_*timestamp*.xlsx” is simply an abbreviated copy of “\_full\_*timestamp*.xlsx”, with just the raw data and final calibrated ratios included. The “\_full\_*timestamp*.xlsx” file contains the following sheets:

* *Reps*, *CPS*, *CPS 1sd*, and *PA* contain the raw, unprocessed data (counts for all replicates, mean counts, 1stdev of the replicates, and detector mode, respectively).
* *Covar* gives the covariances of the analyte replicate counts with the ratio isotope replicate counts.
* *Bracket* and *Bracket 1se* are the drift-corrected data and their errors (see B and in the equations above).
* *S-P cali* and *S-P cali 1se* are the final TE ratios determined using the single-point calibration approach, along with their errors.
* (If calibration curve selected*) Curve cali* and *Curve cali 1se* are the final TE ratios determined using the calibration curve approach, along with their errors.
* (If calibration curve selected) *Curve params* provides all the data used to fit the calibration curve (y data, x data) and the fit output (R2, intercept (beta0), slope (beta1) and their errors).
* *LoD* gives the mean limit of detection (see step 7 of Overview section above). This also includes the LoD as a percentage of the bracketing standard. Note these LoD data are given as TE ratios, so that they are more directly comparable to the samples. This means that when running all samples at 0.5 mM Ca then the LoD will appear twice as large as when compared to a batch run at 1 mM Ca.
* *Long-term precision* includes the archive mean and 2sd of all standards selected in step 13 above that have been analysed to the same bracketing standard as the samples within this batch. If any of those standards were present (based on an automatic text search of the sample name) in the batch, then their means, average internal error (2se) and intermediate error (2sd) will be included. All of the above are included in single-point and calibration curve calibration approaches.
* *theo R rse*, *theo Rbc rse*, *theo B rse, and theo Bbc rse* are the theoretical errors for the ratios, blank-corrected ratios, bracketed samples, and black-corrected bracketed samples. These are largely for more in-depth analysis. More work will be added to these tabs in the future.

**Customisation**

**Adding new standards**

To add new standards (or edit those given) simply edit the stndvals.csv file, making sure to save it with the same name and in the same location (it would be wise to make a copy of the original first). Each entry into the stndvals.csv should be a TE ratio. You can change the ratio isotope by adding new standards, but be sure not to mix standards with different ratio elements within the same batch and do not add duplicates of standard names or elements.

**Rebuilding Pygilent\_Ca.exe and Pygilent\_TE.exe**

The executables of the scripts were made using PyInstaller. To remake the executables simply run PyInstaller with each of the original Python scripts. Then edit the .spec file for the Pygilent\_TE so that the datas=[('stndvals.csv', '.'), (‘AgilentArchive.csv', '.')] and hiddenimports= ['openpyxl.cell.\_writer'], and edit the .spec file for the Pygilent\_Ca so that hiddenimports= ['openpyxl.cell.\_writer']. Then re-run PyInstaller with each of the spec files and the executables should be created.