TI 301F-Level I Validation of Monthly XRF data

Table of Contents

List of Figures

1. PURPOSE AND APPLICABILITY

The subject of this standard operating procedure (SOP) is to describe the Level I validation procedures for monthly XRF data set of the IMPROVE network. This Level I validation includes procedures for creating data sets and reporting final validated XRF results. This SOP aims to ensure good and consistent output for multiple users.

In order to apply this SOP, all the IMPROVE samples for any given month must be analyzed, and the data to be validated must be migrated to the database. In addition, the quality checks of the instruments' performance during analysis of the given month of samples (described in detail in TI-301E) must yield positive results (instruments stable, no contamination detected, etc.)

2. DEFINITIONS

Level I Validation of XRF data: Contains the validation of only XRF results of any given month performing a set of procedures given in this document

XRF Data Management Pages: *(http://webapp.improve.crocker.ucdavis.edu/Xrf/Home)* : The XRF Data Management Pages are webpages related to the administration and processing of XRF data.

cl-SQL Reporting Pages: *(http://cl-sql/Reports/)*:

The cl-SQL Reporting Pages are user-interface webpages that are used to query and view the datasets. Results could then be exported in a workable office-friendly format e.g. .xls, .csv, .doc

Set ID: The monthly XRF data created using XRF Data Management Pages, e.g. set ID 109.

Template xlsm file: This excel workbook contains the following sheets to check for anomaly in XRF data, possible samples swaps and assigning/changing the validity of samples and field blanks (FBs):

QA_Change - lists samples for which validity needs to be changed;

All - includes all the NM, QD and FBs of the monthly data set;

NMQD - lists the samples with NM and QD status and calculates the elemental loadings higher than 3 times of reported detection limits (MDLs) for further calculation/plotting.

FB - lists the field blanks;

Correlations - contains the correlation matrix of monthly data and long-term data, scatterplots of Al vs Si,, Al vs Fe, Si vs Fe, Al vs Ti, Fe vs Ti, PM vs S and PM vs K, and network metrics table (percentage of detection, MDLs, 10^{th}, 50^{th} and 90^{th} percentiles);

MassRatio - calculates and plots the ratio of sum of elements (by XRF) to PM mass (by gravimetric), and highlights the outliers;

Al vs Si - calculates the reconstructed Si loadings based on the Al measurements and Al/Si ratios of:

a) long-term (all 2011-2014 data), and b) long-term studied month. This sheet plots the measured vs constructed Si scatters, and lists the outliers;

Si vs Fe is the same as *Al vs Si*, but calculates and plots for Fe;

Al vs Fe is the same as *Al vs Si*,, but calculates and plots for Fe;

Al vs Ti is the same as *Al vsSi*,, but calculates and plots for Ti;

Fe vsTi is the same as *Al vs Si*,, but calculates and plots for Ti;

PM vs S is the same as *Al vs Si*,, but calculates and plots for S;

PM vs K is the same as *Al vs Si*,, but calculates and plots for K;

Basic_Checks - lists the samples with Fe<0, S≤0 and unusually high trace elements (e.g. Pb, Cu, Zn. Cr, Ni and As);

Outliers - lists all the outliers linked to all worksheets described above.

3. GENERAL GUIDELINES

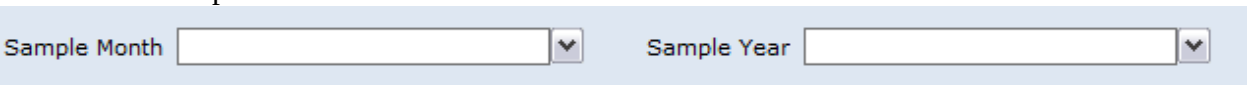
This document is intended to guide users for checking the validity of monthly XRF data, including invalidation of samples with questionable XRF results, evaluation of contamination levels on FBs (If repeated contamination for any given site the maintenance crew is notified) and detecting the anomalies of the data looking for possible sample swap(s) (e.g. FB-sample swaps or PM_{10}-sample swap). The intended audience must have fundamental knowledge of XRF operations and data. A user is required to have access to U drive, XRF Data Management pages and cl-SQL reporting pages.

4. PROCEDURES

The flowchart of procedures for XRF Level I validation is shown in Fig.1 below.

4.1 Creating Set on Webapp: The monthly XRF data set is created on webapp following the procedures given in the instruction how to use the webapp (webapp).

4.2 Accessing the XRF data on cl-sql: To pull the XRF data out, cl-sql/reports is accessed typing http://cl-sql/Reports on web browser. IMPROVE Reports>XRF Analysis>Corrected Sets are clicked. Sample Month and Year are selected on this screen:



Then, View Report button on right corner is clicked. In the next screen, the Set ID is clicked. The report can be viewed in the new screen. The Save button is clicked to select xls data view form.

In the next screen, the Open button is clicked to open the data in xls form.

The xls file is copied to QAtemplate_expt. xls file in U:\IMPROVE_Lab\XRF_Epsilon5\QA directory.

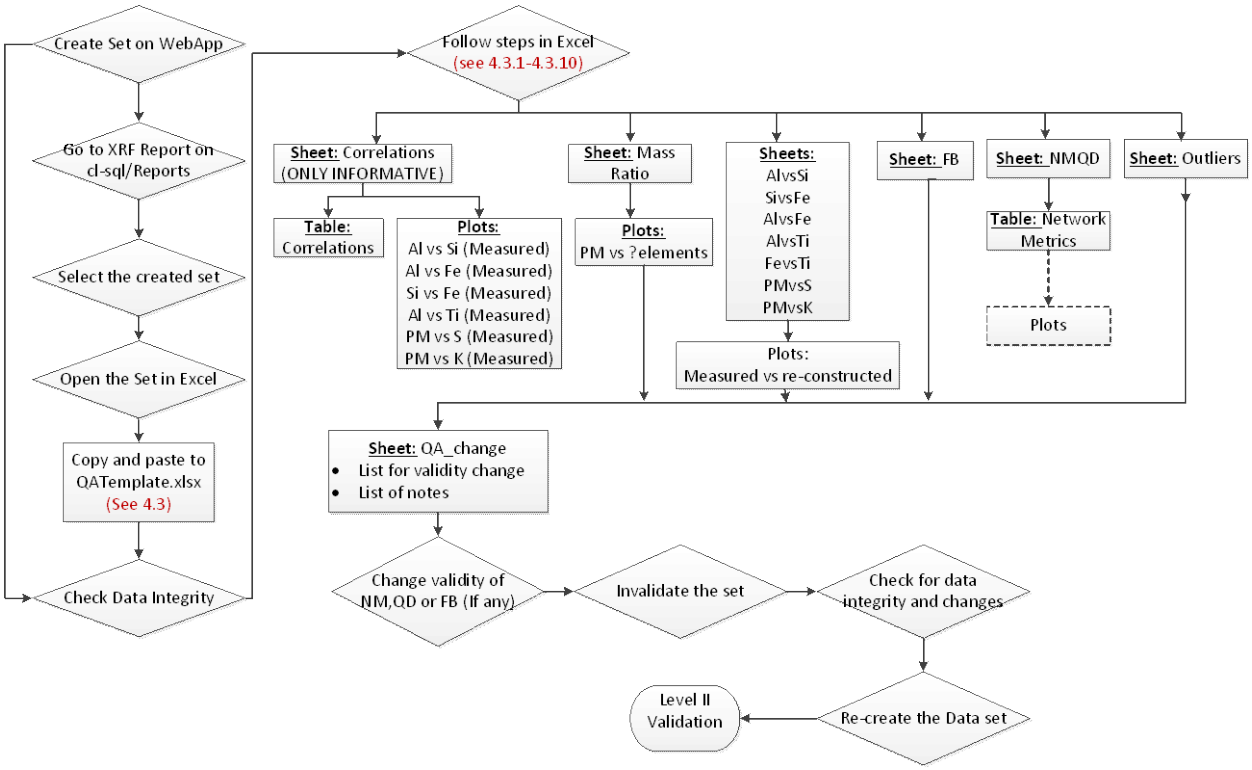


Figure 1. The flowchart of procedures for XRF Level I validation of monthly XRF data

4.3 Work in QAtemplate xls file: The template file, QATemplate_Final.xlsm in U:\IMPROVE_Lab\XRF_Epsilon5\QA, is opened and saved as "MonthYear-work.xlsm" (e.g. Jan2015-work). The copied XRF data in 4.2 is pasted into sheet-*all*. The following steps are employed to check the data:

4.3.1 Data Integrity Check: The macro (Macro1) in the workbook must be first run. This macro will update the sheets *NMQD* and *FB*. The next step is to check the data integrity. The number of samples with NM and QD is shown in cell A11 of sheet-*NMQD*. This number must be the same as the one in the data set created on webapp (Step 4.1). If not, the inconsistency must be cleared with the laboratory assistant and lab personnel. In case of any missing sample(s), e.g. samples in the logs but not XRF analyzed (supposedly must be XRF analyzed), the created data set is invalidated and the missing sample(s) are analyzed to complete the data set. When the data set is invalidated, the clear comment must be entered to explain why the data set invalidated. The FB integrity also must be checked. When the data integrity is assured, the data check steps are followed starting from checking the IMPROVE network statistics located in the cell DC1 of sheet NMQD. The QC checks are performed for all elemental mass loadings higher than 3 times their reported detection limits.

4.3.2 Correlations: The monthly correlated elements are compared with the historical values (for long term data) (2011-2014) in the sheet-*Correlations*. The monthly correlations located in cell BB11 and long-term correlations located in cell CC11 are highlighted in dark for r>0.95 and in light for r between 0.5-0.949. Normally, crustal elements, e.g., Al and Si, are highly correlated. The unusual correlations, e.g. Cu and K, must be noted for any further checks (e.g. *PMvsK* and *basic_check*). The plots in this sheet should be examined for unusual

case, e.g. Al vs Si, sample(s) deviation from Mason ratio in high concentrations. The detailed check of these plots will be performed in the sheets *AlvsSi, SivsFe, AlvsFe, AlvsTi, FevsTi, PMvsS* and *PMvsK*.

4.3.3 Mass Ratios: In this sheet, the ratio of sum_of_elements determined by XRF to particulate mass determined by gravimetric measurement is calculated for each sample. Based on the historical data, these ratios are expected to fall between 4 and 49%.

Sorting the cell AI12 by descending will list the samples outside the acceptance criteria, highlighted in dark for ratio>49% and in light for ratio<4%. The outliers must be checked sample by sample. Generally, the cases of ratios over than 49% result from contribution of sea salt to PM, thus, these cases must be checked for increase in Na and Cl. The cases of ratio lower than 4% are typical for some sites, e.g., FRES, PHOE, BYIS, BIRM or results from fire around the site. Increase in K loadings is a good sign of fire around the site in question. In addition, color of such samples will be brownish, which results from brown carbon in biomass. In case K is low and ratio <4%, samples must be visually inspected to look for dark thick deposition resulted from organic carbon (Typical for FRES and BYIS). The ratio over 100% is a good sign of filter swap. In such cases, filters in question must be reweighed to clarify the possible swaps. Sometimes, filters with ratio>100% can be very low loaded (generally <10 ug), which makes the weighing questionable. In such cases, the other samples of the same site must be checked for similar PM and elemental loading profiles. If the sample in question is different from the rest, then neighboring sites must be checked for the same sampling date sample in question. If it is different, and the re-weighing confirms the post weight, the sample must be reported to Level II validation (comment must be put in the sample editing on webapp). The Comments column on sheet-*all* can be checked for samples in question.

4.3.4 AlvsSi: This sheet plots the measured Si versus reconstructed Si based on the Al measurements and Al/Si ratios of, a) long-term (all 2011-2014 data), and b) long-term studied month (e.g., Jan2011, 2012, 2013 and 2014). Two Si-measured versus Si-constructed plots will be updated automatically except the monthly one linked to column-Q. The column-Q calculates and plots monthly data, and default is linked to slopes of Jan2011-2014 located in sheet-*Correlations,* column-DF. The *Correlations!$DF$11* term in the formula of column-Q must be replaced with studied month (e.g. *Correlations!$DG$11* for Feb), which can be easily done with CTRL+H function of Excel while selecting all column-Q. The blue highlighted cells correspond to values for all data while red ones do for month_2011-2014 data. The plots set linear regression lines with intercepts zero (blue solid lines), upper acceptance limit (red dotted lines) and lower acceptance limit (green dotted lines). The acceptance limits are calculated as 10% of range of Si-measured (cell P29). Sorting descending the cell T31 (outlier according to all data) will list the outliers. Generally the outliers fall close to the acceptance limits. If sample(s) observed very far from the acceptance limits, the reason must be investigated checking the other samples of the same site and samples of neighboring site collected in the same date of sample(s) in question. If the situation cannot be clarified, this should be noted to the Level II validation.

4.3.5 SivsFe: This sheet plots measured vs constructed Fe based on Si measurements, similar to *AlvsSi* in 4.3.4. The same steps of 4.3.4 must be followed to perform this step. The only difference in 4.3.5 than 4.3.4 is the different Si vs Fe pattern of few sites: BIRM, FRES,

PHOE, WASH and BYIS. By the filtering of cell J30, the above listed sites must be unselected to obtain representative plot for the other sites. Generally, more outliers are observed than *AlvsSi* check.

4.3.6 AlvsFe: This is the same as 4.3.5, plotting measured vs constructed Fe based on Al measurements.

4.3.7 AlvsTi: The sheet plots the measured Ti versus reconstructed Ti based on the Al measurements. The only difference than other scatterplots is that it checks the outlier with acceptance limits based on the 15% of the range, due to relatively worse Al vs Ti association than Al vs Si, and Fe.

4.3.8 TivsFe: This plots measured vs constructed Fe based on Ti measurements.

4.3.9 PMvsS: This plots measured vs constructed S based on PM2.5 mass measurements. It checks the outlier with acceptance limits based on the 30% of the range, due to worse association. The samples with high PM but low S must be checked for high K (possible fire) and color by visual inspection. Such filters should not have dark color, either brownish (fire) or grey (crustal contribution).

4.3.10 PMvsK: This plots measured vs constructed K based on PM2.5 mass measurements. It checks the outlier with acceptance limits based on the 30% of the range, due to worse association. The samples with low PM but high K must be checked possible fire by visual inspection. Such filters should have brownish color. In the New Year's eve samples, unusually high K (together with Cu and Sr) can be observed.

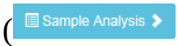
4.3.11 Basic_Checks: The samples with S zero, Fe negative, and trace elements (Cu, Zn, Pb, Cr and Ni) with unusually high loadings should be listed here. The reason of zero S must be investigated. Normally, Fe with negative loadings is no more than 10-20 samples. If more, they should be checked. Normally the trace elemental loadings in BYIS, FRES and PHOE are much higher than the other sites. If unusually high loadings are observed in the other sites, the remaining samples must be checked for similar high loadings. If unusually high loadings are observed randomly (only one sample in a month), the Level II validation must be notified (comment must be put in the sample editing on webapp).

4.3.12 Outliers: This sheet is only informative, and lists automatically all the outliers from *MassRatio* and elemental plots' sheets.

4.3.13 FB: This sheet contains the FB results. If any FB has at least two elements higher than 3 times of detection limit, the Level II validation must be notified (comment must be put in the sample editing on webapp). In case a repetitive contamination of FBs from the same site is observed, the Lab Manager must be notified for further checks and maintenance group should be informed of possible site contamination (need for the site maintenance, instrument malfunction, etc.).

4.3.14 QA_change: Contains the list of all samples with their changed validity codes. All validity changes implemented during Level I validation will have explanations placed in comments (webapp), visible to Level II validation.

4.4 Changing Validity of Samples and Invalidating the Set: On webapp, the samples in the studied set must be accessed to change the validity, if any. If no change/comment requires, no

further action is needed, and the set is ready for Level II validation. To change validity or/and comment, the View Record Details ( in http://webapp.improve.crocker.ucdavis.edu/Xrf/ProcessingCorrectedSets) is clicked. Then, Sample Analysis button (

) is clicked. The required validity changes or/and commenting is done on the samples listed in 4.3.12, if any. After changing/commenting, the current data set must be invalidated (Commenting is optional). To do that, Edit button (

) in http://webapp.improve.crocker.ucdavis.edu/Xrf/ProcessingCorrectedSets/Details/XXX is clicked. In the next screen, Valid must be unselected (

). As the last step, the data set must be re-created, and the changes/comments made must be checked. If everything is OK, then the data set is ready for Level II validation.