Formatting Plots in a Research Report

D. Keffer
Department of Chemical Engineering
University of Tennessee, Knoxville
March 30, 1999

This document provides some pointers on how to format data plots in a technical research report.

Purpose:

The purpose of a data plot is to communicate to the reader the physical consequences of your data. Therefore, there are two ultimate goals of plot formatting. The first goal of plot formatting is to display your data HONESTLY (or objectively). Clever (and deceitful) or just plain careless manipulation of a plot can make data appear to support one hypothesis, when actually it supports another hypothesis. The second goal of plot formatting is to display your data with as much CLARITY as possible.

Guidelines:

I. Titles, axes, captions, legends, and text boxes

- The plot should have a simple, descriptive title.
- The plot should have a simple, descriptive caption, which may contain additional information not given in the plot.
- The plot should have x and y-axes titles. If the x and y-axes have units, the units should be given.
- The range of the x and y axes should correspond to the data. There should not be large blank spaces in the plot.
- The x and y axes should have numerical values. The values should have only the minimum number of significant figures needed to convey to the reader the vicinity of the plot. For example, the numbers on the x-axis should be "0.1" and not "0.100".
- When displaying multiple data sets on the same figure, you need a legend. The legend should be short and sweet. In Figure One, we use the legend to distinguish between experimental and theoretical data.
- If, as in the case, of Figure One, there are some important parameters that are held constant for that plot, they should be stated in a blank space on the plot, if possible, and in the caption of the plot.

II. Plot content

- You should try to display as much related information on a single figure as possible, without compromising the clarity of the plot. Plotting data on the same figure, allows for ease of comparison. In Figure One, we plot both theory and experimental lines on the same plot. We could have plotted them on separate plots but that would have made it more difficult to compare the two. The two sets of data are still clearly displayed in Figure One and Two.
- In Figures Three and Four, six sets of data are clearly displayed. This allows us to see the coupled effect of the initial height and the pipe diameter. This is the best way to display the data, so long as the six data sets do not overlap to such a great extent that the plot becomes an unreadable jumble of lines.

III. Lines, curves, and markers ()

- Your plots should be honest. This means you should only display the data that you obtained. You have not measured how the system behaves between your data points. Therefore, you should not pretend that you know how the system behaves there. The only thing you know is what you measured.
- When displaying experimental data, you should use POINTS where you have a data points. You can connect the points with STRAIGHT lines, if you choose, for the sole purpose of guiding the reader's

eye for clarity in the plot. You should never use a polynomial curve fit because you do not know how the experimental system behaves between your data points.

- If you want to show that data fits a certain model (e.g. linear or quadratic or exponential), you need to indicate that the line is a linear or quadratic or exponential best fit, in the legend. The coefficients of best fit in the model, should be given in the caption or the text of the report.
- When displaying an analytical theory, as is the case for theory 1 in Figures One and Three, you should use a smooth line without points. The analytical theory has been evaluated at many points. STRAIGHT lines between the points appear to the eye as curved. This is what has been done in Figures One and Three.
- Sometimes analytical theories require too much computational effort to evaluate at many data points. In this case the smooth theoretical curve in Figures One and Three cannot be obtained. In this case, we evaluate the theory at a few points. We then have "theory data points", which we must treat in precisely the same manner as "experimental data points". This means we use a marker at the data point and we connect with straight lines if we want to guide the reader's eye. This is what has been done in Figures Two and Four.
- In none of these plots did we connect the experimental points with a straight line. We could have, but we didn't need to because our eye was guided by the theoretical curve.
- In none of these plots did we use "curved" lines.

IV. Lines, curves, and markers (style)

- When displaying multiple data sets on the same figure, you should use different lines and points for the data. This is done in all four plots.
- In Figure Three, notice we used the same line for all three theoretical curves. We could have used a solid line, a dashed line, and a dotted line for the three theoretical curves. It is our choice. I did not use them because I believe the presence of the experimental data clearly distinguishes the theoretical data.
- In Figure Four, we have six sets of data. There are two variables in the data, theory vs. experiment and the initial water height. Notice how I have used open markers for all the experimental sets and filled markers with lines for the theoretical data sets. Note also that I have used the same markers for the same initial water heights. This tells the reader to compare open squares with filled squares, compare open triangles with filled triangles, compare open circles with filled circles, if they want to compare theory and experiment at the same value of initial water height. This also tells the reader to compare only open markers, if they want to look at experimental results as a function of water height, or only filled markers, if they want to look at theoretical results as a function of water height.
- In Figure Four, I distinguished between theory sets by using different filled markers but the same style (solid) of line. There is no need to distinguish between data sets by changing both markers and lines.

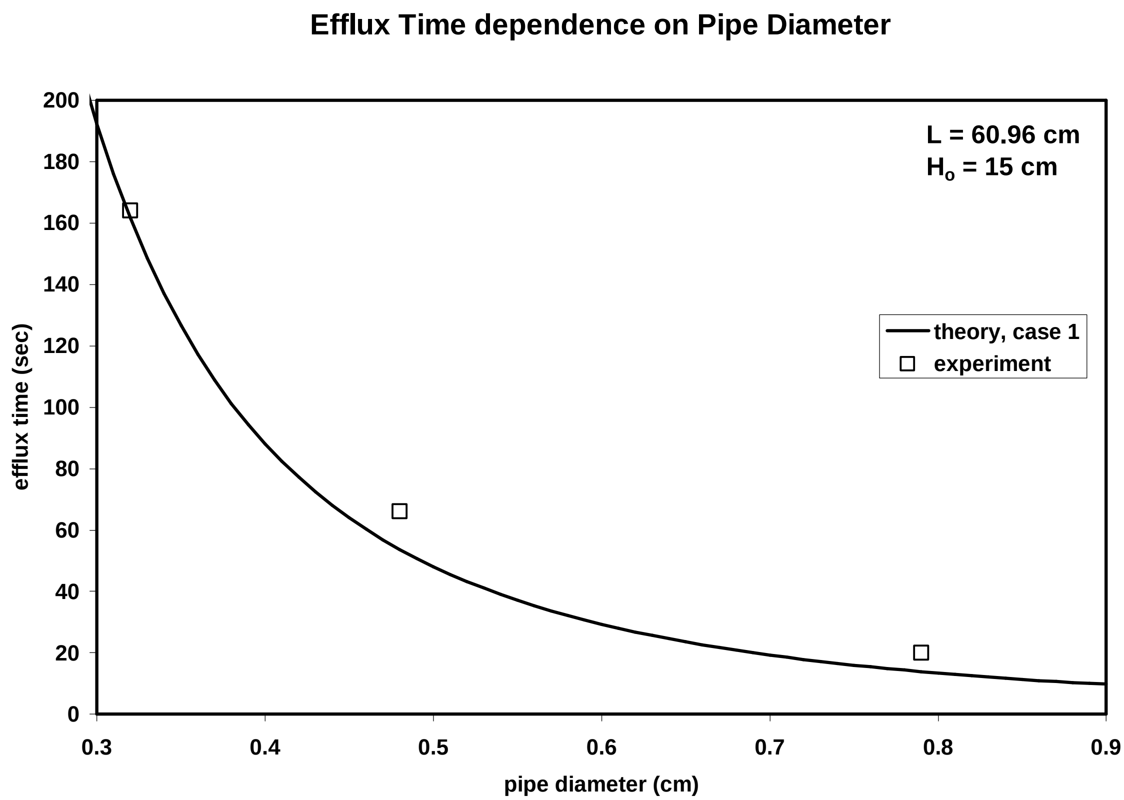

Figure One. Efflux time as a function of pipe diameter, using the pipe of Length 60.96 cm and an initial water height of 15 cm. The theory line represent Case One, equation (8).

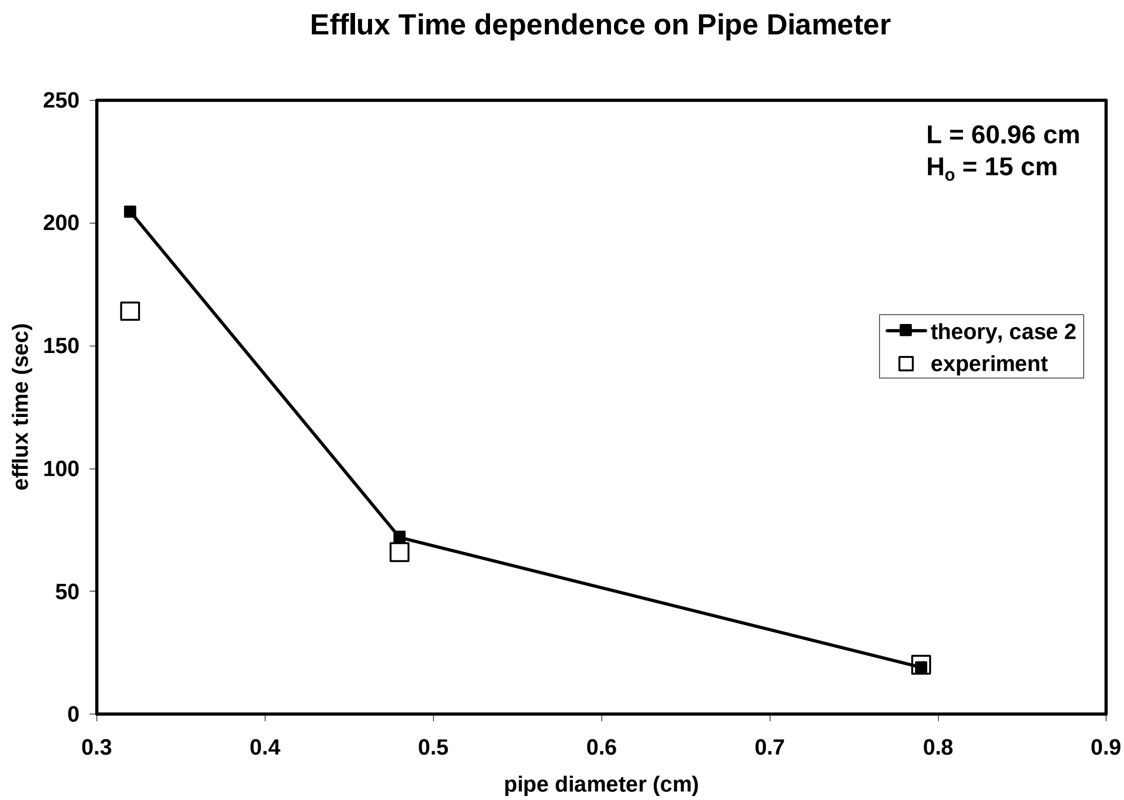

Figure Two. Efflux time as a function of pipe diameter, using the pipe of Length 60.96 cm and an initial water height of 15 cm. The theory line represent Case Two, equation (16).

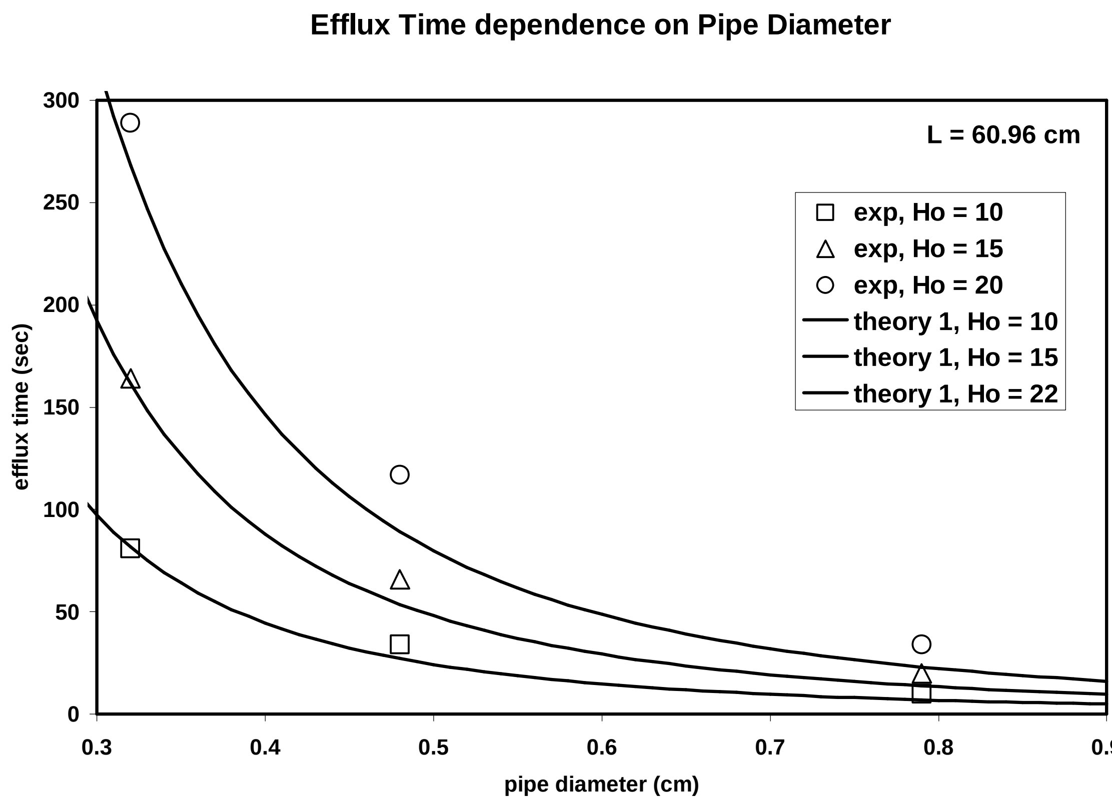

Figure Three. Efflux time as a function of pipe diameter, using the pipe of Length 60.96 cm and initial water heights of 10, 15, and 20 cm. The theory lines represent Case One, equation (8).

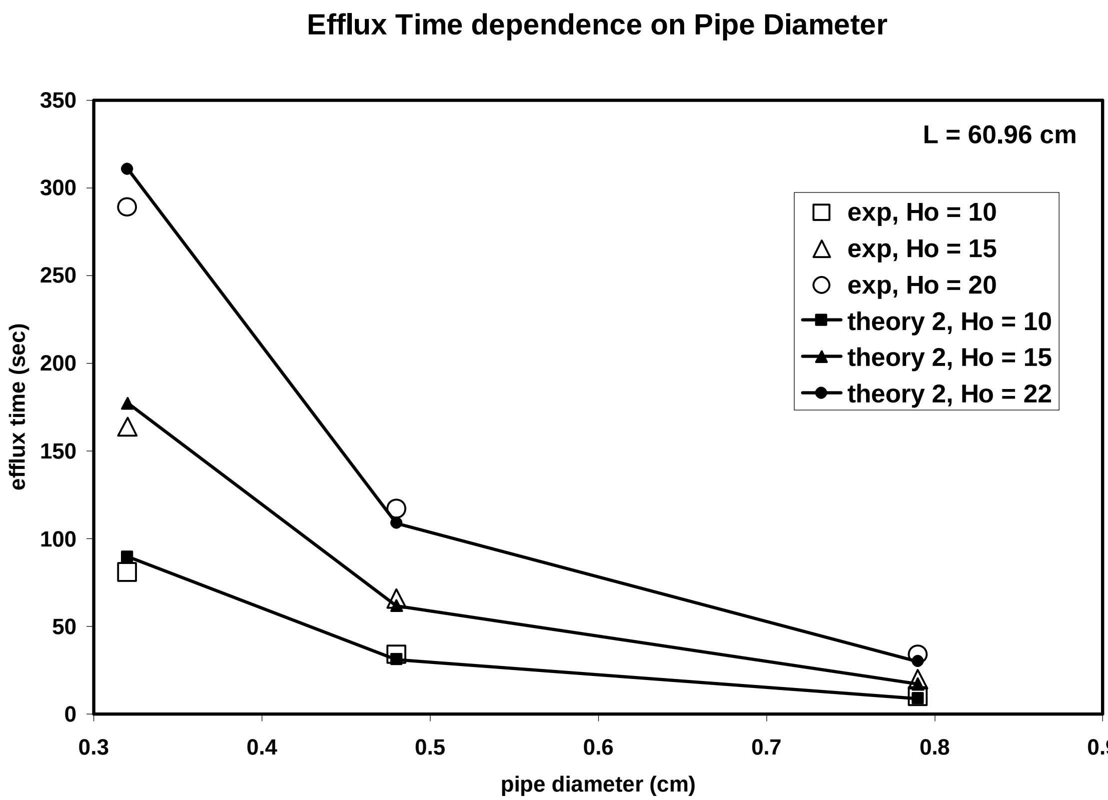

Figure Four. Efflux time as a function of pipe diameter, using the pipe of Length 60.96 cm and initial water heights of 10, 15, and 20 cm. The theory lines represent Case Two, equation (16).