

Importing Text Files in DAX

DAX has the ability to import many different formats of text files. It will search for a number of features in each text file, and will make a best estimate as to how to interpret each file. The user will then be asked for confirmation.

The features DAX looks for are:

- zero or more header lines that do not contain point data, but may contain frequency information
- one or more data lines that contain one or more columns
- columns may contain an index, time (horizontal axis) coordinates, or point values; DAX will try to guess what type of data are in each column

An example

As an example, let us take a look at a file named IRAM1B.A01 containing molecular weight data.

```
= "File Information" 9 18
= "Sample Information" 20 27
= "Comment" 29
= "Peak Data" 31 34
= "Slice Data" 36 245
= "Total Average Molecular Weight" 247 249
= "Sectional Average Molecular Weight" 251 252

= "File Information"
= "Peak Number" 2
= "Slice Number" 208
= "GPC Method File Name" "P41102.GMT"
= "Class Data File Name" "IRAM1B.D01"
= "LC System Number" 1
= "Chromatogram Name" "IRAM1B.C01"
= "Background Chromatogram Name" ""
= "Class Method File Name" "PATRICIA.MET"
= "Date & Time" "02/11/0518:31:10"

= "Sample Information"
= "Operator Name" ""
= "Sample Name" ""
= "Sample ID" ""
= "Sample Type" "U "
= "Attenuation" 3
= "Flow" 0
= "Slice Rate" .5

= "Comment"

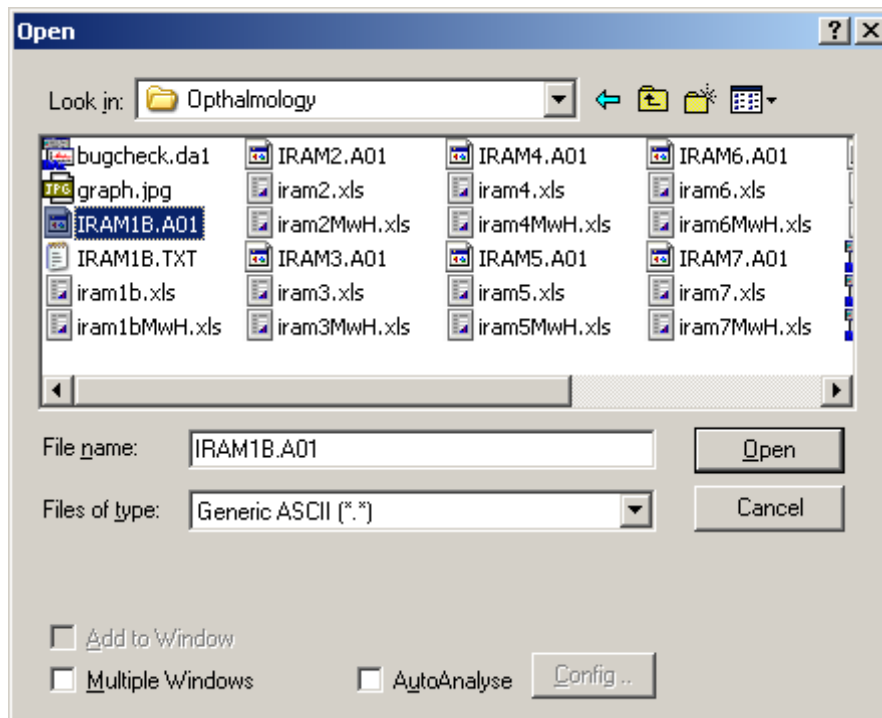
= "Peak Data"
= "Separation" = "Top Height" = "Start Time" = "Top Time"
16 2268 19.933 23.750
16 2847 33.933 35.666

= "Slice Data"
= "Peak Number" = "Retention Time" = "Volume" = "Molecular Weight" = "Height"
1 20,00 0,000 299808 27680
1 20,06 0,000 290759 33880
1 20,13 0,000 281983 50620
1 20,20 0,000 273471 65740
1 20,26 0,000 265217 89240
1 20,33 0,000 257212 109060
<<lines omitted>>
2 37,46 0,000 98 75160
2 37,53 0,000 95 37960
2 37,60 0,000 92 2960
2 37,66 0,000 89 -3100
2 37,73 0,000 86 -12400
2 37,80 0,000 84 -24180
2 37,86 0,000 81 -28080
2 37,93 0,000 79 -16100

= "Total Average Molecular Weight"
= "Mn" = "Mw" = "Mz" = "Mz1" = "Mv" = "I.Visc"
614 38912 94591 131166 0 0.00000

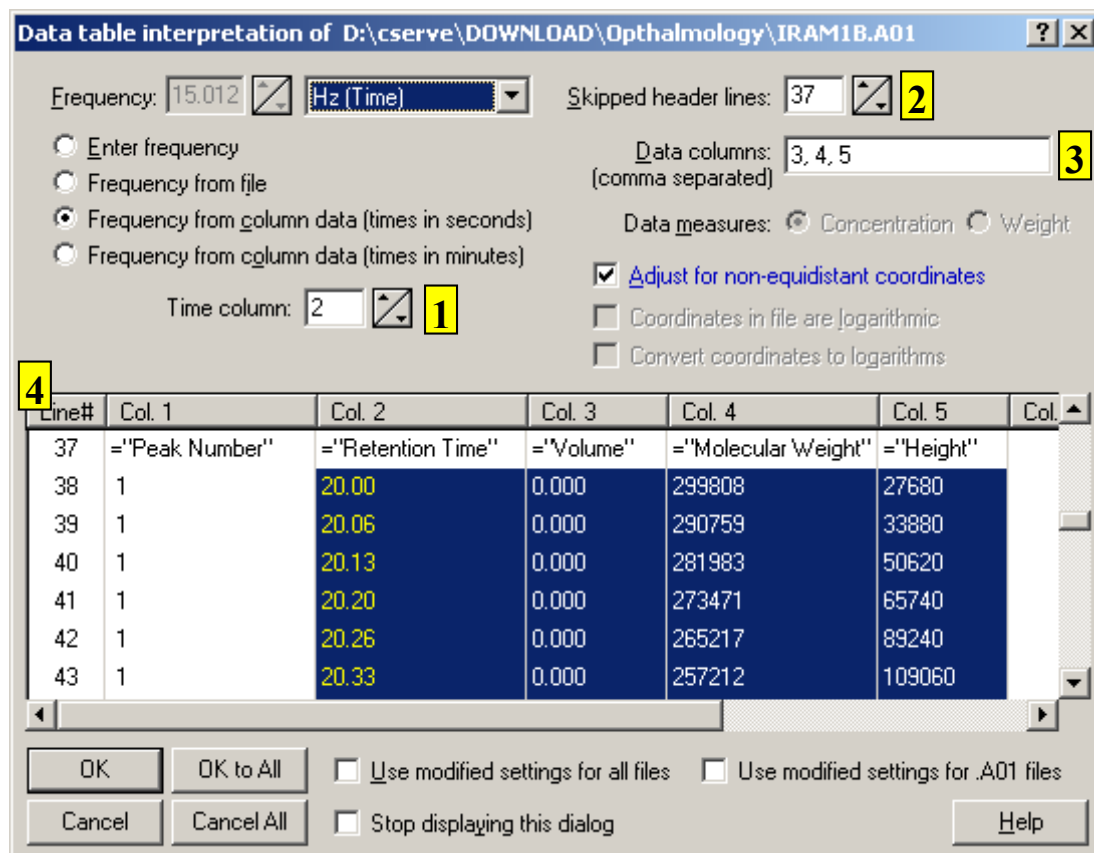
= "Sectional Average Molecular Weight"
= "Area Number" = "Mn" = "Mw" = "Mz" = "Mz1"
```

To read this file in DAX use the **File | Open** menu option. This opens the File | Open dialog box. Under **list files of type**, select **Generic ASCII** files.



Click the **Open** button to import the IRAM1B.A01 file. DAX will display a dialog box to start the interpretation process.

Initially, the dialog box looks like this:



As can be seen, DAX has

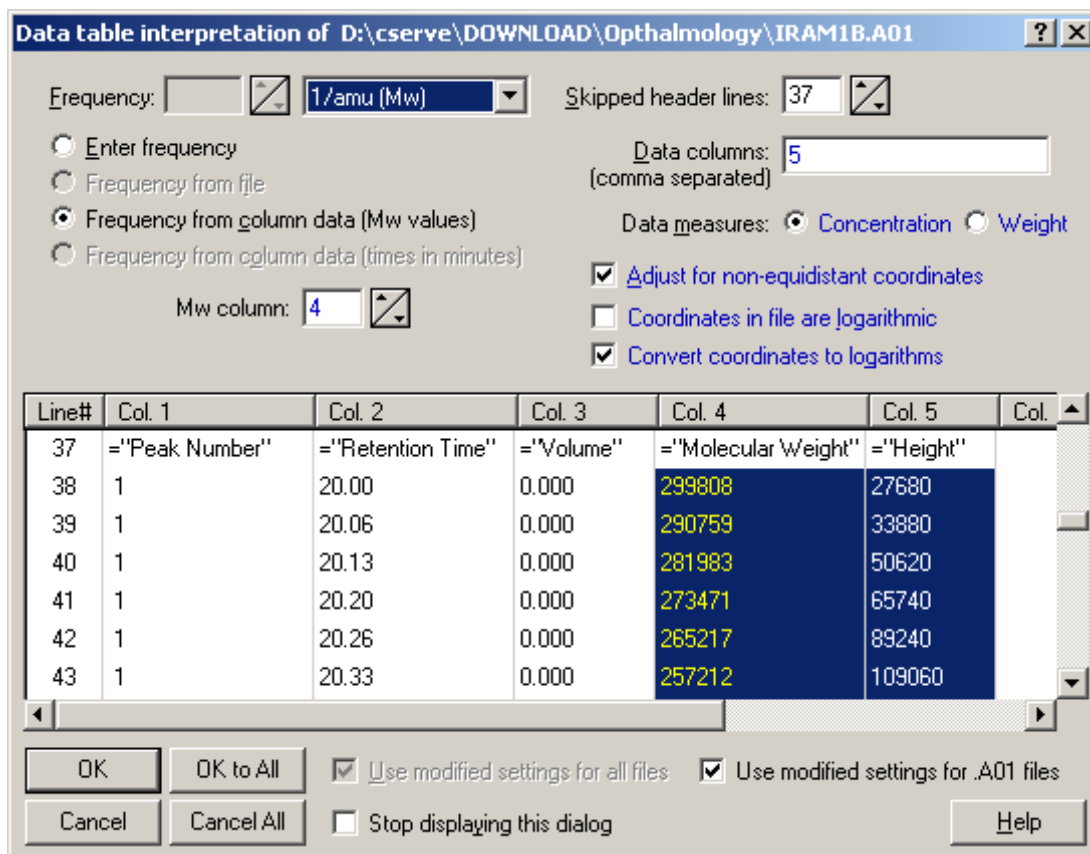
- correctly determined that the second column contains retention time data (1,4)
- correctly determined that the first 37 lines do not contain point data (2)

DAX assumes that the columns following the time coordinates, in this case columns 3, 4 and 5, contain point data (3).

To correctly read the file IRAM1B.A01, A number of adjustments need to be made.

- change the horizontal axis type to **1/amu (Mw)**
- indicate column 4 contains the Mw data
- indicate only column 5 contains point data
- (optional) **convert coordinates to logarithms** causes the Mw data to be converted to logarithmic values, which will create a nicer looking graph

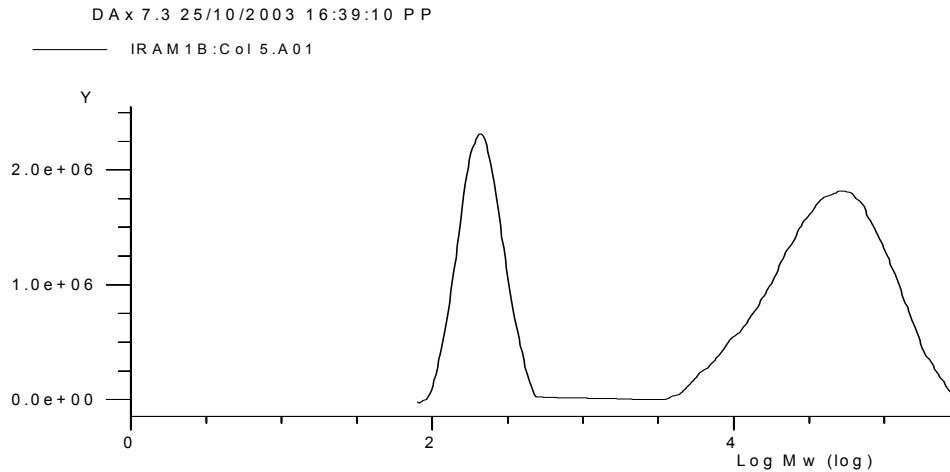
The dialog box ends up looking like this:



Note that values that have been modified are now coloured blue, to indicate that these are **user values**¹. Note that **Use modified settings for .A01 files** has been checked. More on this later.

¹ Some items are always blue, because they are always user values, since DAX cannot estimate them from the contents of the file.

Click the **OK** button to load the IRAM1B.A01 file.



Interpreting other .A01 Files

In the text interpretation dialog box, the **Use modified settings for .A01 files** was checked in the example above. When we now try to open the file IRAM7.A01, the text interpretation dialog box looks like this:

Data table interpretation of D:\cserve\DOWNLOAD\Ophthalmology\IRAM7.A01

Frequency: 1/amu (Mw) Skipped header lines: 52

Enter frequency
 Frequency from file
 Frequency from column data (Mw values)
 Frequency from column data (times in minutes)

Mw column: 4

Data columns: 5 (comma separated)
 Data measures: Concentration Weight

Adjust for non-equidistant coordinates
 Coordinates in file are logarithmic
 Convert coordinates to logarithms

Line#	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
50	1	.	80	0.000	2150999023	379460
51	1	.	86	0.000	2085951694	354760
52	1	.	93	0.000	2022871431	333500
53	1	1.00	0.000	1961698747	319580	
54	1	1.06	0.000	1902375957	304340	
55	1	1.13	0.000	1844847120	287120	
56	1	1.20	0.000	1789057984	293320	

OK OK to All Use modified settings for all files Use modified settings for .A01 files
 Cancel Cancel All Stop displaying this dialog Help

Note that

- DAX has made **skipped header lines** 52, rather than 37. This value was not modified when reading IRAM1B.A01, so DAX was able to adjust it. While lines 50 !! 52 do contain columns, they do not contain valid data, so they were skipped
- DAX is reading Mw data from column 4, and point values from column 5; these settings were kept from IRAM1B.A01

Clearly, no further adjustments are needed to interpret this file.

Automatically Interpreting Text File Formats

The procedure outlined above can be used for up to ten different text file formats, provided the files have different file name extensions. Use the following steps to set up interpretation of a text file format.

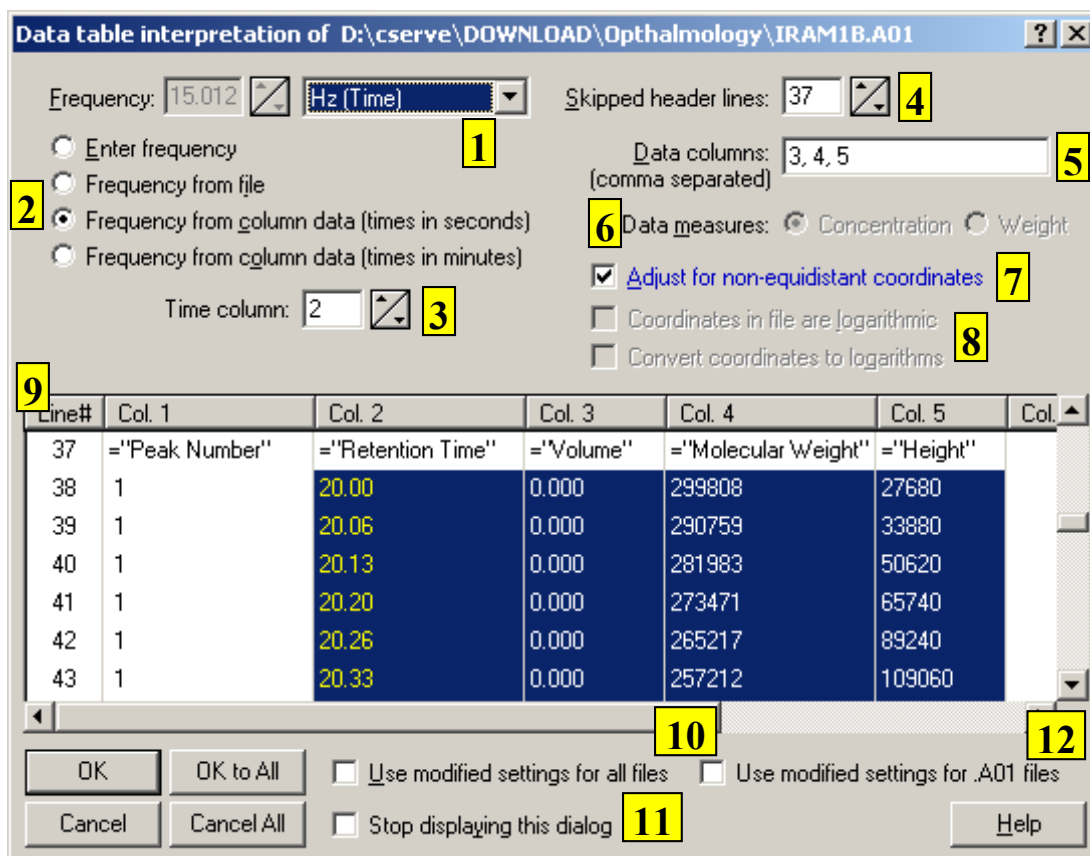
- 1 Open an example file
- 2 Modify settings for interpretation as appropriate in the text file interpretation dialog box; modified settings will be displayed in blue
- 3 Check the **Use modified settings for .XXX files** box
- 4 Click the **OK** button

The next time a text file with a known file name extension is opened, the modified settings will be used.

This behaviour extends to opening files from the command line (e.g. **Dax32 iram7.a01**) or when scanning for files. In these cases the automatic interpretation is particularly important, because no text file interpretation dialog box will be displayed.

To stop automatic interpretation of a certain text file type, un-check the **Use modified settings for .XXX files** box whenever a text file is opened.

A Detailed look at the Text Import dialog box



1. The horizontal axis type to be used with the data can be selected (from time, index, Mw, or BP). If a **point index** axis is used, no frequency is used, since data points are plotted versus their index in the data array.
2. There are several ways in which DAx can determine the measurement frequency.
 - If a frequency is mentioned in the file, identified by the words "frequency" or "rate", DAx will initially use that frequency.
 - If no frequency is mentioned, but there is more than one column of data, DAx will assume that the first or a later column contains time coordinates. The frequency is determined using the time coordinates of the very first and very last points found in the file. For time axes, the user can indicate if time coordinates are in seconds or in minutes.
 - The user may define a frequency.
3. If there is more than one column of numbers, DAx will assume that the first column contains time coordinates. DAx will try to verify this. For instance, if the values in the first column do not increase, DAx will not use this column for time coordinates. Sometimes a file will contain a first column of index numbers and a second column with time coordinates. DAx tries to recognise this situation. If it does, the second column will be indicated as the time coordinates column.
4. The top part of the data is likely to contain a number of header lines. DAx tries to establish how many header lines there are. You may adjust it to a higher value. For instance, if the file contains only a single column of numbers, without any text, DAx will assume all numbers are signal values. If the first number is a frequency value instead, the top line should be skipped.
5. The first column that is not an index or time coordinates column will by default be the data column. It is possible to list more than one data column by entering column coordinates separated by commas, or by clicking the column headers. **Note** that data that are read from different columns are assigned different "trace types", called **Col.1**, **Col.2**, etc. These trace types can be used to limit certain operations, such as recognising peaks, to data from certain columns.
6. Molecular weight data can measure concentration or weight.
7. Check **Adjust for non-equidistant time coordinates** to interpret a text file containing non-equidistant time coordinates (i.e the time difference between consecutive points is not always the same). If you do not check this, and the time coordinates are not equidistant, the resultant data will be distorted. Unchecking this option for equidistant text files slightly speeds up operation.

8. Check **Coordinates in file are logarithmic** if the file contains a column with logarithmic horizontal coordinates (typically Log Mw values). DAX will show the logarithmic coordinates on the horizontal axis of the imported data, but will correctly convert the logarithmic values when calculating peak data (such as Mn, Mz).
Check **Convert coordinates to logarithms** if the file contains non-logarithmic horizontal coordinates, but you want to display logarithmic coordinates on the horizontal axis of the imported data. This is typically used when importing Mw data that span multiple orders of magnitude.
9. A sample of the contents of the text file is displayed here.
10. Check the **Use modified settings for all files** to apply *changed* settings to future files. Only the settings that you modify in the dialog box will be kept; optimum parameters will be derived for all other settings. The settings you modify will be displayed in blue.
This option is particularly useful when you use **OK to All** or check the **Stop displaying this dialog** box. The next time DAX starts, the settings will have been reset.
11. Check the **Stop displaying this dialog** box to stop using the text interpretation dialog box. You may check this box if the text file being read is correctly interpreted by DAX without any user input (and if you know you will only be reading similar text files for the rest of the time you are using DAX) or if you have checked **Use modified settings for all files**.
The next time DAX starts, the text interpretation dialog box will again be displayed.
12. Check the **Use modified settings for .XXX files** to apply *changed* settings to future files with the file name extension **.XXX**. Only the settings that you modify in the dialog box will be kept; optimum parameters will be derived for all other settings. The settings you modify will be displayed in blue. These settings are *kept* between sessions of DAX.