

Sydney Particle Characterisation Study

**PM2.5 PMF Receptor Source Database for Greater Sydney
Metropolitan Area Sites**

(2000-2014)

Instruction Manual

Prepared by

Armand J. Atanacio

Australian Nuclear Science and Technology Organisation (ANSTO)

April 2016

DISCLAIMER

This report was prepared by ANSTO in good faith exercising all due care and attention, but no representation or warranty, express or implied, is made as to the relevance, accuracy, completeness or fitness for purpose of this document in respect of any particular user's circumstances. Users of this document should satisfy themselves concerning its application to, and where necessary seek expert advice in respect of, their situation. The views expressed within are not necessarily the views of the NSW EPA and may not represent NSW EPA policy.

© 2016 State of NSW and NSW EPA

The NSW Environment Protection Authority (NSW EPA) commissioned this report from the Australian Nuclear Science and Technology Organisation (ANSTO).

© 2016 ANSTO

Contents

Figures.....	3
Introduction	5
The Receptor Source PMF Database Macro	5
Extract PMF Data and Plots	6
Export Data and plots as xlsx or pdf file	7
Description of Extracted Worksheets	9
Fingerprint Plots	10
Percentage Plots.....	11
Daily Plots.....	12
Daily Percentage Plots.....	13
Mass Plots	14
Elemental PMF versus Measured Plots.....	15
PMF Master Data	16
F-matrix.....	16
G-matrix.....	17
Tramlines	18
Additional Functions - Interactive Sampling Site Map.....	19

Figures

Figure 1 Microsoft Excel macro security warning and enable content button	5
--	---

Figure 3 Section of the Menu page showing the location of the initialise database button	6
Figure 4 Section of the Menu page showing the site name, site selection buttons and the related comments.....	6
Figure 5 Extract PMF Data and plots button	6
Figure 6 Location of summary information, worksheet navigation buttons, xlsx and pdf file export text box and buttons on Menu page.....	7
Figure 7 Section of Menu page related to xlsx or pdf file export.....	8
Figure 8 Example of the cover page worksheet included in exported xlsx or pdf files.....	9
Figure 9 Example of a set of PMF fingerprints displayed on the FingerprintPlots worksheet. The worksheet navigation button panel is also seen on the right side of the image.	10
Figure 10 Example of a set of PMF percentage plots on the PercentPlots worksheet. The worksheet navigation button panel is also seen on the right side of the image.	11
Figure 11 Example of plots on the DailyPlots worksheet showing the daily time series contribution of each PMF fingerprints in ng/m^3 . The worksheet navigation button panel is also seen on the right side of the image.....	12
Figure 12 Example of plots on the DailyPlots% worksheet showing the each PMF fingerprints daily percentage (%) contribution. The worksheet navigation button panel is also seen on the right side of the image.....	13
Figure 13 Example of plots on the MassPlots worksheet comparing PMF mass with gravimetric mass. The worksheet navigation button panel is also seen on the right side of the image.	14
Figure 14 Example of plots on the PMF versus IBA plots worksheet comparing PMF fitted concentration (ng/m^3) against the ion beam analysis (IBA) measured concentration (ng/m^3) for each element. The worksheet navigation button panel is also seen on the right side of the image.	15
Figure 15 Example section of the PMF Data worksheet showing the cell reference method, for example cell E10, as show in the image. Please note: the worksheet navigation button panel is in a slightly different format/location to previous worksheets and is seen along the top of the worksheet.	16
Figure 16. Location of hyperlinks that will open an interactive Google map showing sampling site locations in your internet browser. NB: This function requires an active internet connection.	19

Introduction

This document accompanies the Sydney Particle Characterisation Study - Receptor Source Positive Matrix Factorisation (PMF) Database File.

The aim of this document is to provide instructional steps and related information necessary for navigating and utilising the PMF database. It is important to note that interpretation of the receptor source PMF fingerprints and apportionment contained in the database is beyond the scope of this document and can be found in the full study report.

The Receptor Source PMF Database Macro

The receptor source PMF fingerprint database is provided as the following zipped file: *SydneyParticleStudy2000-14Macro.zip*. This zipped file contains the following three files: (1) the main VBA macro-enabled excel file: *SydneyParticleStudy2000-14.xlsm*, (2) *SamplingSiteMap.html*, and (3) an electronic copy of this instruction manual. Please note, for the remainder of this document, the Excel macro file will be referred to as the *database*. In an effort to minimise the size of this database macro, it only contains the PMF data (without the associated PMF plots) from each of the sites (15 years of PMF data for one site alone is over 2Mb!). The Excel visual basic for applications (VBA) macro functions are then utilised to automatically generate the associated PMF plots from the data for each site as required.

The macro options are accessed from the **Menu** worksheet and their functions described in the following sections of this document. Please note: these macro functions have been written and tested to operate correctly in Excel version 2007-2010.

Important: Depending on your Microsoft Excel settings, you may encounter a “Security Warning Macros have been disabled” alert when you open the PMF Master database files (see Error! Reference source not found.). You will need to press the “Enable Content” button for the macros to function correctly.

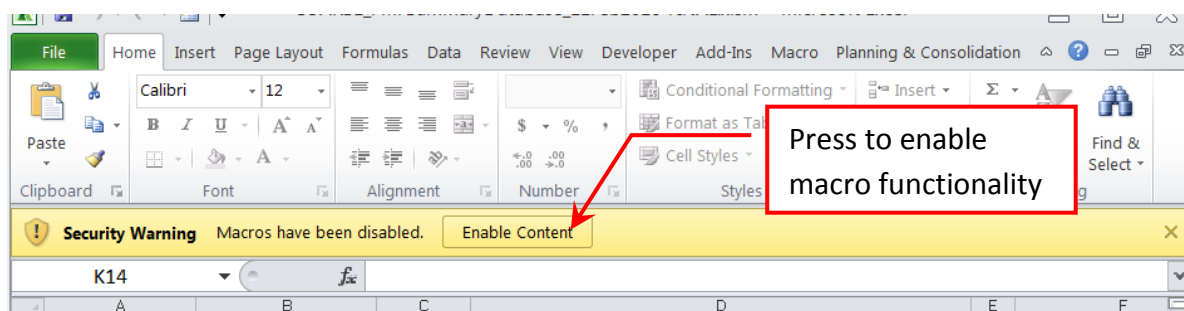


Figure 1 Microsoft Excel macro security warning and enable content button

When you open the database for the first time, you will land on the “Menu” worksheet (See **Figure 2**). For convenience, a brief version of instructions is also available at the top of this page (**Figure 2**). To begin using the database, the program and data must first be initialised. This process clears any previous data and plots and also prepares the required macro functionality. Initialisation is performed by pressing the button labelled “**(1) INITIALISE**”

DATABASE” (Figure 2). A progress bar will be displayed in bottom left of the database window.

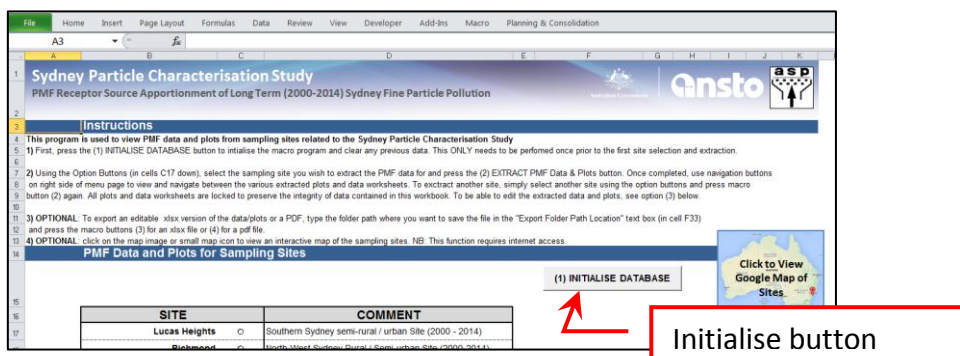


Figure 2 Section of the Menu page showing the location of the initialise database button

Extract PMF Data and Plots

Now select the site to extract the PMF data. This is done by clicking on one of the site selection buttons (**Figure 3**). Please note, only one site can be selected for extraction at a time.

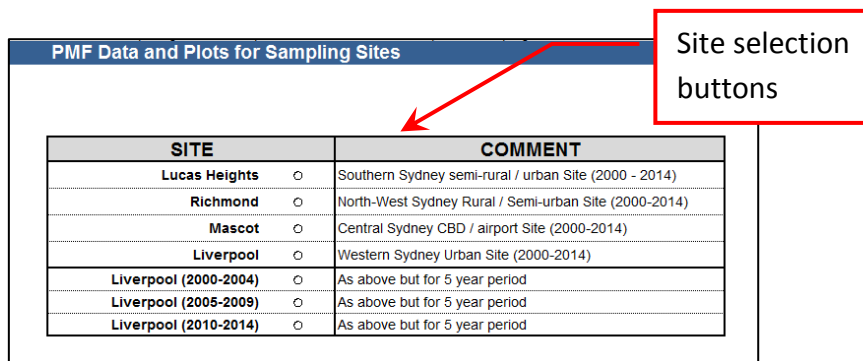


Figure 3 Section of the Menu page showing the site name, site selection buttons and the related comments

Once a site has been selected, press the **“(2) EXTRACT PMF Data & PLOTS”** button (**Figure 4**). As there is a significant amount of data for each site, the extraction and plotting may take up to a minute to complete depending on the processing speed of your computer. The progress bar will be shown in the bottom left corner of the excel window during the extraction process.

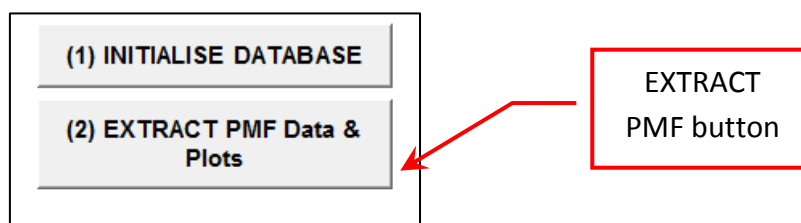


Figure 4 Extract PMF Data and plots button

Once the extraction is complete, the Menu worksheet will be updated with additional SUMMARY INFORMATION, worksheet navigation buttons and file export buttons (**Figure 5**).

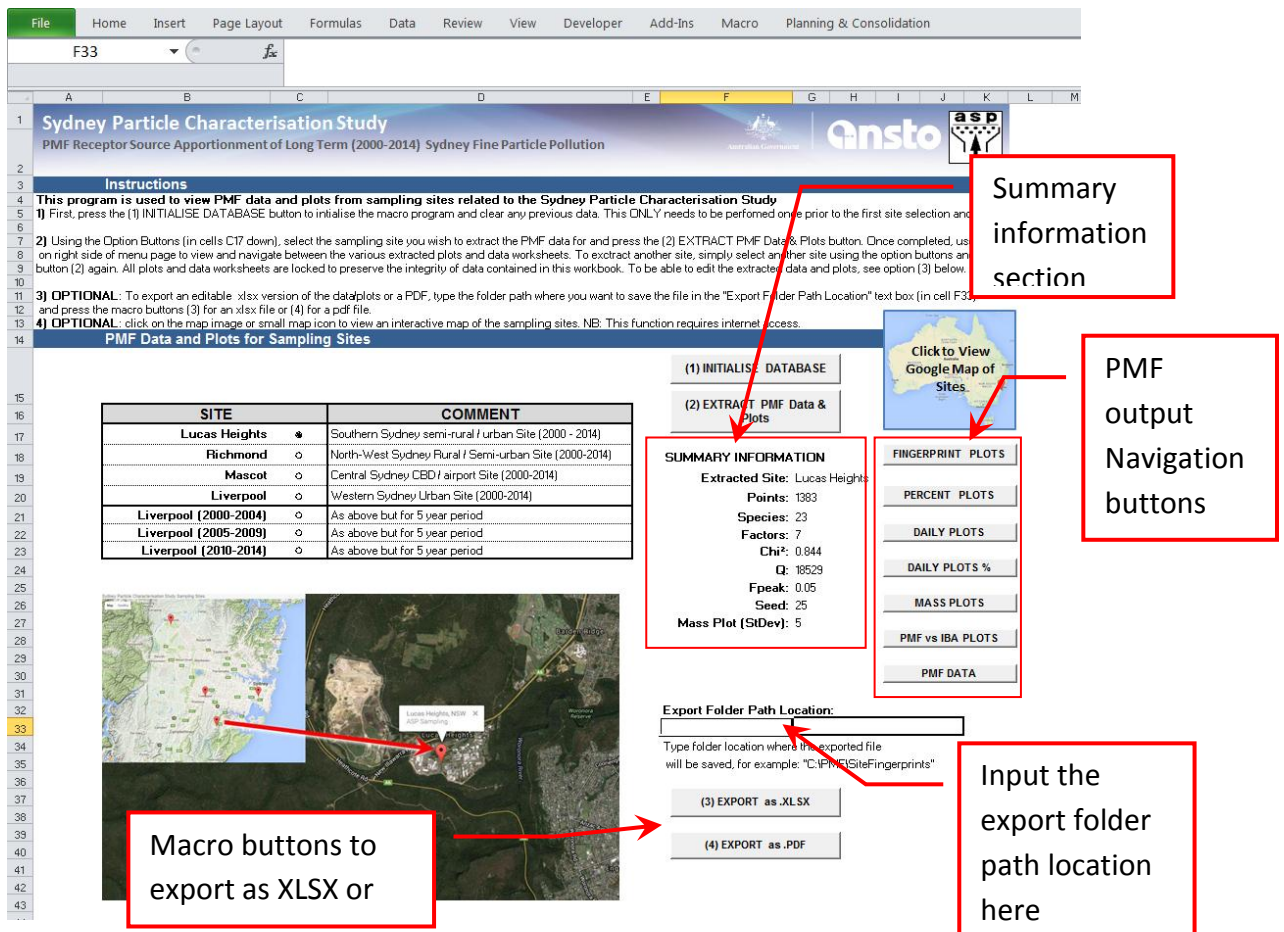


Figure 5 Location of summary information, worksheet navigation buttons, xlsx and pdf file export text box and buttons on Menu page

This summary information provides the following details: the extracted site, number of points (i.e. sampling days) used in the PMF analysis, number of species (or elements) used in the PMF analysis, number of fingerprints obtained with PMF analysis, chi-squared (χ^2) value, the Q-value, F-peak value, Seed value and the number of standard deviations used for the tramlines on the Mass Plot (Figure 5).

The Menu worksheet will also be updated with additional navigational buttons which can be used to go between the sites PMF data and plots (Figure 5). These navigation buttons are also available on each of the generated worksheet.

Export Data and plots as xlsx or pdf file

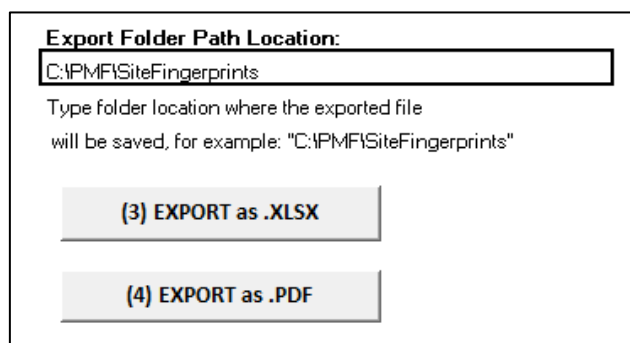
XLSX

The various data and plots can be viewed within the database using the navigational button. However, all of the database worksheets are **locked** and cannot be edited. This is to maintain the integrity of the data contained in the database program itself. If you would like to edit the extracted data or plots (for example, to re-plot the data in another program), you will need to export the extracted data as an **unlocked** xlsx file. To do this, type an existing folder path location where you would like this .xlsx file to be exported to in cell F33, for example, "C:\PMF\Australia\fingerprints". Now press the macro button "(3) EXPORT as

.XLSX” to export the file (see **Figure 5**). The filename of the exported xlsx file will be automatically generated as the selected site name. This xlsx file will contain the following: a summary information cover page (CoverPage) (see **Figure 7**), worksheet containing the PMF data (named as the site that was extracted), worksheet of fingerprint plots (FingerprintPlots), worksheet of percent plots (PercentPlots), worksheet of daily plots (DailyPlots), worksheet of daily plot percent (DailyPlots%), worksheet of PMF Mass vs gravimetric Mass plots (MassPlots) and a worksheet of PMF vs IBA plots for each specie.

PDF

To export a pdf file, follow the same process described above for the xlsx file but press the macro button “**(4)EXPORT as .PDF**”. The generated pdf file contains the same cover page and plots as the xlsx file, however, it does not include the PMF data worksheet – as converting this worksheet to pdf can fill more than 20 pages depending on the amount of data associated with the selected site. The filename of the pdf file will be automatically generated to be the same as the selected site.



Export Folder Path Location:
C:\PMF\SiteFingerprints
Type folder location where the exported file will be saved, for example: "C:\PMF\SiteFingerprints"
(3) EXPORT as .XLSX
(4) EXPORT as .PDF

Figure 6 Section of Menu page related to xlsx or pdf file export

Cover Page

A cover page is automatically generated and included when you export the data as either an xlsx or pdf file (**Figure 7**). This cover page provides the date the data was extracted, site, comment (with space to input additional comments if required), PMF analysis summary information, and a map showing the immediate vicinity of the site (**Figure 7**).

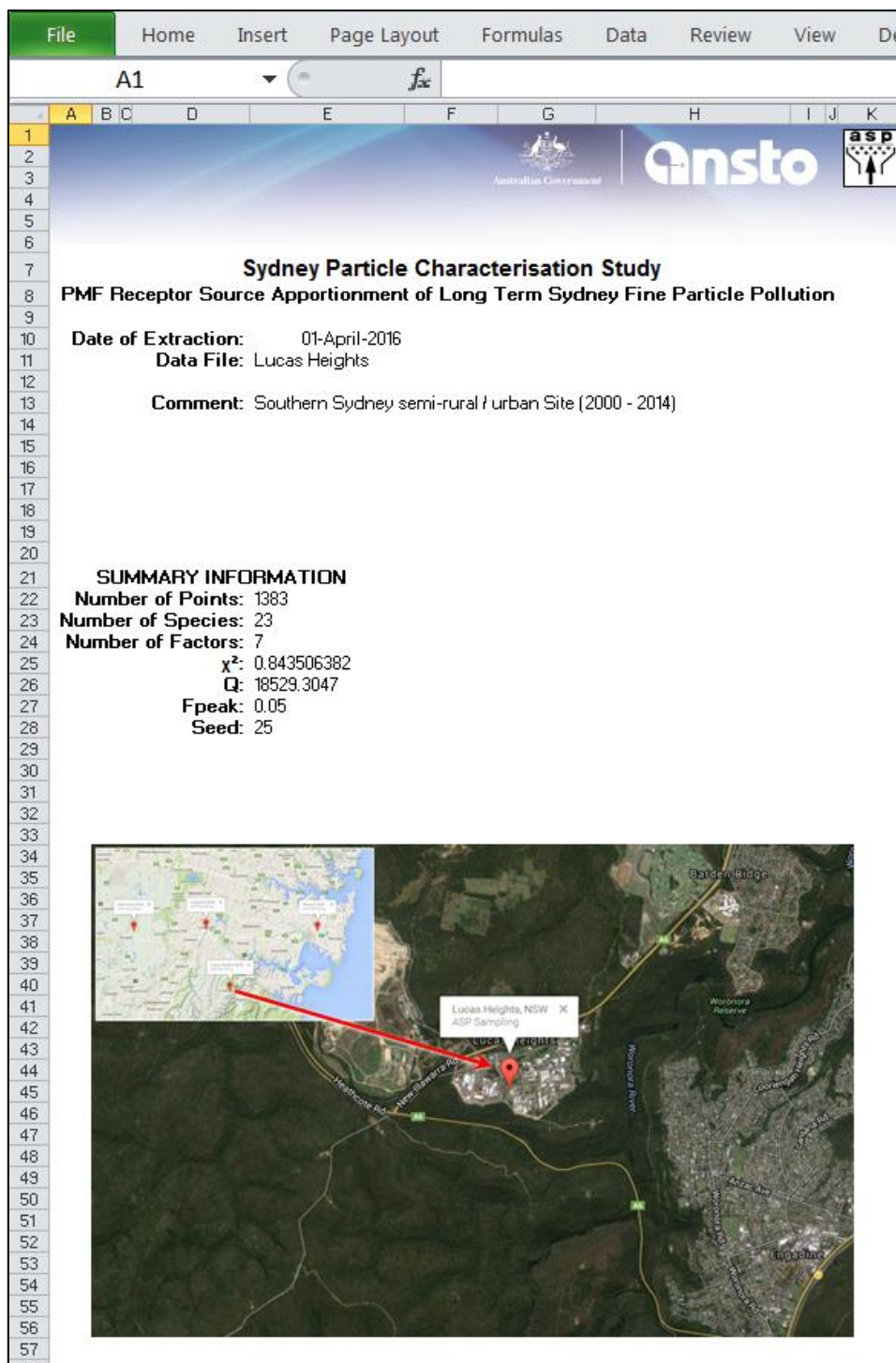


Figure 7 Example of the cover page worksheet included in exported xlsx or pdf files

Description of Extracted Worksheets

The following sections describe the various worksheets generated from the PMF data extraction process described earlier.

Fingerprint Plots

The PMF fingerprint plots are displayed on the FingerprintPlots worksheet (**Figure 8**). Each fingerprint is comprised of the fractional ratios for each correlating element. The elemental ratios in each fingerprint have been normalised to have the maximum element with a value of 1. This clearly identifies the main driving element/s for a particular fingerprint and assist the data analyst in assigning possible fingerprint names. The Y-axis of each fingerprint is a 4-decade log plot which allows trace elements to also be easily identified. The error bars relate to 3 standard deviations calculated by the PMF analysis codes. The percentage value in brackets next to the title of each fingerprint denotes the percentage of that fingerprint to the total 'fitted' mass, not the total gravimetric mass.

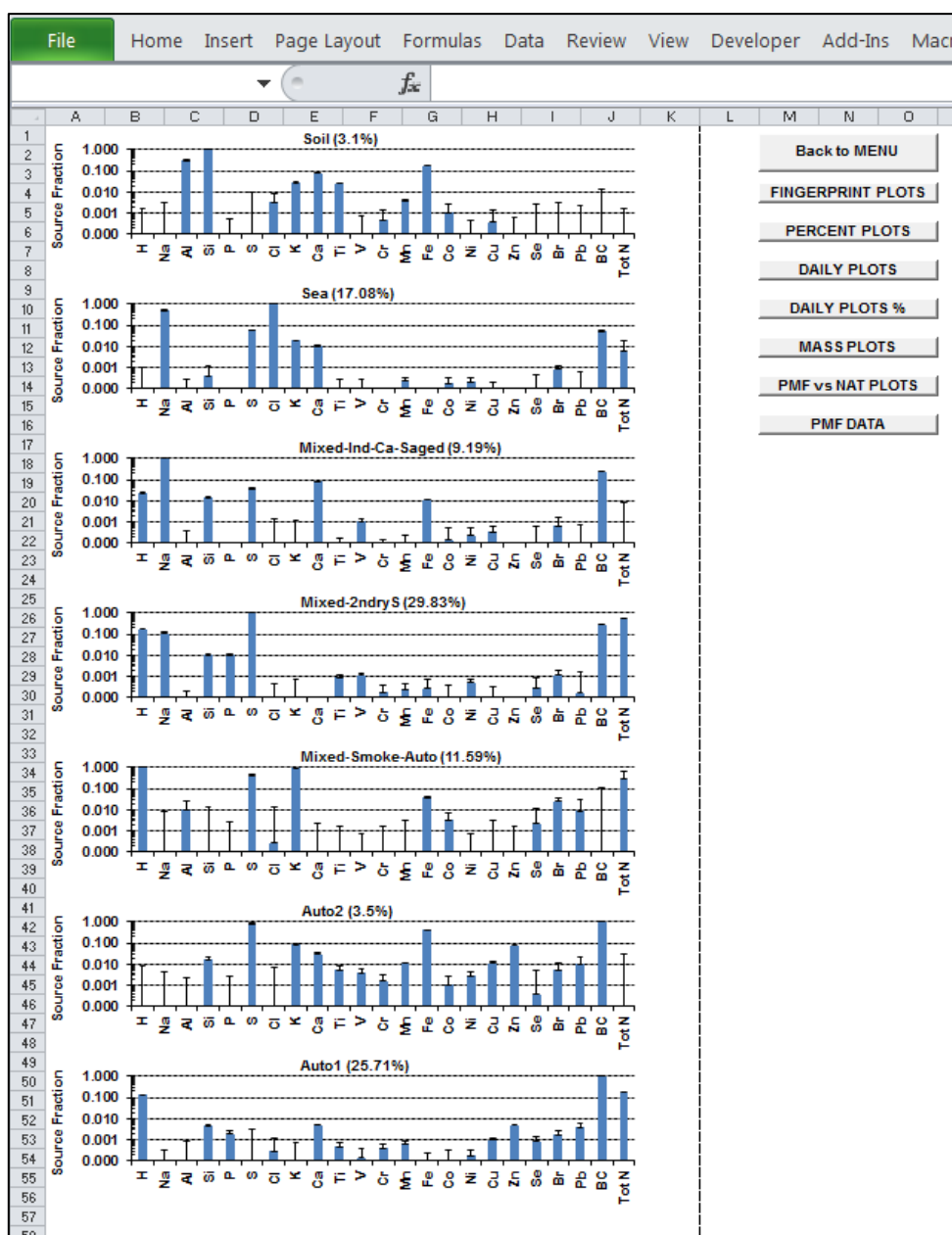


Figure 8 Example of a set of PMF fingerprints displayed on the FingerprintPlots worksheet. The worksheet navigation button panel is also seen on the right side of the image.

Percentage Plots

The plots on the PercentPlots worksheet represent the percentage distribution of each of the elements across all identified fingerprints. For example, adding the Na percentage shown in each of the 7 fingerprints (see **Figure 9**) will account for 100% of the Na fitted mass. These plots are NOT source fingerprints and should not be used as such. They are useful to assess where the major contributions of each element reside in terms of fingerprints. Correct PMF fingerprint interpretation requires the use of both the fingerprints plots (**Figure 8**) and the percentage plots (**Figure 9**) in conjunction to best to identify meaningful receptor sources.

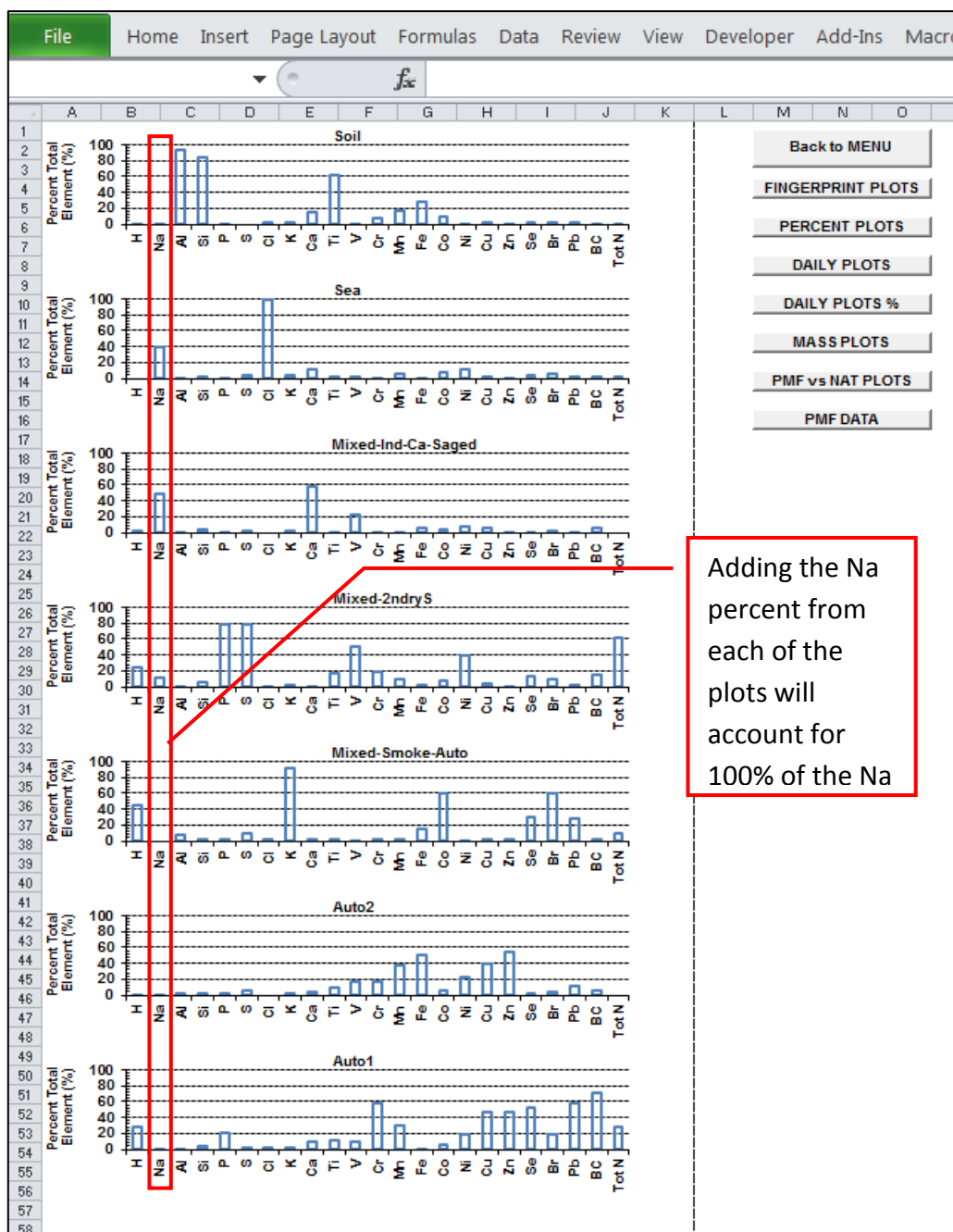


Figure 9 Example of a set of PMF percentage plots on the PercentPlots worksheet. The worksheet navigation button panel is also seen on the right side of the image.

Daily Plots

The plots on the DailyPlots worksheet (**Figure 10**) represent the daily contribution of each fingerprint in nanograms (ng/m^3). The date on the x-axis has the format DD-MMM-YY, for example 02-Jan-14 denotes Day = 02, Month = January and Year = 2014.

These plots are useful to observe daily, monthly and yearly trends for each identified fingerprint. Such trends may include: gradual yearly reduction in receptor sources as a result of the implementation of pollution reduction policies, regular monthly seasonal variations of particular receptor sources (e.g. summer-winter), or extreme daily events such as dust storms or bush fires.

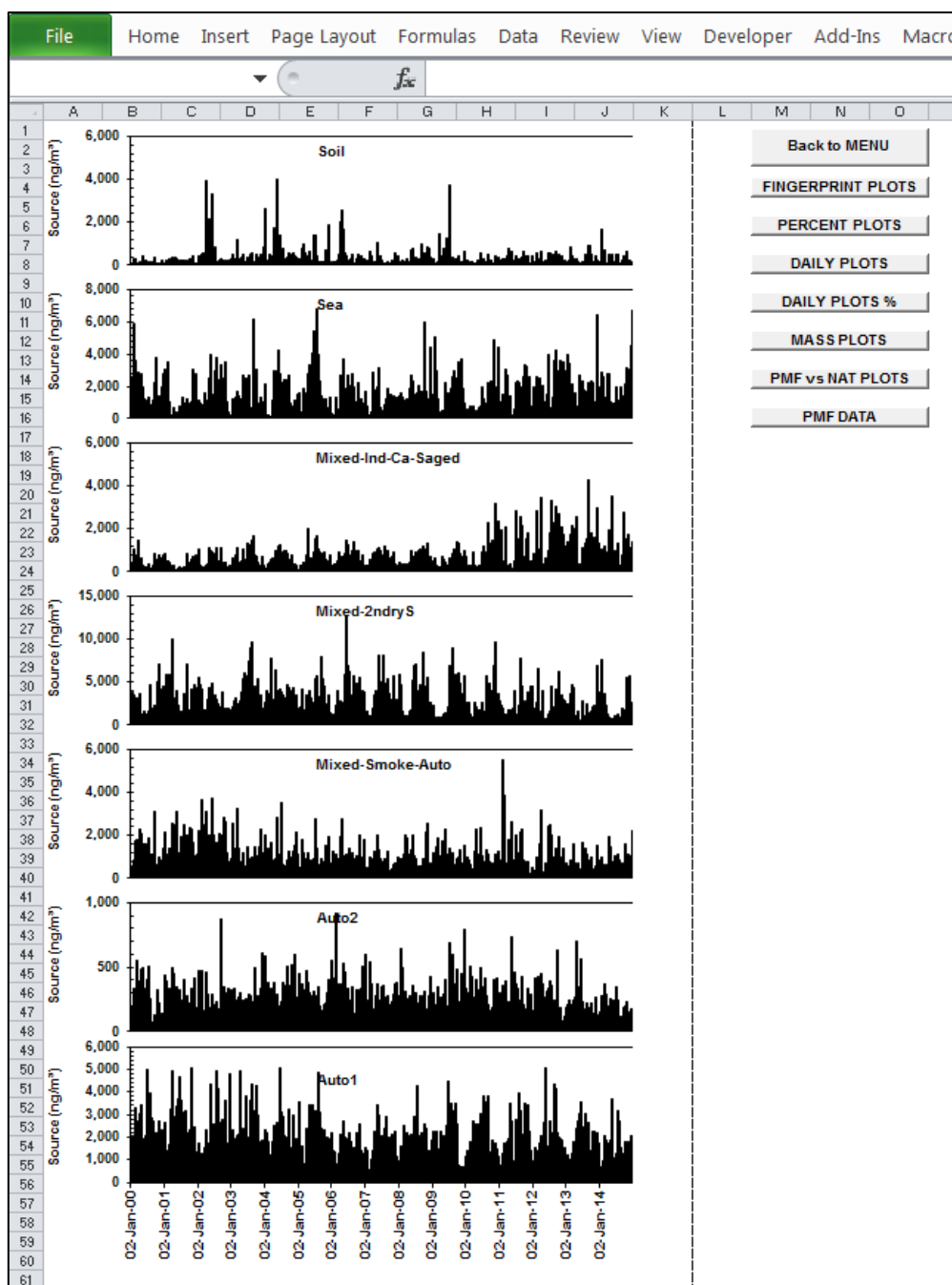


Figure 10 Example of plots on the DailyPlots worksheet showing the daily time series contribution of each PMF fingerprints in ng/m^3 . The worksheet navigation button panel is also seen on the right side of the image.

Daily Percentage Plots

The plots on the DailyPlots% worksheet (**Figure 11**) are similar to the daily plots (**Figure 10**) but instead of daily concentration in ng/m^3 , these plots represent the percentage (%) contribution of each identified receptor source to total pollution as a daily time series. The date on the x-axis has the format DD-MMM-YY, for example 02-Jan-14 denotes Day = 02, Month = January and Year = 2014.

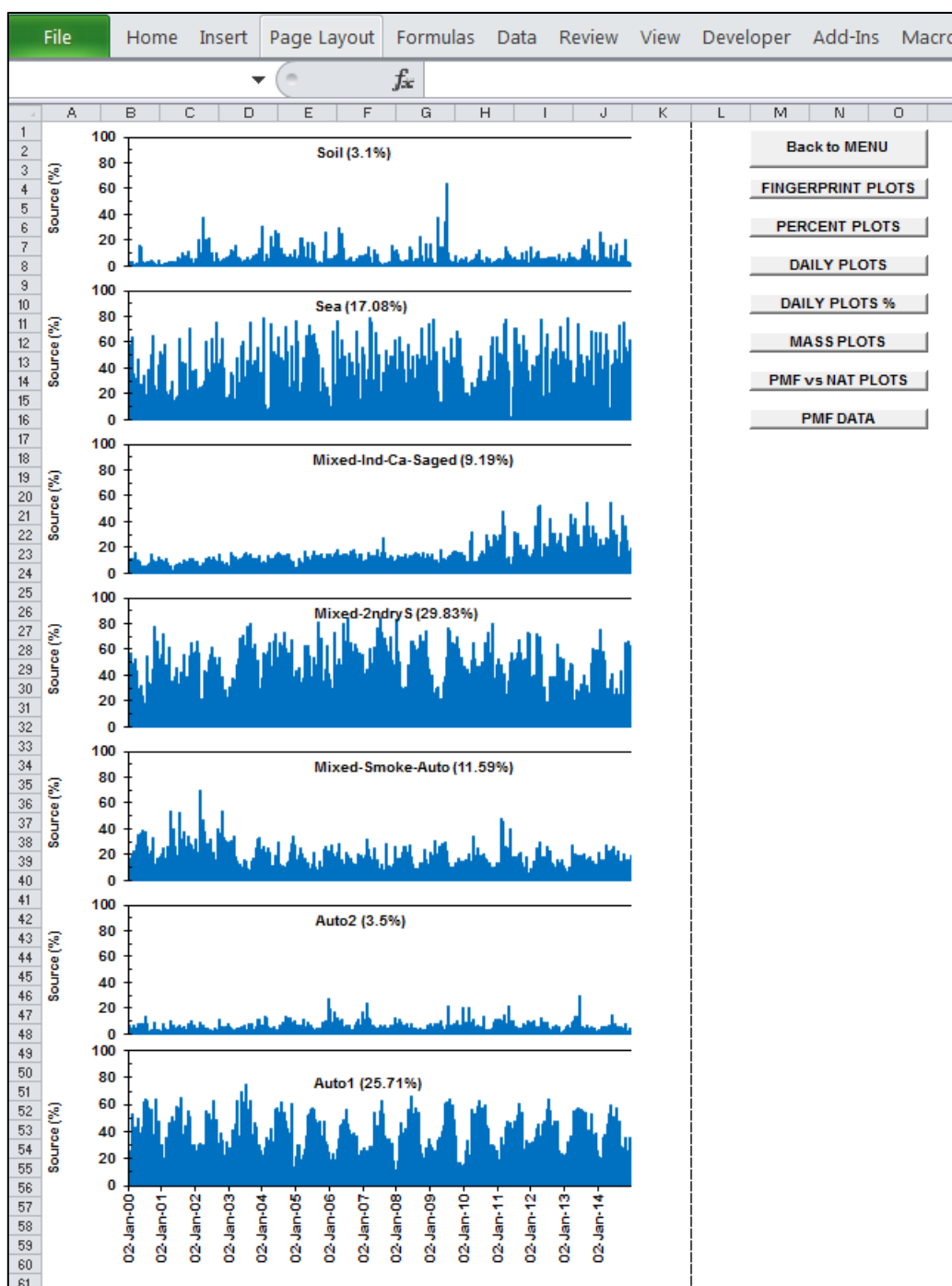


Figure 11 Example of plots on the DailyPlots% worksheet showing the each PMF fingerprints daily percentage (%) contribution. The worksheet navigation button panel is also seen on the right side of the image.

Mass Plots

The MassPlots worksheet contains two mass related plots. The first plot compares the calculated PMF mass with the measured gravimetric mass, both in ng/m^3 . Ideally, both the linear fit and R^2 value should be close to 1. However, this linear fit can be skewed significantly by outlying points causing the fitted data to be representative of neither the majority of data nor the outlying point, but somewhere in-between. To avoid this, the data analyst generally tries to remove most of the outlying or extreme points outside of the “tramlines” represented by the dotted lines in the top plot of **Figure 12**. These tramlines represent a certain number of standard deviations (SD) from the fitted line, generally between 3 to 6 standard deviations. The tramline SD value which was applied in each site’s PMF analysis is provided in the SUMMARY INFORMATION section (**Figure 5**).

The second plot on the MassPlots worksheet is a time series comparison of both the PMF mass and gravimetric mass. It is used to check that the daily PMF fit obtained from statistical analysis matches the measured daily gravimetric mass.

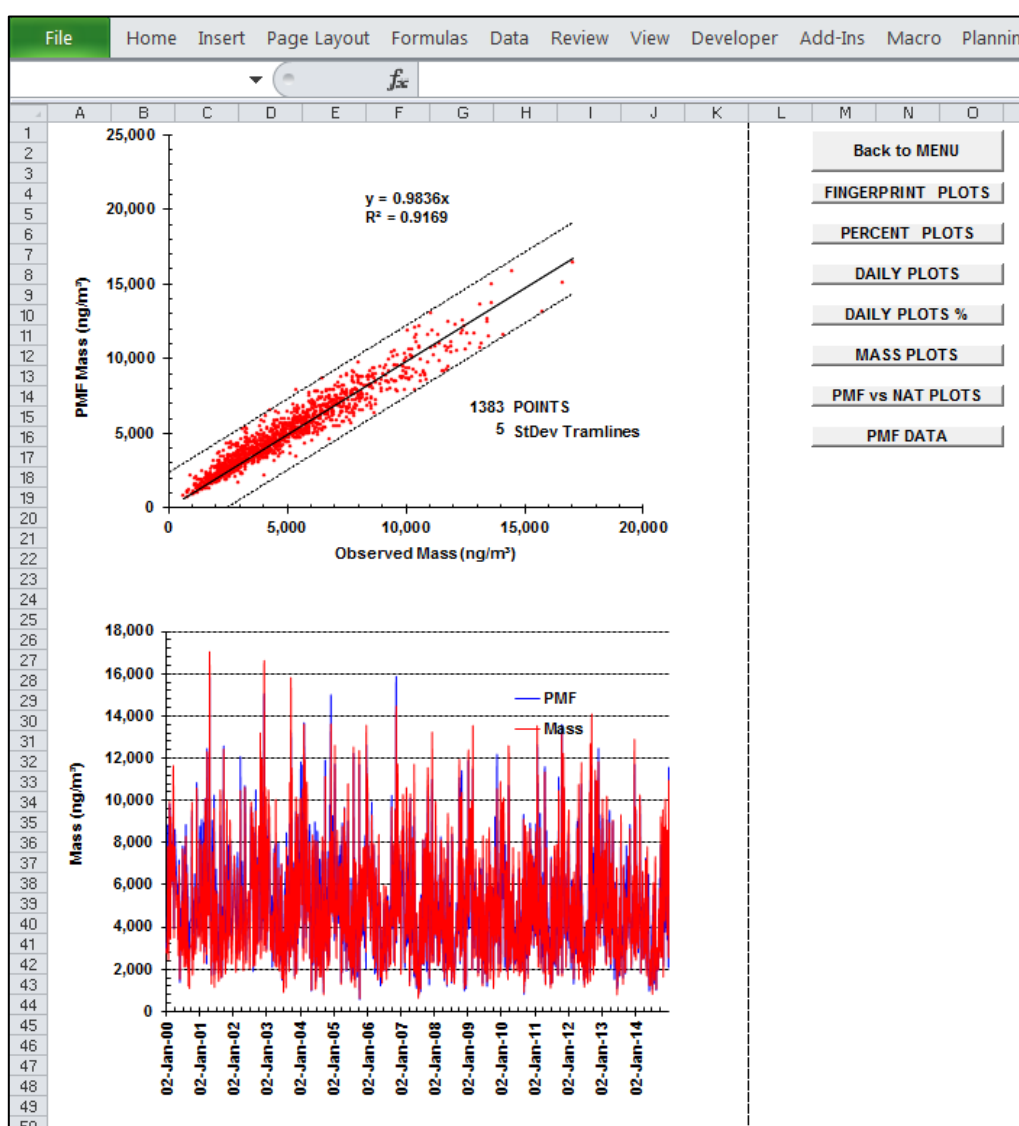


Figure 12 Example of plots on the MassPlots worksheet comparing PMF mass with gravimetric mass. The worksheet navigation button panel is also seen on the right side of the image.

Elemental PMF versus Measured Plots

The plots on the PMFvsIBApots worksheet (**Figure 13**) represent the correlation of each elements PMF fitted value against its measured concentration value obtained with accelerator ion beam analysis (IBA). These plots provide a clear visualisation, element by element, of which elements have been fitted well by the PMF process. Ideally, the obtained gradient and R^2 for each element would be 1 if the PMF fit was perfect. This linear fit can be skewed significantly by outlier points resulting in a solution that is representative of neither the majority of data for that element nor the outlying point, but somewhere in-between. Therefore, during the PMF analysis process, the data analyst is able to systematically remove outliers from these plots in order to get closer to this ideal condition. However, for trace elements or elements with very large associated errors/MDLs, this is not often possible nor even advantageous as the PMF process will always be driven by elements with more accurate and precise measurements. These plots are also useful during the PMF analysis process in determining which elements can be removed from the PMF analysis as they have minimal impact on the final fits.

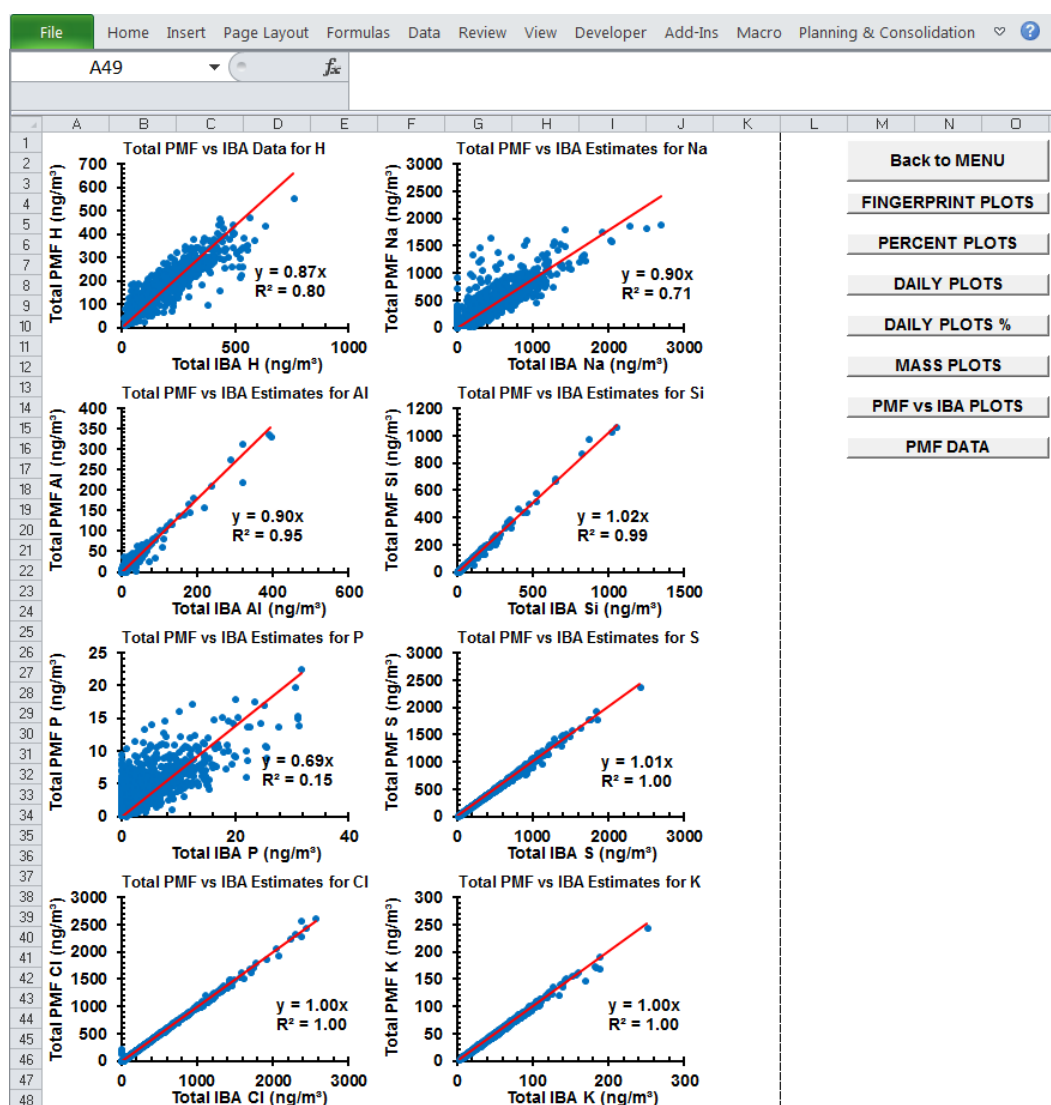


Figure 13 Example of plots on the PMF versus IBA plots worksheet comparing PMF fitted concentration (ng/m^3) against the ion beam analysis (IBA) measured concentration (ng/m^3) for each element. The worksheet navigation button panel is also seen on the right side of the image.

PMF Master Data

The data contained on the worksheet with the selected sampling site name (e.g. Lucas Heights) represents the PMF Master Data related to that site and from which all of the previously described plots are generated. The following section of the present document will describe the various sections of data located on the PMF Master Data worksheet using the excel cell reference method of letter and column, e.g. "E10" refers to column E and row 10 (see Figure 14).

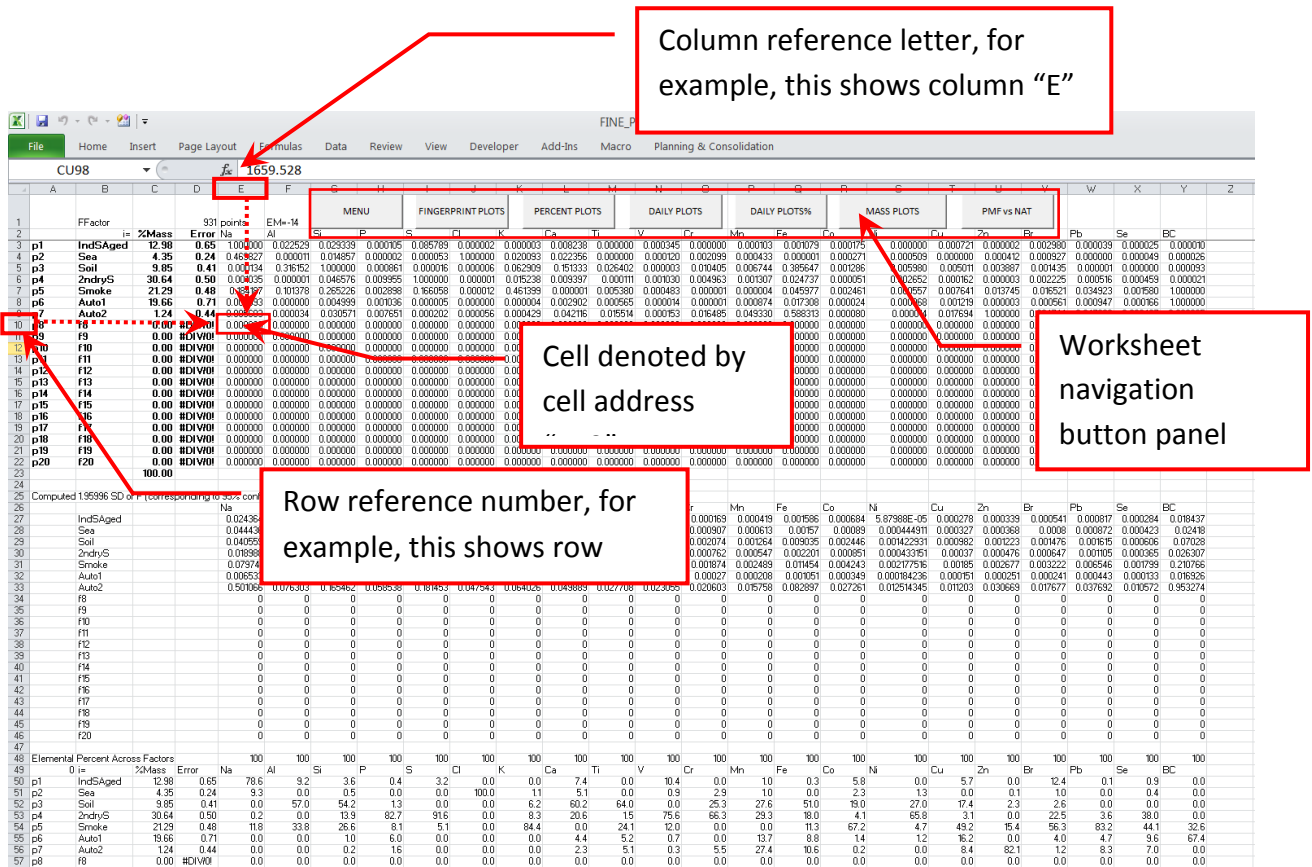


Figure 14 Example section of the PMF Data worksheet showing the cell reference method, for example cell E10, as show in the image. Please note: the worksheet navigation button panel is in a slightly different format/location to previous worksheets and is seen along the top of the worksheet.

Important Note

When viewing the PMF Master Data worksheet for each member state, it should be noted that the PMF analysis program which generates this Master Data worksheet is capable of processing up to 20 different PMF fingerprints, but in reality we would rarely use that many fingerprints. Most sites to date are adequately accomodated by between 5-10 fingerprints. Therefore, you will notice many cells containing either zero or #DIV/0! values as these are related to unused factors of the available 20.

F-matrix

Rows 3 to 22 (across): Data related to fractional ratio contribution of each element in each fingerprint. It represents the F-matrix. In this data, the element with the largest contribution in each fingerprint has been normalised to 1. The fingerprint names are listed in cells B3 to

B22, with their corresponding percentage mass contributions and associated error listed in cells C3 to C22 and D3 to D22, respectively. The elements included in the PMF analysis are listed in row 2 from column E and onwards across the worksheet (depending on how many elements were included in the analysis). This data is used to generate the fingerprint plots.

Cell D1: Number of points (i.e. days) included in the PMF analysis.

Rows 26 to 46 (across): Data related to calculated 95% confidence interval of the F-matrix data from column B and onwards across the worksheet depending on how many elements included in the analysis. This data is used to generate the error bars on each element in the fingerprint plots.

Rows 48 to 69 (across): Data related to the percentage distribution of each element across all of the fingerprints from column A and onwards across the worksheet depending on how many elements included in the analysis. Therefore, the sum of each element's percentage contribution in each fingerprint, e.g. cells E50 to E69, results in a value of 100 percent as show in row 48 from columns E onwards (depending on how many elements were included in the analysis)

Rows 70 to 90 (across): Data related to manual Fkey pull-down strength used during the PMF analysis from column A and onwards across the worksheet depending on how many elements included in the analysis. Column B70 to B90 lists the fingerprint names related to that row of data. Columns E (for rows 70-90) onwards relate to the elements in each fingerprint (as listed in row 2). For example, cell E71 would refer to the Fkey value applied to the first element in the first fingerprint.

The optional Fkey option available during PMF analysis refers to the use of known external information to impose additional control on the rotation of the PMF analysis. For example, if specific elements in specific values are **known** to be zero, such as black carbon (BC) in a pure sea spray fingerprint, then assigning an Fkey value to that element can be used to force the PMF solution toward zero for BC values in that sea spray fingerprint. A range of Fkey pulldown strength may have been used between one, which has minimal pull-down strength, to nine which has the strongest Fkey pulldown strength. A value of zero, denotes no Fkey pull-down has been used.

Cell A96: Shows the year range of the dataset used in the PMF analysis

G-matrix

Cells A96 to AA97 (down): Data related to daily contribution of each fingerprint to total mass. This data represent the G-matrix. The names of the fingerprints are listed horizontally from cell F96 to Y96 with their data listed in the corresponding column below each name. Cells A97 to D97 (downward depending on the number of analysed days) relate to the corresponding Site, Day, Month and Year of each G-matrix data row. Column E97 and Z97 relate to the gravimetric mass concentration (Cmass) and fitted PMF mass concentration (FitCmass), respectively. Column AA97 relates to the date (DD-MMM-YY) used in several of the plot worksheets described earlier.

Rows E92 and E92 (across): Data related to the average (mean) and standard deviation values, respectively, for the columns in “Cells A96 to AA97 (down):” described above.

Cell E94: Q-value obtained from analysis

Cell G94: Fpeak value used for analysis

Cell J94: Chi-squared value obtained from analysis

Cell M94: Seed value used for analysis

Row E95 to AA95 (across): Least squares linear regression coefficient values corresponding to each of the G-matrix columns from F98 to Y98 (down).

Rows AB96 AU96 (down): Daily percentage of each fingerprint to the total PMF fitted mass concentration (i.e. FitCmass). Summation of each row, e.g. AB98 to AU98 which represents the individual percentage contribution of each fingerprint for that day, should equal a value of 100%.

Tramlines

Columns AW97 to AZ97 (down): Data used to create the standard deviation tramlines on the PMF mass vs. Observed Mass plot (see MassPlots worksheet, **Figure 12**).

Cell BA97: Standard deviation tolerance for the upper and lower tramlines

Cell BB97: Standard deviation of the values listed in the “Distances” column A97 (down)

Additional Functions - Interactive Sampling Site Map

NB: this function requires **internet connection** and the "SamplingSiteMap.html" file must be located in the same folder as the Database file.

Clicking the interactive map button on the Menu worksheet accesses generates a map of the sampling sites which it displays in a new internet browser window using Google Maps.

NB: If this function fails to open correctly from within the database, please try opening the SamplingSiteMap.html file directly using an internet browser other than Internet Explorer (.e.g. Chrome).

Instructions

1) First, press the (1) INITIALISE DATABASE button to initialise the macro program and clear any previous data. This ONLY needs to be performed once prior to the first site selection and extraction.

2) Using the Option Buttons (in cells C17 down), select the sampling site you wish to extract the PMF data for and press the (2) EXTRACT PMF Data & Plots button. Once completed, use navigation buttons on right side of menu page to view and navigate between the various extracted plots and data worksheets. To extract another site, simply select another site using the option buttons and press macro button (2) again. All plots and data worksheets are locked to preserve the integrity of data contained in this workbook. To be able to edit the extracted data and plots, see option (3) below.

3) OPTIONAL: To export an editable .xlsx version of the dataplots or a PDF, type the folder path where you want to save the file in the "Export Folder Path Location" text box (in cell F33) and press the macro buttons (3) for an .xlsx file or (4) for a pdf file.

4) OPTIONAL: click on the map image or small map icon to view an interactive map of the sampling sites. NB: This function requires internet access.

SITE	COMMENT
Lucas Heights	Southern Sydney semi-rural / urban Site (2000 - 2014)
Richmond	North-West Sydney Rural / Semi-urban Site (2000-2014)
Mascot	Central Sydney CBD / airport Site (2000-2014)
Liverpool	Western Sydney Urban Site (2000-2014)
Liverpool (2000-2004)	As above but for 5 year period
Liverpool (2005-2009)	As above but for 5 year period
Liverpool (2010-2014)	As above but for 5 year period

(1) INITIALISE DATABASE
(2) EXTRACT PMF Data & Plots

SUMMARY INFORMATION
Extracted Site: Lucas Heights
Points: 1282

Click to View Google Map of Sites

Click on either map to open a Google map of the sampling sites in your default web

(3) EXPORT as
(4) EXPORT

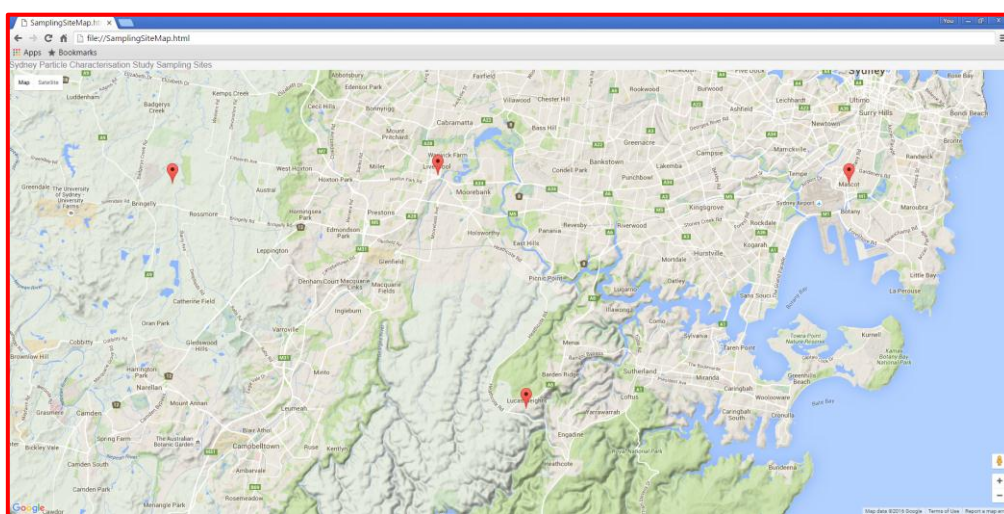


Figure 15. Location of hyperlinks that will open an interactive Google map showing sampling site locations in your internet browser. NB: This function requires an active internet connection.



Nuclear-based science benefiting all Australians