Title - Data Rejection Criteria and Estimates of Standard Deviations of Diffraction Data

Author - F.H.Moore

Affiliation - AINSE (Australian Institute of Nuclear Science and Engineering)

Abstract - (for Kyoto paper XXV-15 in Acta Cryst A28 Supplement 1972 pS256)

A brief survey of publications of structure analyses is sufficient to realise that there are major differences in the methods used by different laboratories in the treatment of diffractometer data. Some of the methods used are of questionable validity and the actual inconsistency between treatments leads to difficulties in estimating the accuracy of published work. The discrepancy index R which could be used as an overall guide to the accuracy of a structure determination is virtually useless in the current situation.

Text

The problem of making a choice as to whether or not to reject diffractometer data we believe to be inaccurate appears, from the inconsistencies in the literature, to be a difficult one. This is in contrast with x-ray photographic data in which the choice is relatively simple: if a reflection cannot be seen on the film it is classed as unobservable and when least squares refinements of structure parameters is used some attempt is usually made to recognise the importance of the weak data by assigning values as appropriate as possible.

Now that diffractometers have come into general use, for some reason I haven't fathomed, we appear to be maintaining a very strong desire to class our observations as "observable" or "unobservable" even though the "unobservable" data has actually been observed and we set up criteria for making a choice between observed data and unobserved data. The 3σ criterion seems to have pride of usage but there are many who use 1σ, 2σ, or even 4σ. In addition, we have a strong group of experimentalists who insist on looking on these "unobserved" reflections as a completely different class and replacing the values which they have actually observed by values which they haven't observed. There is a further group with computer-controlled diffractometers which automatically reject weak reflections. Perhaps these scientists are hazily aware of the ethics of the situation and are relieved by the decision being taken away from them and placed in the hands of a machine.

Personally I am very much opposed to these attitudes. First, they assume there is a rigorous criterion which can be used when in fact there is none. Secondly, there can never be any excuse, in any realm of Science, for replacing observed values by values the experimenter might feel more appropriate no matter how directed is his guess. Scientifically, this is dishonesty in the extreme. The position is illustrated in Slide 1. Thirdly, the further I look into the matter, the more convinced I am that we are becoming almost paranoic in our worship of the R index. We realise that the more weak data we reject the lower our R index becomes, but we conveniently forget or don't even realise that (i) the weak reflections are the most sensitive to parameter shifts (ii) the strong reflections which we are so fond of accepting as good data can have enormous systematic errors and (iii) when we reject weak data we systematically alter our data set.

Let us look at the whole problem a little more closely. First we must realise that statisticians strongly object to the concept of data rejection, unless there is some good explanation, outside the realm of statistics, for the divergent value. This is because such rejection violates the fundamental postulate of statistical interpretation of random errors. The onus is therefore upon us to establish whether or not the individual values we obtain for our weak reflections lie outside the realm of probability. In general the data we obtain for weak reflections, although having a large % error, in fact reflects very adequately what we expect from probability arguments. Many experimenters, however, reject this advice and I can only attribute such mental aberration, to three possibilities: (1) the experimenter is ignorant, (2) the experimenter wants to obtain as good an R index as possible without appearing dishonest, after all, everyone else does it, or (3) he honestly believes that this weak data is useless or even detrimental to his parameter determination. I don’t think I need elaborate on the first possibility however let us consider the second. If we look at Slide 2, we can see an example of how by restricting the data set we can get nice low R indices. We all know full well however that this is not the object of a refinement and in fact by restricting our data set our estimates of parameter errors becomes progressively worse as can also be seen from Slide 2.

Let us then look at the third possibility which asks the question: is this weak data useless data or even detrimental to the estimates of the parameters? I am going to restrict myself to four lines of evidence, although there are more, to establish that the answer to this question is NO.

The first line of evidence I can explain with the aid of Slide 3 which gives some very familiar equations. R is the residual we minimise in least squares, R_1 and R_2 are the conventional and weighted R indices respectively. It takes only a brief glance at these equations to realise that the important factor as far as least squares is concerned (and this holds for the difference fourier too) is the magnitude of the error in each observation and not the percentage error. On the other hand the important factor in the estimates of the R indices is the %error and not the magnitude of the error. Another important point is emphasised by the equation for the least squares residual. The magnitude of the error for our weak reflections is usually smaller than that for the strong reflections and the weight given to the weak reflections in this equation is therefore greater than the weight given to the stronger reflections.

If that isn’t enough evidence, let us go to the second line of evidence which is illustrated in Slide 4. This slide shows two plots, one of variance, (based solely on Poisson statistics) as a function of intensity (dotted line) and the full line plot of the variance as determined from an analysis of equivalent reflections. The difference between the curves represents a systematic error, which in fact is very common and probably the major systematic error in diffractometer data. It is very obvious from these plots that the very weak reflections contain very little of this systematic error whereas in the very strong reflections the error can be predominant. Again the argument against throwing away weak data is obvious. What of the third line of evidence then. Let us look at Slide 5 which gives two Wilson Type plots. The full line gives an idealised result from a full set of data while the dotted line represents that from a restricted set of data in which the weak reflections were rejected and averages calculated ignoring these reflections. It is very obvious from this plot that when weaker data is rejected the data set is interfered with SYSTEMATICALLY. A random interference could be tolerated but a systematic interference is dangerous in the least.

Let us then go over to the fourth line of evidence given in Slide 6, which gives the results of two refinements of potassium hydrogen oxalate, the oxalate group only being shown. It can be seen that there is good agreement of bond angles and lengths between the values obtained from the full set of data and the restricted set of data. The important point to note about this result is that the restricted set of data was restricted in the sense that it contains only those reflections which would normally be rejected.

The final point I wish to raise is the treatment of negative intensities. Again, honesty necessitates that we retain the –ve sign. One might be tempted to say that we are at liberty to change all negative values to zero. Even if all the negative values were truly zero this would not be satisfactory because we have forgotten the positive side of the true zero distribution. This is illustrated in Slide 7. The shaded area is given to depict the true zero intensities. It can also be seen from this slide that the true zero reflections are not the only reflections for which –ve values can be obtained. We therefore have no alternative but to maintain the sign of our –ve measurements throughout our calculations. Anything else is contrary to ethics. In fact, there is no problem in maintaining the sign – the observant amongst you will have noticed that throughout this talk and in my slides I have always used F^2 rather than F. This is because the only way to handle –ve data is to use F^2. $-F^2$ is meaningful whereas –F is not.