International Tables for Crystallography (2006). Vol. F, Section 11.5.6, p. 238.

11.5.5. Generalization of the procedure for averaging reflection intensities

Once the scale factors of all frames are determined, they need to be applied to the reflection intensities and error estimates. The reflection intensities with the same reduced Miller indices can then be averaged.

When method 2 is used for averaging, the determination of $\langle I_h \rangle$ is more complicated as there are as many estimates of the full intensity I_{hi} as there are partial reflections h_{im}. Therefore, intensity averaging of reflection h has to be done in two steps. First, for every reflection h_i, the intensity estimates from all partial observations will be the weighted mean, where the weights are based on the estimated standard deviations of each intensity measurement. In the second step, the average is taken over the i different scaled intensities for the observed reflections.

The selection of reflections useful for averaging is the same as for scaling (Table 11.5.3.1), except that it is no longer necessary to reject reflections that have insignificant intensities. Applying a σ cutoff while averaging the scaled intensities will lead to a statistical bias of the weaker reflection intensities.

For samples of three or more equivalent reflections, it is necessary to consider the absolute values of the differences between individual intensities and the median of the sample: $|I_{hi} - I_{\rm median}|$. The outliers can be detected by several statistical tests and, once detected, can be either down-weighted or rejected. When the sample consists of only two reflections, they can be considered a 'discordant pair' if the difference between their intensities is not warranted by the estimated errors and, hence, both reflections can be rejected (Blessing, 1997).

Averaging intensities estimated according to method 2 has an advantage over method 1 as outliers and discordant pairs can be 'screened' at two levels: firstly, when the estimates of the full reflection intensity I_{hi}, calculated by expression (11.5.2.2) from different parts of the same reflection, are considered, and secondly when the mean intensities $\langle I_{hi} \rangle$ from different reflections are considered.

11.5.6. Estimating the quality of data scaling and averaging

A commonly used estimate of the quality of scaled and averaged Bragg reflection intensities is $R_{\rm merge}$. Useful definitions of R factors are:

$$R_{\rm merge} = \quad R_1 = \left[\left(\sum_h\sum_i |I_{hi} - \langle I_h \rangle|\right) \Big/ \sum_h\sum_i |I_{hi}|\right] \times 100\%, \tag{11.5.6.1}$$

$$R_2 = \left\{\left[\sum_h\sum_i (I_{hi} - \langle I_h \rangle)^2\right] \Big/ \sum_h\sum_i I_{hi}^2\right\} \times 100\% \tag{11.5.6.2}$$

and

$$R_w = \left\{\left[\sum_h\sum_i W_{hi}(I_{hi} - \langle I_h \rangle)^2\right] \Big/ \sum_h\sum_i W_{hi} I_{hi}^2\right\} \times 100\%. \tag{11.5.6.3}$$

The linear (R_1), square (R_2) and weighted (R_w) R factors can be subdivided into resolution ranges, intensity ranges, reflection classes, frame number and regions of the detector surface. When method 1 is used, reflections h_i can be grouped in terms of the sums of partialities of contributing partial reflections h_{im}.

The R-factor variation depends on the properties of the detector with respect to intensities. Generally the R factor decreases as intensity increases. Thus, the R factor generally increases with resolution. Any deviation from this behaviour might indicate a problem in the data collection due to nonlinearity of the detector response, ice diffuse diffraction, or any other stray effects superimposed on the crystal diffraction.

A useful indicator of the quality of the intensity estimates of partial reflections is the mean ratio of calculated partiality to observed partiality:

$$r_p = \langle p_{him}^{\rm calc} / p_{him}^{\rm obs} \rangle = \langle p_{him}^{\rm calc} \langle I_h \rangle / I_{him} \rangle. \tag{11.5.6.4}$$

The deviation of this ratio from unity can be examined as a function of the reflection intensity, resolution and calculated partiality.

The comparison of R factors for centric and noncentric reflections can be used to determine the significance of an anomalous-scattering effect. The quality of the anomalous-dispersion signal can be assessed by calculation of the scatter, σ_{Ih}, where

$$\sigma_{Ih} = \left\{\left[1/(n-1)\right] \sum_n (\langle I_h \rangle - I_{hn})^2\right\}^{1/2} \tag{11.5.6.5}$$

and $\langle I_h \rangle$ is the average of the n measurements of the full reflection intensities I_{hn}. The σ_{Ih} values for noncentric reflections can be compared to the scatter, σ_{Ih}^+ or $\sigma_{\bar{I}h}^-$, of reflections differing only in absorption while excluding Bijvoet opposites. The mean scatter is calculated from all σ_{Ih} values,

$$\langle \sigma_{Ih} \rangle = (1/h) \sum_h \left\{\left[1/(n-1)\right] \sum_n (\langle I_h \rangle - I_{hn})^2\right\}^{1/2}. \tag{11.5.6.6}$$

The ratios $\langle \sigma_{Ih} \rangle / \langle \sigma_{Ih}^+ \rangle$ and $\langle \sigma_{Ih} \rangle / \langle \sigma_{\bar{I}h}^- \rangle$ should be larger than unity for significant anomalous-dispersion data.

11.5.7. Experimental results

11.5.7.1. *Variation of scale factors versus frame number*

If scale factors are to make physical sense, their behaviour with respect to the frame number has to be in accordance with the known changes in the beam intensity, crystal condition and detector response.

The scaling of a φX174 procapsid data set (Dokland *et al.*, 1997) was performed using methods 1 and 2 as described here and using *SCALEPACK* (Otwinowski & Minor, 1997) (Fig. 11.5.7.1). Graphs (*a*) and (*b*) in Fig. 11.5.7.1 have four segments corresponding to four synchrotron beam 'fills'. All three methods give scale factors within 5% of each other (Figs. 11.5.7.1*c* and *d*). However, for the first and last frame of each 'fill' the results can differ by as much as 15%. Both method 1 and *SCALEPACK* produce physically wrong results in that the scale factors of these frames look like outliers compared to the scale factors of the neighbouring frames. By contrast, method 2 provides consistent scale factors for these frames. Although the algorithm used by *SCALEPACK* for scaling frames with partial reflections has never been disclosed, the similar results obtained by method 1 and *SCALEPACK* suggest that *SCALEPACK* might be using an algorithm similar to that of method 1 (Fig. 11.5.7.1*d*).

Attempts at scaling a data set of a frozen crystal of HRV14 (Rossmann *et al.*, 1985, 1997) failed with method 1 as a result of gaps in the rotation range for the first 20 frames, causing singularity of the normal equations matrix. When frames without useful neighbours were excluded, the cubic symmetry of the crystal was sufficient for successful scaling. In contrast, method 2 did not have any problems with the whole data set, and the results obtained with method 2 showed greater consistency than those obtained with method 1 or *SCALEPACK* (Fig. 11.5.7.2).