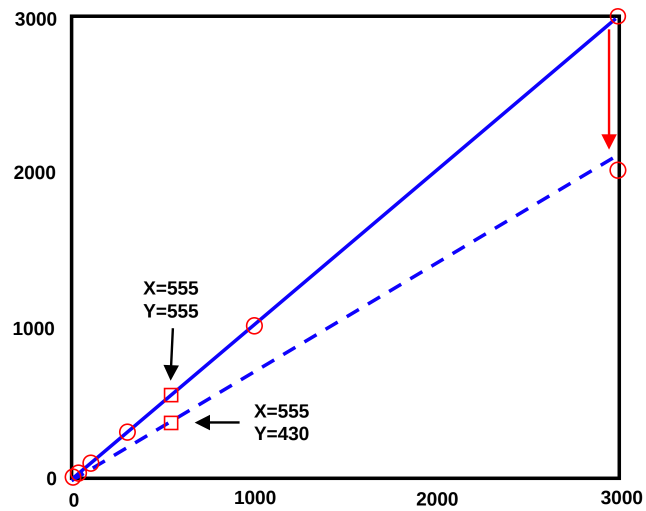
Curve Weighting: Effects of High-End Non-Linearity

A significant deviation from linearity by a single high point completely dominates all other points in the calculations.



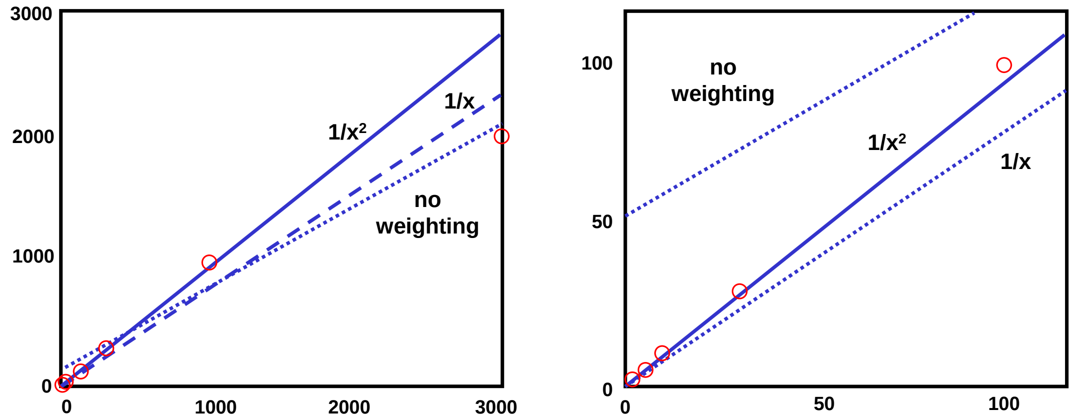
7-28

The illustration shows the effect of a high-end point on a calibration curve running from 1 to 3000 in concentration. The response is assumed to be equal to the concentration (to make it simple). The data pairs are all perfect except the highest one (at x=3000, y=2000).

The defective high point does two things: first, it significantly changes the position of the mean value of Y. Second, it dominates the calculation of the slope.

Again, if we use the idea of a center of gravity for the data, that outlying point is like a weight on the end of a lever, and remember that the longer the lever, the more effect a weight at its end has.

Calibration Curves:
The Effect of Different Weighting Factors



7-29

The illustrations here show the effect of different weighting factors on the calibration curve from the previous slide, running from 1 to 3000 in concentration.

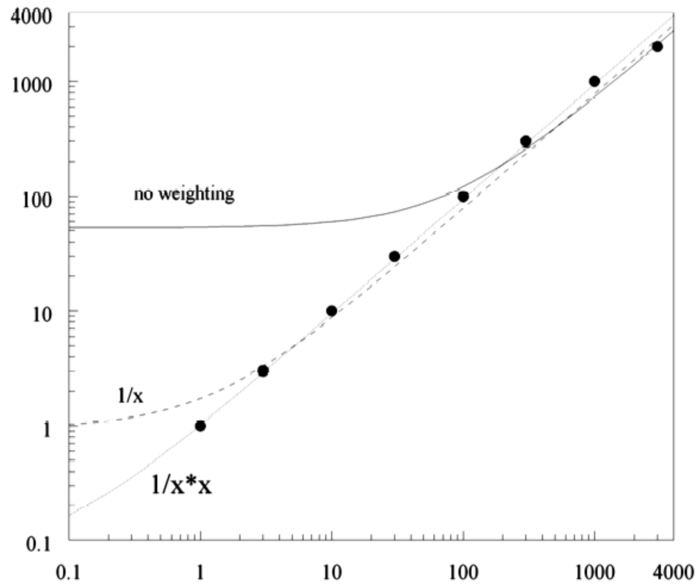
With no weighting, the slope of the line is completely dominated by the highest term. The expanded scale shows that the lowest five points are essentially ignored by the calculation. When weighting of 1/concentration (1/x) is used, the slope more closely approximates the majority of the points, and the intercept approaches zero.

When a weighting of $1/x^2$ is used, the slope approaches unity and the intercept is not significantly different from zero. The fit also essentially ignores the high-end point.

This tends to bother people, that last point hanging out there, ignored. Part of the problem is that it's a visual artifact of the way we normally present data.

Calibration Curves:
A Closer Look at Weighting Effects

Now which fit is better?



7-30

LC RESOURCES

Here the same data and graph are used, except that the axes are now logarithmic. Graphing this way gives a visual approximation of equal weighting for each data point.

The first thing you notice is that the deviation of the highest point no longer predominates the visual distribution of the data points; it's really pretty minor.

The second thing you notice is that the non-weighted best fit line completely ignores the low end of the point set, while the $1/x^2$ weighting looks like a much better choice.

We operate in a linear, Cartesian world for the most part. If you show both the linear and log-log plots to most people the response will likely be "Yeah, but aren't you cheating by squeezing the top part down like that?" The reply might be "What would you have said if I showed you the second plot first?".

Most of us, unconsciously, think "small size, small importance". In analysis, plotting our calibration curves in linear coordinates exclusively may be one of the dumber things we do.