AUTOMATED XY PLOTS FROM MICROSOFT EXCEL

DAVID M. CHRISTIE¹ and CHARLES H. LANGMUIR²

¹College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331-5503 and
²Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, U.S.A.

(Received 16 June 1993; revised 27 August 1993)

Abstract—We present a macro program which greatly simplifies the process of making X–Y (scatter) and
triangular plots using the spreadsheet program Microsoft Excel®. The program draws on user-defined
"type-plots" to establish plot format, allowing a great amount of flexibility in style and presentation.
Predefined templates can be used to define fields or lines for specific purposes and these can be retained
easily from plot to plot facilitating rapid comparison of different data sets. The program requires only
the addition of some simple control characters to be able to draw data from existing spreadsheets with
little or no reformatting.

Key Words: Spreadsheets, Microsoft Excel, Plotting, Graphing, X–Y plots, Scatter plots, Triangular plots.

INTRODUCTION

The spreadsheet program Microsoft Excel® is used widely for data storage and manipulation. The pro-
gram also is capable of producing high-quality x–y plots ("scatter plots" in Excel parlance) in a variety
of formats, but the obscurity of the process is such that many users prefer to export data to other
programs for plotting purposes. We have used the macro functions of Excel to automate substantially
the plotting process, with few restrictions on data format and with considerable flexibility of output
format. The output format is taken from, and exactly reproduces, the format of a "type plot" set up by the
user. The macro should be useful even to novices, once they have mastered the rudiments of Excel
jargon. We have tested the macro extensively on Macintosh computers using versions 2.2a and 3.0 of
Microsoft Excel. Recent tests with Excel version 4.0 and Macintosh System 7.0.1 have encountered no
difficulty. The macro should be transferable to the IBM Windows environment, but we have not tested
this feature.

DESCRIPTION OF THE MACRO

The text of the macro "Multiplot" is given in the Appendix. It should be stored within an Excel macro
sheet, defined as a "command macro", named and assigned a command-option-letter combination
using the "Formula Define Name" command. If the macro sheet is renamed "Excel Startup" and placed
in the "Preferences" folder (System 7 only) of the System folder, it will open automatically whenever
Excel is started.

In writing the macro, we have used a three-column format. The middle column contains the actual
macro code. The left column carries variable names; each variable name refers to the cell immediately
to the right, allowing the program to be followed easily. The variable names are assigned to the rel-
cent cells using the "Formula Create Names" command, enabling their use as actual variables in later
macro functions and as cell addresses for "GOTO" statements. Finally, the right-hand column contains
comments.

HOW TO USE THE MACRO

"Multiplot" takes "x" and "y" values from designated vertical columns of an Excel Worksheet, with
columns designated by numbers, not letters ("R1C1" format). Numbers or text can be taken from a third
column and written onto the plot next to each data point. Names of the "x" and "y" variables can be
taken from any horizontal row within the worksheet, usually from the column headings. Plots are named
by concatenation of the name of the worksheet with the names of the "x" and "y" variables.

To use the macro with an existing worksheet, it is necessary to insert two blank columns at the left edge
of the sheet. These are used to group the data and to name the groups. If data are arranged in rows, a new
sheet can be created easily using the transpose function under the "Edit Paste Special" command.

Data may be grouped into series of two or more rows. The first and last line of each series is design-
ated by placing an integer in column 1 of the first

CAGEO 20(1-D) 47
must be numbered consecutively, starting with "1", but they need not be in order down the worksheet (e.g. series "1" may be below series "2"). Series may be nested or overlapping.

**Plot formats**

A wide range of plot sizes and formats are available through the "Format" and "File Page Setup" functions of Excel. Figures 2 and 3 are examples. Because some aspects of plot size and format can be controlled directly by macro functions, Multiplot begins by opening a preexisting "type plot" set up by the user. In the Appendix, it can be seen at line 46 that the plot "xy" is called from the folder (directory) "TypePlots" within the folder "Excel" on drive "80 MB HD". This sequence must be modified to match each individual computer. An error subroutine (TP.Addresser—see Appendix) is provided to do this automatically.

**Obtaining plots of uniform size**

A major deficiency of Excel is its failure to provide a mechanism for designating relative or absolute axis lengths. Nevertheless, with care (and

---

**Figure 1.** Elements of typical worksheet include: (i) data arranged horizontally for each sample, (ii) column headings, in single row (row 3 in this example), (iii) series delimiters in column 1 at first and last row of each series, and (iv) series names in column 2 at first row of each series.

**Figure 2.** Typical xy plot using data from Figure 1. Labels are sample numbers from column 3.
frustration) a series of plots of identical size can be produced.

The final printed size of an Excel plot is determined by numerous features including the shape of the plot on screen, a variety of settings in “Page Setup”, the size and font of labels on “x” and “y” axes and the lengths of series names in the explanation. All these variables must be kept constant if plots of identical size are to be achieved. If the explanation is beside the plot, its size is determined by the longest series name it contains. Titles can be made constant in length using spaces or periods, if a nonproportional font is used. Explanations beneath the plot generally are less troublesome. Numerical labels on axes obtain their format from the original data sheet, so a uniform format usually can be achieved using the “Format Number” command to temporarily modify the format of the data. No precision is lost by this procedure. The size of standard text labels on the axes of Excel plots is difficult to control. A simple solution is to add axis labels as “unattached text” which can be located at any convenient position within the plot. Multiplot assigns the y-axis name to “Text 1” and the x-axis to “Text 2”. These labels may be moved around the plot by the user. This approach has the added advantage of greatly enlarging the area of the actual plot on the screen.

**Triangular diagrams and other special templates**

Triangular diagrams can be made easily as special “xy” plots in Excel. For a triangle labeled $ABC$, clockwise from the apex, with components $A$, $B$, and $C$ summing to 100, the “y” coordinate is equal to $A$ and “x” equals $B + A/2$. (This may seem counter-intuitive, but it works because the units on the “x” and “y” axes are of unequal length; see Fig. 3.) The triangle outline can be incorporated “permanently” into a new “type plot” as the first data series. Using the “Format Patterns” command, points may be made invisible and connecting lines drawn, in a variety of formats. A wide variety of permanent templates that define areas, trend lines, or other reference markers, may be created for specific purposes in this manner, using as many series as necessary to create the required lines.

Multiplot can be used in conjunction with special templates in the following way, which is described in reference to Figure 3.

A type plot is created and renamed “xy3” for the duration of the plotting exercise. Essential attributes of this type plot are (1) that the triangle is plotted as the first series, and (2) that the coticect (gray, shaded lines) is plotted as the second series. Data for both series can be stored permanently in a separate worksheet.

When Multiplot is run, the variable “Plot-Type” is designated by the user as number “3” (see Appendix, Line 8). This causes the type plot “xy3” (Appendix, Line 49) to be opened instead of “xy” as described previously, and the first series of variable data becomes series 3. The macro is written specifically for values up to 3, but higher values can be used provided an appropriate type plot is provided and named “xy3”.

![Figure 3. Typical triangular plot. Triangle (series 1) and coticect (gray lines; series 2) are stored as first two series. First data series is series 3.](image-url)
Controlling symbol size

On printed plots, the final size of symbols and lettering can be controlled to some extent by selecting “fit to page” and specifying a percentage value for reduction (or enlargement) as part of the “Page Setup” process. For plots with many data points, reductions of 50% or more produce satisfactory results. For plots intended for reduction prior to publication, enlargements up to 200% can be used to produce simple plots of publishable quality. Use of this feature slightly affects overall plot size, so uniform reductions should be used if identical plots are desired.

Running the macro

With the worksheet from which the data are to be plotted as the active (uppermost) window, the macro can be initiated using the “Macro Run” command or, more conveniently, by typing the assigned command–option-letter key sequence. The user then must respond to requests for several items of information:

1. the number of series to be plotted,
2. a variable integer termed “PlotType” corresponding to the number of the first data series to be plotted (usually 1, but higher for special plots: e.g. 3 for the triangular plot as described),
3. the row number in which column headings are written,
4. column number for x axis,
5. minimum and maximum column numbers to be plotted as x-axes on successive plots, and
6. column number for labels if points are to be labeled individually.

For most of these requests, a default value is offered.

Execution of the macro terminates with a request for a new x-axis value. Plotting can be continued indefinitely or terminated by typing a zero at this point. All plots are saved in the folder from which the data originated.

Common problems

We have encountered five common problems in working with Multiplot:

1. If a series delimiting number is absent from column 1, the program will search all 16,384 rows. All macros can be halted by typing “Command-Period”.
2. Text for axis labels, column headings, and file names cannot contain spaces or certain punctuation and symbol characters. This seems to be a bug in the Excel macro programming as spaces are not normally a problem.
3. Long filenames may cause errors. There seems to be a limit to the length of text string which can be handled. We have not explored the exact value of this limit.
4. Blank column headings terminate the macro.
5. An attempt to generate a plot with the same name as an existing open window will cause an error. User must close or delete spurious windows generated by run failures before trying again.

Debugging

If the first line of Multiplot is replaced by the macro function “=STEP(“), the macro will execute one line at a time, allowing progress to be monitored, and the point at which an error occurs to be identified. The step function can be inserted at any point in any macro.

CONCLUSION

The macro “Multiplot” provides a rapid, versatile plotting capability for Microsoft Excel Spreadsheets. Copies of the macro in Macintosh format will be provided by the authors on receipt of a formatted 3.5" diskette or via FTP on request to dchristie@oce.orst.edu (Internet).

Acknowledgments—The authors acknowledge support from the National Science Foundation, including OCE 90-00595, OCE 89-11826 to Christie and OCE 90-14298 to Langmuir. We thank two anonymous reviewers for comments.
APPENDIX

Text of Macro "Multiplot" and Associated Subroutines

© Copyright 1993 by
Charles H. Langmuir and
David M. Christie

A
B
C

1 MULTI PLOT Command Option X

2 NAME

3 =ECHO(FALSE)

4 INDEX(WINDOWS(1,1,1)

5 NSERIES

6 =INPUT("How many series are there?",1, "Series?",1)

7 =IF(NSERIES=FALSE,HALT(),

8 ploctype

9 =INPUT("Plot type (for templates)?",1,1)

10 axisname

11 =INPUT("In which row are the axis names?",1,3)

12 x

13 =INPUT("Column number of x-axis",1)

14 y

15 =INPUT("Input min column number for y axis",1)

16 ymax

17 =INPUT("Input max column number for y axis",1)

18 ybiclu

19 =INPUT("Input column number for labels. 0 if none",1,0)

20 Asstart

21 =IF(x=x,y,GOTO(Afinish))

22 folder

23 =GET.DOCUMENT(2,"&NAME&")

24 XNAME

25 =SELECT("R"&axisname&"C"&x&")

26 YNAME

27 =ACTIVE.CELL()

28 =SELECT("R"&axisname&"C"&yc&")

29 endf1

30 =IF(ybiclu=0,GOTO(endf1))

31 RNAME

32 =SELECT("R"&axisname&"C"&ybiclu&")

33 lbyname

34 =ACTIVE.CELL()

35 count

36 =SET.NAME("count",1)

37 ibname

38 =ACTIVE.CELL()

39 startseries

40 =SET.NAME("plotseries",iplottype)

41 startpoint

42 =SELECT("R"1")

43 getpoint

44 =SELECT("R"1")

45 Rstartseries

46 =ROW(ACTIVE.CELL())

47 setseries

48 =SELECT("RC1")

49 Rendseries

50 =ROW(ACTIVE.CELL())

51 Rrendseries

52 =GET seriesseries+counter,GOTO(Rstartseries)

53 count

54 =GET seriesseries+counter,GOTO(Rstartseries)

55 Ibname

56 =GET seriesseries+counter,GOTO(Rstartseries)

57 Directory

58 =DIRECTORY("(80 MB HD:Excel Folder :TypePlots")

59 =ERROR(2,TP.Addresser)

60 Directory

61 =DIRECTORY("(80 MB HD:Excel Folder :TypePlots")

62 =ERROR(2,TP.Addresser)

63 =FORMULA("=SERIES("&seriesname&","&NAME&","IR"

64 &R&startseries&"C"&x&","R&endseries&"C"&x&","&NAME&","IR"

65 &R&startseries&"C"&y&","R&endseries&"C"&y&","plotseries&")")

66 =ERROR(1)

67 =DIRECTORY(folder)

68 =SAVEAS("&NAME&","&XNAME&","&NAME&","1",FALSE)

69 label

70 =IF(ybiclu=0,GOTO(nobill))

71 =SET.NAME("plotpoint",1)

72 label

73 =SET.NAME("points",Rendseries-Rstartseries+1)

NOTE: Errors at this point are usually due to spaces or characters in text.

Returns to normal error response

Back to original folder

Giving the chart a name

LABEL ROUTINE:

Set counter

Determines number of points in series

Continued overleaf
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>62</td>
<td>WHILE(pointno&lt;=points)</td>
<td>Label loop</td>
</tr>
<tr>
<td>63</td>
<td>lbitxt =GET CELL(5,TEXTREF NAME=&quot;IR&quot; &amp; Rstartseries &amp; pointno-1 &amp; &quot;C&quot; &amp; ilbitxt, FALSE))</td>
<td>Get label text</td>
</tr>
<tr>
<td>64</td>
<td>ERROR(0)</td>
<td>Allows plot to continue if there is an error</td>
</tr>
<tr>
<td>65</td>
<td>ATTACH TEXT(4, plotseries, pointno)</td>
<td>Attach to plot</td>
</tr>
<tr>
<td>66</td>
<td>ERROR(1)</td>
<td>Returns to normal error response</td>
</tr>
<tr>
<td>67</td>
<td>FORMULA(&quot;** &amp; lbitxt &amp; **&quot;)</td>
<td>Attaches label to plot</td>
</tr>
<tr>
<td>68</td>
<td>SET NAME (&quot;pointno&quot;, pointno+1)</td>
<td>Goes to next point</td>
</tr>
<tr>
<td>69</td>
<td>NEXT()</td>
<td>End loop. End label.</td>
</tr>
<tr>
<td>70</td>
<td>FORMAT FONT(0,1, TRUE,&quot;Geneva&quot;, 9, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE)</td>
<td>Sets font</td>
</tr>
<tr>
<td>71</td>
<td>NOBIT</td>
<td>Are there more series?</td>
</tr>
<tr>
<td>72</td>
<td>NEW SERIES =SET NAME (&quot;counter&quot;, counter+1)</td>
<td>This tracks the series number</td>
</tr>
<tr>
<td>73</td>
<td>&amp; SET NAME (&quot;plotseries&quot;, plotseries+1)</td>
<td>This tracks the series number as plotted.</td>
</tr>
<tr>
<td>74</td>
<td>&amp; ACTIVATE (&quot;** NAME**&quot;)</td>
<td>Activate data sheet</td>
</tr>
<tr>
<td>75</td>
<td>&amp; SET NAME (&quot;switch&quot;, 1)</td>
<td>Switch for second or higher series</td>
</tr>
<tr>
<td>76</td>
<td>&amp; GOTO (startseries)</td>
<td>Go get the delimiters for the next series</td>
</tr>
<tr>
<td>77</td>
<td>NEW SERIES =ACTIVATE (&quot;** NAME** &amp; NAME** &amp; NAME** &amp; NAME**&quot;)</td>
<td>Activate plot</td>
</tr>
<tr>
<td>78</td>
<td>SELECT (&quot;PLOT&quot;)</td>
<td>Select plot</td>
</tr>
<tr>
<td>79</td>
<td>&amp; FORMULA (&quot;** SERIES** &amp; seriesname** &amp; NAME**&quot;)</td>
<td>Add second or higher series to plot</td>
</tr>
<tr>
<td>80</td>
<td>&amp; Rstartseries &amp; C &amp; x &amp; R &amp; Rendseries &amp; C &amp; y &amp; &amp; NAME** &amp; IR</td>
<td>NOTE: Errors at this point are usually due</td>
</tr>
<tr>
<td>81</td>
<td>&amp; Rstartseries &amp; C &amp; y &amp; R &amp; Rendseries &amp; C &amp; y &amp; &amp; plotseries** &amp;&quot;)</td>
<td>to spaces or characters in text</td>
</tr>
<tr>
<td>82</td>
<td>&amp; GOTO (label)</td>
<td>Go to label routine</td>
</tr>
<tr>
<td>83</td>
<td>&amp; IF(plottype=1, GOTO (end))</td>
<td>More series? Or are we done.</td>
</tr>
<tr>
<td>84</td>
<td>&amp; SELECT (&quot;Text 2&quot;)</td>
<td>Select x-axis label</td>
</tr>
<tr>
<td>85</td>
<td>&amp; FORMULA (&quot;** &amp; NAME**&quot;)</td>
<td>Write name of x-axis on plot</td>
</tr>
<tr>
<td>86</td>
<td>&amp; SELECT (&quot;Text 1&quot;) &amp; FORMULA (&quot;** &amp; NAME**&quot;)</td>
<td>Select y-axis label</td>
</tr>
<tr>
<td>87</td>
<td>&amp; IF (lbitxt=0, GOTO (end2)) &amp; IF (lbitxt=0, GOTO (end2))</td>
<td>Write name of y-axis on plot</td>
</tr>
<tr>
<td>88</td>
<td>&amp; ERROR(0) &amp; ERROR(0)</td>
<td>Error check off in case Title or Text3 missing</td>
</tr>
<tr>
<td>89</td>
<td>&amp; SELECT (&quot;Title&quot;) &amp; SELECT (&quot;Title&quot;)</td>
<td>Select attached plot title if present</td>
</tr>
<tr>
<td>90</td>
<td>&amp; FORMULA (&quot;** &amp; NAME**&quot;) &amp; SELECT (&quot;Text 3&quot;)</td>
<td>Write filename as plot title</td>
</tr>
<tr>
<td>91</td>
<td>&amp; SELECT (&quot;Text 3&quot;) &amp; FORMULA (&quot;** &amp; NAME**&quot;)</td>
<td>Write name of label on plot</td>
</tr>
<tr>
<td>92</td>
<td>&amp; ERROR(1)</td>
<td>Normal error response</td>
</tr>
<tr>
<td>93</td>
<td>&amp; IF(SERIES=1, LEGEND, FALSE)</td>
<td>Remove legend for single series</td>
</tr>
<tr>
<td>94</td>
<td>&amp; MOVE ((y-ymin)/5, (y-ymin)/5)</td>
<td>Stacks new window</td>
</tr>
<tr>
<td>95</td>
<td>&amp; END12 &amp; SAVE12</td>
<td>Save the final chart</td>
</tr>
<tr>
<td>96</td>
<td>&amp; IF (y=max, GOTO (newchart), GOTO (AStart))</td>
<td>On to the next y-axis for the same x-axis</td>
</tr>
<tr>
<td>97</td>
<td>&amp; AFINISH &amp; SET.VALUE (yy+1)</td>
<td>Make another one?</td>
</tr>
<tr>
<td>98</td>
<td>&amp; INPUT(&quot;For another x-axis value type 1, 1, 0&quot;) &amp; IF (newchart=1, GOTO (x),)</td>
<td>Finish</td>
</tr>
<tr>
<td>99</td>
<td>&amp; IF (newchart=1, GOTO (x),)</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>&amp; RETURN()</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>*** Asterisks lines are continued from the cell above, solely for</td>
<td>NOTE: Each macro must be named</td>
</tr>
<tr>
<td>102</td>
<td>legibility. To function correctly, the entire formula should be</td>
<td>(Define Name) as shown in bold</td>
</tr>
<tr>
<td>103</td>
<td>&amp; RETURN()</td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>SUBROUTINES:</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>TP ADDRESSER</td>
<td>Changes Multiplot to find TypePlots at new address</td>
</tr>
<tr>
<td>109</td>
<td>&amp; ALERT (&quot;Can't find TypePlots. Please Open the plot 'xy', 1)</td>
<td>Message to screen</td>
</tr>
<tr>
<td>110</td>
<td>&amp; OPEN (Y,0)</td>
<td>Produces the 'Open' dialog box</td>
</tr>
<tr>
<td>111</td>
<td>&amp; GET DOCUMENT (2, &quot;xy&quot;)</td>
<td>Stores address of 'xy'</td>
</tr>
<tr>
<td>112</td>
<td>&amp; CLOSE</td>
<td>Subroutine to call in case of error</td>
</tr>
<tr>
<td>113</td>
<td>&amp; ERROR(2, TP ADDRESSER)</td>
<td>Normal error response</td>
</tr>
<tr>
<td>114</td>
<td>&amp; ACTIVATE (&quot;MULTIPLATFORM&quot;) &amp; IF (newchart=1, GOTO (x),)</td>
<td>Amends address for TypePlots in Multiplot</td>
</tr>
<tr>
<td>115</td>
<td>&amp; ALERT (&quot;Address now corrected on line labeled Directory. Try again&quot;, 1)</td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>&amp; ERROR(1)</td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>&amp; SELECT (&quot;Directory&quot;)</td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>&amp; FORMULA (&quot;MULTIPLATFORM&quot; &amp; ENTFOLDER &amp; &quot;&quot;)</td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>&amp; RETURN()</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>TEXTERROR</td>
<td>TextError</td>
</tr>
<tr>
<td>121</td>
<td>&amp; ALERT (&quot;TextError, Remove spaces,symbols from sheet</td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>or column names&quot;, 3)</td>
<td></td>
</tr>
<tr>
<td>123</td>
<td>&amp; ALERT (&quot;Be sure to close (not save) any odd plots that have appeared&quot;, 3)</td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>&amp; HALT</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>&amp; ALERT (&quot;Sorry, Macro is only self correcting if it is named Multiplot Rename and try again&quot;, 3)</td>
<td></td>
</tr>
<tr>
<td>126</td>
<td>&amp; HALT</td>
<td></td>
</tr>
</tbody>
</table>