Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 18 June 2013
Accepted 10 August 2013

More about residual values

Julian Henn* and Andreas Schönleber

Laboratory of Crystallography, University of Bayreuth, 95440 Bayreuth, Germany. Correspondence e-mail: julian.henn@uni-bayreuth.de

The usual residual values are complemented by expectation values based solely on the experimental data and the number of model parameters. These theoretical R values serve as benchmark values when all of the basic assumptions for a least-squares refinement, *i.e.* no systematic errors and a fully adequate model capable of describing the data, are fulfilled. The prediction of R values as presented here is applicable to any field where model parameters are fitted to data with known precision. For crystallographic applications, F^2-based residual benchmark values are given. They depend on the first and second moments of variance, intensity and significance distributions, $\langle\sigma^2\rangle$, $\langle {I_o}^2\rangle$, $\langle {I_o}^2/\sigma^2\rangle$. Possible applications of the theoretical R values are, for example, as a data-quality measure or the detection of systematic deviations between experimental data and model predicted data, although the theoretical R values cannot identify the origin of these systematic deviations. The change in R values due to application of a weighting scheme is quantified with the theoretical R values.

1. Introduction

Take a list of experimentally observed data together with their standard uncertainties (s.u.'s), and adjust model parameters against the experimental data in a least-squares refinement. The deviation between observed and calculated data (data predicted from the model) is expressed by a residual factor, the R value. What is an appropriate R value, given the data? And is it possible to predict a benchmark value from the experimental data alone or is a model required in any case?

A residual factor can only describe the fit between model and data; it does not say anything about the degree of agreement that is expected. Therefore it is a good descriptor of the fit rather than a quality indicator. For a quality indicator a reference value is necessary. If no explicit reference value is given one might involuntarily refer to zero, which is not appropriate: as experimental data are always affected by noise an R value of zero would be very suspicious. But how much deviation can be accepted? How much deviation is too much or too little to be 'appropriate'? For example what does an R value of 0.05 tell us? This value might appear satisfying if it is not known that with the given data an R value of, say, 0.02 is attainable. In this case the actual or *de facto* R value is larger than the expected one. This may indicate systematic sources of error, a poor model, underfitting or incorrect s.u.'s. If, in contrast, the attainable R value is determined to be 0.10, while the *de facto* value is 0.05, this means that the refinement is overfitting, or that again the s.u.'s are incorrect.

The experimental data are assumed to be of high accuracy and their precision is assumed to be given by the s.u.'s. Assuming that the model is fully adequate to describe the data a Gaussian distribution in the residuals should result from the least-squares fit (Prince, 2004). Residual distributions that are very different from a Gaussian may indicate that the residuals are not statistically independent, but correlated.

The prediction of R values as presented here is completely general and applies to any situation where model parameters are fitted to experimental data, as is the case in many scientific fields. In the following, however, we refer without loss of generality to crystallographic R values.

Many different conventions for crystallographic residual values are applied. They may be based on the structure-factor amplitudes, $|F|$, or on the observed corrected intensities, F^2, with unit or statistical weights or with application of a weighting scheme. For the present purpose it is important to refer to the R value that corresponds to the residual sum that is minimized in the least-squares procedure. The observed intensities are the experimental data; therefore the discussion is focused on the corresponding R values based on F^2. The R values of interest are discussed in detail in the next section.

Comparative R values that might serve as a reference exist, for example the R-free (Brünger, 1992), which needs the specification and refinement of a model. Another comparative R value is the theoretical R value obtained under the special assumption of a Poisson distribution of the detected particles (Henn & Meindl, 2010) of the photon or neutron beam. Also empirically predicted R values for macromolecular refinements exist (Urzhumtsev *et al.*, 2009). In this work we want to derive R values based on the experimental data without the necessity to actually specify model parameters and to conduct a least-squares refinement, and

without any additional assumptions, apart from those necessary for a least-squares refinement.

In the following sections, the individual R values are presented with their definitions, along with a brief discussion of the properties and their relations to other R values. The predicted R value is introduced, which is an R value calculated from the experimental data without explicitly specifying a model. It gives a realistic estimate of the attainable R value based on the s.u.'s, provided all the above-mentioned assumptions are met. A discussion of the predicted R values follows, where these are applied to experimental data.

2. R values

2.1. The R_2 value

2.1.1. Definition.

$$R_2 = \left[\frac{\sum (I_\mathrm{o} - I_\mathrm{c})^2}{\sum I_\mathrm{o}^2}\right]^{1/2}. \quad (1)$$

It is called R_2(obs) or R_2(all) depending on whether it refers to the observed data with significance exceeding some threshold value $I_\mathrm{o}/\sigma(I_\mathrm{o}) > x$ or to all data. The threshold is a matter of convention; a typical value is for example $x = 2$ for standard structure determination and $x = 3$ for charge-density studies.

For least-squares refinements usually all data are used as otherwise data would be suppressed systematically, thereby leading to bias in the model parameters (Hirshfeld & Rabinovich, 1973).

The R_2 value has the important property that it is by definition independent of the standard uncertainties, which do not enter equation (1). In this sense R_2 is a 'hard' measure of the difference between observed and calculated intensities. In contrast, measures including a weight yield different results when the weights are changed. The R_2 value can therefore be seen as the expected value when all data are treated alike and independent of their individual precision.

2.1.2. Prediction. The residuals $(I_\mathrm{o} - I_\mathrm{c})/[\sigma(I_\mathrm{o})]$ from a least-squares fit with weights $1/\sigma^2$ are normally distributed when the following conditions are fulfilled: (*a*) the data are accurate, (*b*) the (hypothetical) structure model is appropriate such that the calculated intensity values I_c in equation (1) are unbiased on the true intensities and (*c*) the s.u.'s from the reflection file describe adequately the dispersion of the observed intensity mean values about the true intensities.

From a normal distribution of residuals it follows that the typical squared unweighted residual is of the order of magnitude of the corresponding variance,

$$(I_\mathrm{o} - I_\mathrm{c})^2 \simeq \alpha \sigma^2(I_\mathrm{o}), \quad (2)$$

with $\alpha \leq 1$ being a constant that takes into account the number of independent model parameters N_par and the number of reflections used in the least-squares refinement, N_ref,

$$\alpha = \frac{N_\mathrm{ref} - N_\mathrm{par}}{N_\mathrm{ref}}. \quad (3)$$

Individual values of unweighted squared residuals in equation (2) may be larger by a factor 3 to 5 (depending on the number of reflections in the data set), or smaller than the corresponding variance. However, these independent statistical fluctuations tend to cancel each other more efficiently the more reflections $i = 1, \ldots, N$ are involved. Therefore, if equation (2) applies, the following equation will apply too:

$$\sum_{i=1}^{N} \left(I_{\mathrm{o},i} - I_{\mathrm{c},i}\right)^2 \simeq \alpha \sum_{i=1}^{N} \sigma^2(I_{\mathrm{o},i}). \quad (4)$$

From this follows an equation we will refer to repeatedly by dividing both sides by N on

$$\left\langle (I_\mathrm{o} - I_\mathrm{c})^2 \right\rangle \simeq \alpha \left\langle \sigma^2 \right\rangle, \quad (5)$$

where $\sigma^2(I_\mathrm{o})$ is abbreviated by σ^2. From equation (5) the predicted value is obtained:

$$R_2^\mathrm{pred} = \left(\alpha \frac{\left\langle \sigma^2 \right\rangle}{\left\langle I_\mathrm{o}^2 \right\rangle}\right)^{1/2}. \quad (6)$$

Note the logic of the arguments: from a Gaussian distribution of weighted residuals it follows that the squared unweighted residuals group around the corresponding variances as expressed in equation (5). The inverse is *not* true: from the validity of equation (5), in any given particular case, one cannot conclude that the underlying distribution is a normal distribution.

When a weighting scheme is applied, the corresponding $\hat{\sigma}$ values are taken (the 'reduced σ values', for details see §2.3 and equations therein):

$$R_2^\mathrm{pred} = \left(\alpha \frac{\left\langle \hat{\sigma}^2 \right\rangle}{\left\langle I_\mathrm{o}^2 \right\rangle}\right)^{1/2}. \quad (7)$$

When there is a danger of confusion we will refer to R_2^pred either as $R_2^\mathrm{pred}(\sigma^2)$, equation (6), or as $R_2^\mathrm{pred}(\hat{\sigma}^2)$, equation (7).

Although it has been stated that the R_2 value is independent of the s.u.'s, they now enter the R_2^pred value *via* equation (5). This equation reflects the assumptions of the existence of an adequate model and the adequateness of the experimental values for the s.u.'s. No model parameters have to be specified. The predicted R_2 value becomes smaller when the experimental s.u.'s are smaller, for example by adding redundancy, or when the significance $\left\langle I_\mathrm{o}/[\sigma(I_\mathrm{o})] \right\rangle$ increases in any other way. Setting $\alpha = 1$ in equation (6) may serve as an upper bound for the actual R_2 and can be calculated directly from a data set without any model. Alternatively, one may derive α from the number of atoms, with four parameters per atom for coordinates and isotropic atomic displacement parameters, or nine parameters per atom for coordinates and anisotropic atomic displacement parameters. Of course, parameter restraints, *e.g.* from special positions, have to be taken into account too. In this way it can be used as a quality descriptor of the data set when the assumptions for least-squares fitting are met.

2.2. The R_{F^2} value

2.2.1. Definition.

$$R_{F^2} = \frac{\sum |I_o - I_c|}{\sum |I_o|}. \quad (8)$$

This residual value is also independent of the s.u.'s, in the sense that these do not enter the formula. Therefore it is a robust absolute measure like the above-mentioned R_2 value. It has the same information content as R_2, as it again measures the difference between observed and calculated intensities without weights but with a different metric. It is called for example $R2$(obs) or $R2$(all) in the *SHELX* software package (Sheldrick, 2008) and $R(F^2)$ or $R\mathrm{all}(F^2)$ in the *XD2006* software package (Volkov *et al.*, 2006).

2.2.2. Prediction. Using the standard assumption of unbiasedness of calculated intensities on the true intensities and an equation similar to equation (5), namely

$$|I_o - I_c| \simeq \left(\frac{2}{\pi}\right)^{1/2} \alpha\sigma, \quad (9)$$

the R_{F^2} value is predicted directly from the experimental data by

$$R_{F^2}^{\mathrm{pred}} = \left(\frac{2}{\pi}\right)^{1/2} \frac{\alpha\langle\sigma\rangle}{\langle|I_o|\rangle}. \quad (10)$$

The factor $(2/\pi)^{1/2}$ is necessary to convert the expectation value from the variance to the standard deviation. The assumption entering here is the normal distribution of residuals with mean value zero and standard uncertainty σ. The factor α takes again the degrees of freedom of the model into account: a more flexible model should result in a closer agreement between observed and calculated intensities.

2.3. The $wR(F^2)$ value

2.3.1. Definition.

$$wR(F^2) = \left[\frac{\sum w(I_o - I_c)^2}{\sum wI_o^2}\right]^{1/2}. \quad (11)$$

This is the R value relevant for least-squares refinements on experimental data, as the counter contains the residual sum that is minimized. Different weights are in use. When $w = 1/\sigma^2(I_o)$ is chosen, this is referred to as 'statistical weights', whereas weights chosen according to a weighting scheme have no special name. A popular weighting scheme applied for example in *SHELX* (Sheldrick, 2008) is of the type

$$w = \frac{1}{\sigma^2 + (aP)^2 + bP} \quad (12)$$

with

$$P = \frac{I_o + 2I_c}{3}. \quad (13)$$

One may interpret the inverse weight as a new variance that is larger or equal to the original one:

$$\hat{\sigma}^2 = \frac{1}{w} \geq \sigma^2. \quad (14)$$

These values were already used in equation (7) to obtain a prediction for R_2 under the assumption that the $\hat{\sigma}^2$ values instead of the σ^2 values adequately describe the measurement errors. The $wR(F^2)$ value has the interesting property that it is invariant under a multiplicative scaling transformation of the weights. Setting for example $w = 1/\sigma'^2$ with $\sigma' = \beta\sigma$ and $\beta \in \mathbb{R}$, $\beta > 0$ and putting this into the definition leads to $wR(F^2)\big|_{w=1/\sigma'^2} = wR(F^2)\big|_{w=1/\sigma^2}$. A scaling transformation of this type would of course affect other R values and the goodness of fit (GoF), as well as the precision of the refined model parameter and the number of reflections classified as 'observed'. The average significance of the reflections $\langle I_o/[\sigma(I_o)]\rangle$ would be affected too; however, the resulting model parameter values would be the same for a least-squares refinement with $w = 1/\sigma'^2$ and for $w = 1/\sigma^2$.

When the s.u.'s in the reflection file describe the variance of the data adequately (and do not over- or underestimate it) and if the structure model is adequate and no systematic errors are in the data (like neglected absorption or extinction correction), a weighting scheme is not necessary. Note that in this case also the GoF will be automatically correct without enforcing it to assume a value close to 1, *i.e.* by deliberately refining the values a and b in the weighting scheme, equation (12), such that the GoF equals 1. As a weighting scheme allows for larger differences [all contributions in equation (12) are positive] between observed and calculated intensities, it will also increase the R_2 and R_{F^2} values when the F_c put into these definitions stem from a least-squares refinement with a weighting scheme. For both parameters a, b approaching zero, the inverse weights approach the variance again, $\lim_{a,b\to 0} \hat{\sigma}^2 = \sigma^2$. Setting the weights to the inverse variances, $w = 1/\sigma^2$, is the standard setting for least-squares refinement on accurate data. This means that data with higher significance have a larger influence on the model parameters than data with lower significance when the intensity is the same. The application of a weighting scheme is the admission of differences between observed and calculated intensities that are inconsistent with the s.u.'s, but no information is given about the origin of these differences.

The $wR(F^2)$ value is equal to the R_2 value when the weights are set to $w = 1$, or *any* other constant value, like $w = 1000$: $R_2 = wR(F^2)\big|_{w=\mathrm{const}}$. Setting the weights to a constant value means that all data are treated the same way, regardless of how precisely they are measured, *i.e.* no data are preferred over other data.

2.3.2. Prediction. To obtain a prediction we assume again equation (5) in the case of statistical weights, or, for employing a weighting scheme,

$$\langle (I_o - I_c)^2 \rangle \simeq \alpha\langle\hat{\sigma}^2\rangle. \quad (15)$$

For statistical weights this leads to the prediction

$$wR(F^2)\big|_{w=1/\sigma^2}^{\mathrm{pred}} = \frac{\alpha^{1/2}}{\left(\langle I_o^2/\sigma^2\rangle\right)^{1/2}}, \quad (16)$$

Table 1
Statistical descriptors of the data sets.

The angle brackets ⟨⟩ indicate averaging over all data used for the refinements in the data sets. From left to right: number of data set, mean intensity, mean s.u., mean squared intensity, mean variance, mean significance and mean squared significance.

Data set	$\langle I_o \rangle$	$\langle \sigma(I_o) \rangle$	$\langle I^2(I_o) \rangle$	$\langle \sigma^2(I_o) \rangle$	$\langle I_o/[\sigma(I_o)] \rangle$	$\langle I_o^2/[\sigma^2(I_o)] \rangle$
1	3.00×10^4	809.90	5.33×10^9	1.44×10^6	23.47	1338.36
2	3.22×10^4	663.63	5.44×10^9	1.08×10^6	30.26	1877.82
3	3.30×10^4	440.47	6.54×10^9	6.75×10^5	44.21	4161.13
4	3.03×10^4	608.16	3.25×10^9	4.94×10^5	48.39	4342.47
5	1.93×10^3	74.90	4.92×10^7	9.32×10^3	15.50	877.81
6	5.62×10^3	318.70	1.02×10^8	2.29×10^5	13.70	254.26
7	3.27×10^3	74.98	8.67×10^7	2.00×10^4	29.00	1999.62
8	1.90×10^3	48.02	2.75×10^7	1.54×10^4	27.54	1410.21
9	1.13×10^3	81.88	8.05×10^6	2.18×10^4	8.77	128.80
10	1.95×10^3	71.53	2.25×10^7	1.36×10^4	16.31	486.64
11	6.01×10^2	20.92	1.65×10^6	7.90×10^2	23.63	1225.00
12	4.43×10^2	41.32	3.65×10^6	1.45×10^4	6.88	93.44
13	1.77×10^3	65.44	2.43×10^7	1.86×10^4	16.35	527.63
14	1.13×10^4	425.08	6.92×10^8	4.83×10^5	17.86	528.74
15	5.19×10^2	9.27	2.10×10^6	4.60×10^2	37.29	2092.63
16	1.08×10^3	26.62	8.35×10^6	1.52×10^3	27.39	1541.70
17	3.82×10^2	8.98	3.25×10^6	1.03×10^3	21.72	836.19
18	6.14×10^1	2.84	3.41×10^4	9.53	16.31	1070.49
19	7.69×10^1	1.81	1.22×10^5	8.64	21.66	1757.36
20	6.14×10^2	30.87	1.00×10^7	5.70×10^4	11.73	286.27
21	1.28×10^2	4.03	6.81×10^5	8.27×10^2	16.31	721.03
22	4.91×10^3	364.17	1.47×10^8	1.79×10^5	16.12	1127.78
23	3.89×10^2	35.59	4.31×10^6	1.83×10^3	7.76	256.54
24	1.89×10^2	9.85	5.49×10^5	4.00×10^2	10.22	231.22
25	2.48×10^2	2.14	1.49×10^6	4.95×10^1	64.24	6167.76
26	4.15×10^2	10.29	1.12×10^6	2.96×10^2	29.10	2411.91
27	1.34×10^3	50.23	1.32×10^7	3.75×10^3	22.11	1337.07
28	8.73×10^2	28.15	1.14×10^7	1.50×10^3	23.75	2210.08
29	4.84×10^2	13.80	2.31×10^6	3.38×10^2	25.97	2457.13
30	3.08×10^2	10.82	1.65×10^6	6.25×10^2	18.94	912.62
31	8.49×10^2	28.02	1.14×10^7	2.19×10^3	19.88	1135.52
32	8.24×10^1	1.48	1.33×10^5	2.38×10^1	33.59	2580.03
33	9.95×10^1	3.47	1.69×10^5	6.91×10^1	18.10	1363.91
34	2.04×10^2	8.20	8.91×10^5	6.66×10^2	13.11	482.94
35	2.48×10^2	16.41	8.64×10^5	1.79×10^3	10.70	267.86
36	8.17×10^2	36.43	4.03×10^7	2.84×10^4	13.84	520.60
37	2.01×10^2	9.37	2.96×10^5	1.22×10^2	16.14	656.99

References: (1–3) He *et al.* (2013), (4) Jennene Boukharrata *et al.* (2013), (5) Zhou *et al.* (2013), (6) Deng *et al.* (2013), (7–8) Wang *et al.* (2013), (9–10) Zhang *et al.* (2013), (11) Huang (2013), (12) Zhou & Wang (2013), (13) Wu & Jin (2013), (14) Luo (2013), (15–17) Krause *et al.* (2013), (18–19) Arkhipov *et al.* (2013), (20–21) Su *et al.* (2013), (22) Arderne (2013), (23–25) Bats *et al.* (2013), (26–27) Smith & Wermuth (2013*a*), (28–29) Smith & Wermuth (2013*b*), (30–31) Castillo *et al.* (2013), (32–36) Gomes *et al.* (2013), (37) Görbitz & Yadav (2013).

and for application of a weighting scheme with a reduced σ value $\hat{\sigma}$ correspondingly to the prediction

$$wR(F^2)\big|_{w=1/\hat{\sigma}^2}^{\text{pred}} = \frac{\alpha^{1/2}}{\left(\left\langle I_o^2/\hat{\sigma}^2 \right\rangle\right)^{1/2}}. \qquad (17)$$

The denominator in equation (16) contains the mean squared significance and the denominator in equation (17) the reduced mean squared significance.

3. Discussion of the predicted *R* values

Although it is expected that strong reflections lead to small *R* values, none of the predicted *R* values depends on the observed intensity alone. The observed intensity is always related somehow to the s.u.'s or variances. This reflects the fact that intensities are on an arbitrary scale. When the intensities are rescaled by multiplication with a scale factor, the standard uncertainties have to be multiplied with the same scale factor in order to keep the significance constant (Henn & Meindl, 2010).

To find confidence in the predicted *R* values, the prediction is compared to the actual outcome based on published experimental data after a brief discussion of the non-standard statistical descriptors of the data sets. Data published online between 9 and 18 April 2013 were taken from *Acta Crystallographica Section C*, **69**, May 2013. All data sets with refinements against F^2 were analysed; the data set with *F* values (Mesto *et al.*, 2013) is omitted from the analysis. Several data descriptors important for the theoretical *R* values and references for the data sets are given in Table 1. Table 2 shows these data descriptors after application of the weighting scheme and the corresponding parameters. The list of predicted and published *R* values is given in the supplementary material.[1]

The predicted *R* values can be derived from the statistical descriptors of the data sets given in Tables 1 and 2. The average significance of the data sets comprises the range from 6.88 (data set No. 12) to 64 (data set No. 25), *i.e.* approximately one order of magnitude, whereas the mean intensities and the mean variances each comprise several orders of magnitude (columns 2 and 3 of Table 1). Table 2 shows the parameters *a* and *b* from the weighting scheme. The effect of these parameters is given in Table 2 in the form of the ratio $\langle \sigma^2/\hat{\sigma}^2 \rangle$ in column 6. For data set No. 25 ($a = 0.0460$, $b = 0.1500$) the weighting scheme produces $\langle \sigma^2/\hat{\sigma}^2 \rangle = 0.08$, which is by far the lowest value. The experimental σ values are clearly underestimated as can be seen from the corresponding normal probability plot in the supplementary material. This underestimation of experimental s.u.'s leads to the low predicted *R* value. The weighting scheme effectively reduces the mean significance of this data set from 64.24, which is the largest mean significance, to 12.80. For data set No. 14, with the largest value for $a = 11.2420$, the reduction in significance is comparably moderate from 17.86 to 12.17. Why are the data sets so differently affected? For the parameters *a* and *b* from the weighting scheme to create an impact, for example, on the significance, more than the mere numerical

[1] A list of predicted and published *R* values, and normal probability plots for all data sets as obtained from the DRK-plot software implemented in *WinGX* (Farrugia, 2012) are available from the IUCr electronic archives (Reference: KX5020). Services for accessing these data are described at the back of the journal. For more information on normal probability plots, see Abrahams & Keve (1971).

Table 2
Statistical descriptors of the data sets after application of the weighting scheme.

The angle brackets ⟨⟩ indicate averaging over all data used for the refinements in the data sets. From left to right: number of data set, mean reduced s.u., mean reduced variance, mean reduced significance, mean reduced squared significance, mean reduction factor of variance, and parameters a and b from the weighting scheme as given in equations (12) and (13).

Data set	$\langle\hat{\sigma}(I_o)\rangle$	$\langle\hat{\sigma}^2(I_o)\rangle$	$\langle I_o/[\hat{\sigma}(I_o)]\rangle$	$\langle I_o^2/[\hat{\sigma}^2(I_o)]\rangle$	$\langle\sigma^2(I_o)/\hat{\sigma}^2(I_o)\rangle$	a	b
1	1045.47	3.21×10^6	16.90	516.76	0.83	0.0180	0.0000
2	910.02	2.69×10^6	21.57	724.39	0.79	0.0170	0.0000
3	594.06	1.34×10^6	31.49	1628.79	0.75	0.0097	1.3211
4	716.40	7.55×10^5	36.62	2047.28	0.78	0.0082	1.3622
5	125.47	8.69×10^4	8.09	117.64	0.69	0.0374	1.7477
6	433.61	4.61×10^5	9.88	118.81	0.62	0.0421	4.9899
7	263.00	3.60×10^5	7.28	76.05	0.34	0.0626	8.0375
8	139.87	9.36×10^4	8.16	93.04	0.28	0.0521	5.2961
9	143.16	9.17×10^4	5.09	35.70	0.59	0.0941	1.6163
10	110.18	4.77×10^4	10.95	171.22	0.71	0.0387	0.2140
11	26.64	1.57×10^3	16.52	441.35	0.79	0.0217	0.0000
12	43.32	1.47×10^4	6.15	75.87	0.89	0.0000	0.6236
13	72.52	2.33×10^4	14.06	345.45	0.87	0.0132	0.2094
14	577.51	9.12×10^5	12.17	225.48	0.62	0.0209	11.2420
15	31.38	5.29×10^3	10.63	147.00	0.17	0.0451	0.9370
16	77.84	2.71×10^4	8.71	106.13	0.38	0.0548	1.8743
17	36.60	2.25×10^4	5.73	45.65	0.25	0.0836	0.9747
18	4.01	4.61×10^1	9.16	165.10	0.83	0.0324	0.0000
19	5.40	3.13×10^2	6.86	75.75	0.41	0.0487	0.2145
20	50.97	8.94×10^4	6.21	66.51	0.57	0.0550	0.7500
21	8.63	1.86×10^3	5.75	66.54	0.38	0.0380	0.5000
22	399.58	2.30×10^5	11.11	306.49	0.89	0.0187	0.0000
23	40.80	5.51×10^3	5.73	91.31	0.92	0.0300	0.0000
24	14.05	1.50×10^3	7.30	95.16	0.81	0.0450	0.0000
25	12.78	3.27×10^3	12.80	193.03	0.08	0.0460	0.1500
26	24.74	3.06×10^3	11.18	173.59	0.59	0.0488	0.0000
27	78.79	2.22×10^4	11.97	208.69	0.71	0.0372	0.0000
28	65.55	3.91×10^4	8.07	103.90	0.66	0.0571	0.0000
29	33.66	7.37×10^3	8.97	119.82	0.61	0.0541	0.1041
30	24.54	6.09×10^3	6.88	78.69	0.50	0.0544	0.6625
31	61.51	2.25×10^4	6.71	84.66	0.48	0.0376	3.2029
32	6.11	5.90×10^2	8.14	91.77	0.36	0.0656	0.0688
33	10.53	1.51×10^3	5.66	44.00	0.53	0.0898	0.0839
34	30.95	1.65×10^4	3.92	23.02	0.57	0.1408	0.0000
35	33.35	8.68×10^3	4.24	28.55	0.48	0.0855	1.4865
36	86.10	2.66×10^5	5.21	38.07	0.35	0.0773	2.8776
37	18.26	1.38×10^3	7.38	78.80	0.62	0.0625	0.1448

values of a and b play a role. For example, a and b will affect the data sets differently when the intensities and/or the s.u.'s are on a different scale. The parameter a will affect the strong reflections most. Therefore the distribution of intensities plays a role too. The most significant reflection in data set No. 25 has a significance of $I_o/[\sigma(I_o)] = 339.0$ before and $I_o/[\hat{\sigma}(I_o)] = 24.5$ after application of the weighting scheme (data not shown), whereas for data set No. 14 the change in the maximum significance is from 58.8 to 35.1, *i.e.* the maximum significance in data set 14 is much smaller than that of data set No. 25. Application of $a \neq 0$ leads to a larger R value and to a smaller GoF. There are obviously many interesting links between the data descriptors; however, we prefer to stop exploring these links in order to proceed with the theoretical R values.

To explore the relation between the theoretical R values in more detail, the questions to be answered are: is there a hierarchy of the predicted R values in the sense that one is always larger than the other? How does the weighting scheme affect the predicted R values? What is the meaning of the predicted R values?

For a first test, the $wR(F^2)\big|^{\rm pred}_{w=1/\hat{\sigma}^2}$ value [equation (17)] was calculated from the set of experimental I_o and $\hat{\sigma}(I_o)$ for all data sets. It is expected that the *de facto* $wR(F^2)\big|_w$ value as calculated from the data set including calculated intensities [equation (11)] is close to the predicted one. The curves are shown in Fig. 1. Recalculation of the published $wR(F^2)$ values ensures that the same definitions are used, and that parameter values for the weighting scheme and for the number of refined parameters of the structure model are correct.

The published R values were recalculated from the published diffraction data. The agreement is almost perfect, with one minor exception for data set No. 12: the $wR(F^2)$ values are 0.114 (published) and 0.119 (recalculated). The reasons for these differences could not be tracked down.

Fig. 1 shows that the predicted values are close to the *de facto* values, with a tendency for the predicted values to stay below the actual ones. Only data sets Nos. 18, 22, 23, 27 and 28 differ in this respect, where the predicted values are larger than the *de facto* ones. These are the data sets with $S < 1$ (see Table 3). A large difference between *de facto* and predicted values is observed only for data set No. 34. For this data set the value for the GoF is with 1.24 the largest in the whole table. Multiplying $wR(F^2)\big|^{\rm pred}_{w=1/\hat{\sigma}^2}$ with the respective values for the GoF results in virtually indistinguishable *de facto* and predicted values.

It can be thus concluded that the prediction, which assumes GoF = 1, works well. This confirms the assumptions made for deriving the predictions, in particular equations (5) and (15).

3.1. $wR(F^2)\big|^{\rm pred}_{w=1/\sigma^2}$ and $wR(F^2)\big|^{\rm pred}_{w=1/\sigma^2}$

The denominator in equation (17) can be interpreted as a mean squared weighted significance. The weighted significance will always be smaller than or equal to the usual significance as $\hat{\sigma}^2 \geq \sigma^2$. Therefore the inequality holds:

$$wR(F^2)\big|^{\rm pred}_{w=1/\sigma^2} \leq R(F^2)\big|^{\rm pred}_{w=1/\hat{\sigma}^2}, \qquad (18)$$

where the equal sign applies when the parameters a and b for the weighting scheme approach zero, *i.e.* for statistical weights.

Fig. 2 shows that the reduced significance $w = 1/\hat{\sigma}^2$ increases the $wR(F^2)$ value. The space between the red and the green curve can be interpreted as an estimate for the costs of application of the weighting scheme, as predicted and *de facto* $wR(F^2)\big|_{w=1/\hat{\sigma}^2}$ values agree fairly well (Fig. 1). In other words: if application of a weighting scheme was not necessary, $wR(F^2)$ values close to the green line were possible with the given

Table 3
Comparison between reduced variances and unweighted squared residuals: ratios of mean values, coefficient β [equation (25)] from a fit of unweighted squared residuals against $\alpha\hat{\sigma}^2$, correlation coefficient *cc* between $(I_o - I_c)^2$ and $\alpha\hat{\sigma}^2$, and published values of GoF.

Data set	$\langle(I_o - I_c)^2\rangle/\langle\alpha\hat{\sigma}^2\rangle$	β	*cc*	GoF
1	2.33	4.52	0.95	1.00
2	2.69	4.29	0.74	1.07
3	1.43	1.95	0.85	1.09
4	2.42	9.81	0.92	1.12
5	4.85	11.4	0.94	1.04
6	7.30	26.4	0.82	1.11
7	1.88	2.74	0.79	1.05
8	2.51	3.91	0.96	1.05
9	1.15	1.41	0.59	1.10
10	1.15	2.36	0.50	1.04
11	1.57	2.46	0.62	1.02
12	0.53	0.38	0.93	1.08
13	1.05	0.90	0.64	1.06
14	0.99	0.49	0.47	1.08
15	1.36	1.31	0.98	1.06
16	1.62	1.99	0.84	1.05
17	1.01	0.57	0.50	1.06
18	2.22	4.42	0.69	0.93
19	0.33	0.06	0.27	1.05
20	0.83	0.89	0.96	1.01
21	0.85	0.76	0.59	1.01
22	0.73	0.33	0.21	0.88
23	4.30	8.68	0.85	0.92
24	0.42	0.32	0.74	1.04
25	11.5	18.1	0.95	1.05
26	1.23	1.14	0.58	1.10
27	1.23	1.16	0.87	0.96
28	0.53	0.36	0.89	0.95
29	0.97	2.10	0.89	1.04
30	2.69	3.89	0.96	1.06
31	5.46	9.83	0.94	1.08
32	0.32	0.26	0.91	1.17
33	0.52	0.45	0.91	1.10
34	0.88	1.85	0.86	1.23
35	1.87	1.16	0.32	1.10
36	0.29	0.03	0.20	1.08
37	1.53	1.82	0.60	1.05

experimental data sets. But there are discrepancies between model predicted and experimental data that are inconsistent with the experimental s.u.'s. This makes the application of a weighting scheme necessary (or alternatively a change of the structure model and/or data processing), and therefore the resulting $wR(F^2)$ values are larger.

3.2. $R_2^{\rm pred}$ and $wR(F^2)|_{w=1/\sigma^2}^{\rm pred}$

In the preceding paragraph it was shown that $wR(F^2)|_{w=1/\hat{\sigma}^2}^{\rm pred} \geq wR(F^2)|_{w=1/\sigma^2}^{\rm pred}$. Is there a similar relationship between $R_2^{\rm pred}$ and $wR(F^2)|_{w=1/\sigma^2}^{\rm pred}$?

From the definitions of the predicted values it follows that R_2 is not always smaller than $wR(F^2)|_{w=1/\sigma^2}^{\rm pred}$. It is smaller when the following inequality holds:

$$\left\langle \frac{I_o^2}{\sigma^2} \right\rangle < \frac{\langle I_o^2 \rangle}{\langle \sigma^2 \rangle}. \qquad (19)$$

What does this inequality mean and when is it fulfilled?

For the case of only weak correlation between the intensities and their respective variances it is expected that

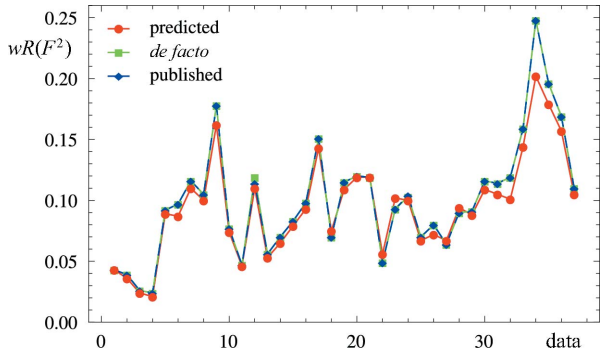

Figure 1
Predicted and *de facto* $wR(F^2)$ values and $wR(F^2)$ values from the literature. The *de facto* $wR(F^2)|_{w=1/\hat{\sigma}^2}$ value was calculated according to equation (11) with $\hat{\sigma}^2$ from equation (14) from the published diffraction data (green line). It should exactly match the values taken from the literature (blue, dashed). Predictions (red) are according to equation (17). Numerical values are given in the supplementary material. References for the data sets are given in Table 1. The individual values are connected with a line as a guide for the eye.

$$\left\langle \frac{I_o^2}{\sigma^2} \right\rangle > \frac{\langle I_o^2 \rangle}{\langle \sigma^2 \rangle}, \qquad (20)$$

because then the summands in which an accidentally low variance σ^2 coincides with an accidentally large value of I_o^2 will increase the term on the left-hand side of equation (20).

Other factors may be important too. For example, when the correlation between observed intensities and variances is strong and the variance is in proportion to the observed intensity [which is a minimum variance (Henn & Meindl, 2010), both the variance from the beam-inherent Poisson statistics as well as fluctuations in the beam stability should be in proportion to sufficiently large I_o], the inequality for the squared significance, equation (19), turns into an inequality for the observed intensity:

$$\langle I_o \rangle^2 \geq \langle I_o^2 \rangle. \qquad (21)$$

The validity of this inequality depends on the distribution of intensities. In the unrealistic cases where the intensity is about the same for all reflections the equal sign should hold. For other cases it is expected that $\langle I_o \rangle^2 < \langle I_o^2 \rangle$, which violates

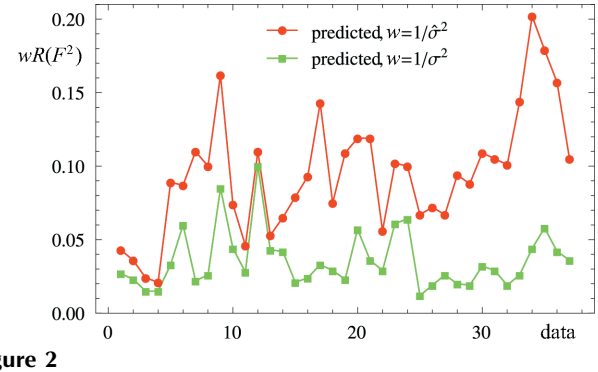

Figure 2
Predicted $wR(F^2)$ values. For all data sets it is predicted that $wR(F^2)|_{w=1/\hat{\sigma}^2}^{\rm pred}$ (red) $\geq wR(F^2)|_{w=1/\sigma^2}^{\rm pred}$ (green) according to equation (18). The space between red and green curves represents the costs of the weighting procedure.

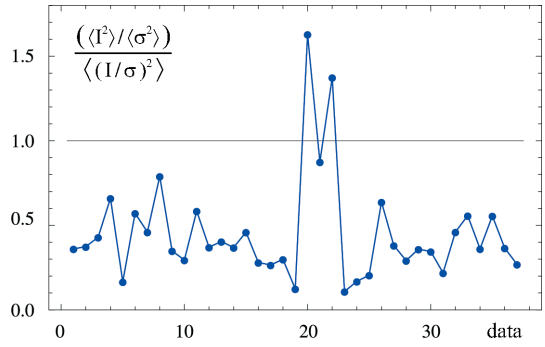

Figure 3
Ratio of $\langle I_o^2\rangle/\langle\sigma^2\rangle$ and $\langle I_o^2/\sigma^2\rangle$. Additionally the horizontal line is plotted, where the ratio equals 1. For more information see the text.

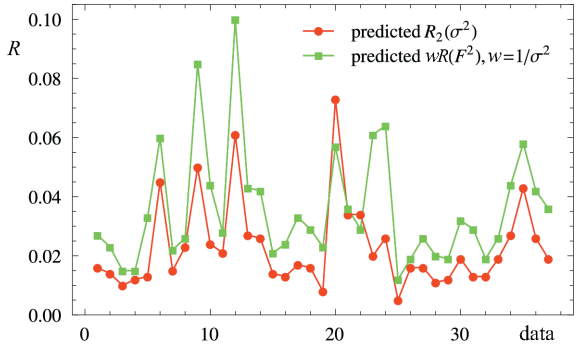

Figure 4
The predicted R values $R_2^{\rm pred}$ (red) and $wR(F^2)\big|_{w=1/\sigma^2}^{\rm pred}$ (green). For more information see the text.

the inequality [equation (21)]. We are not going further into this discussion here.

A plot of the ratio $\langle I_o^2\rangle/\langle\sigma^2\rangle$ and $\langle I_o^2/\sigma^2\rangle$ is shown in Fig. 3. For the data sets above the constant line (Nos. 20 and 22) it is predicted that $R_2^{\rm pred} \leq wR(F^2)\big|_{w=1/\sigma^2}^{\rm pred}$ according to equation (19).

A plot of $R_2^{\rm pred}$ and $wR(F^2)\big|_{w=1/\sigma^2}^{\rm pred}$ is shown in Fig. 4. The $R_2^{\rm pred}$ value tends to be lower than the corresponding weighted R value with statistical weights. Only in the cases of data sets Nos. 20 and 22 is the $R_2^{\rm pred}$ value larger than the $wR(F^2)\big|_{w=1/\sigma^2}^{\rm pred}$ value. This was predicted by equation (19) and is visualized in Fig. 3.

It is beyond the scope of the present work to track down the exact differences between the data sets.

4. Discussion of the *de facto* and the predicted *R* values

For a comparison between *de facto* R values and predicted ones one has to consider a few points: the *de facto* R values are obtained by fitting model parameters to the experimental data with a given metric, *i.e.* with weighted or not weighted differences or squared differences of $|F|$ or F^2 values. A least-squares refinement minimizes the residual sum and relies on the assumption that the resulting residuals follow the normal distribution (Prince, 2004). A Gaussian distribution results if the contributing errors are random. If there are systematic errors, however, a non-Gaussian distribution of residuals may result.

For the interpretation of predicted and *de facto* R values it is helpful to keep in mind that the prediction *also* relies on the assumption of a Gaussian distribution of residuals. Deviations between predicted and actual R values (of the same metric that was used for least-squares residual minimization) therefore indicate that not all assumptions hold. But the inverse is not true: if the predicted and actual R values of the metric that was used for least-squares residual minimization coincide, this does *not* prove that all assumptions are met (that 'everything is okay'). For example, the distribution of residuals may follow a different distribution than a Gaussian distribution and yield by chance the same R value as the predicted one.

4.1. *De facto* and predicted $wR(F^2)\big|_{w=1/\hat\sigma^2}$

As already shown and discussed in Fig. 1, there is a fair agreement between observed and predicted values.

4.2. *De facto* and predicted $wR(F^2)\big|_{1/\sigma^2}$

The *de facto* $wR(F^2)\big|_{1/\sigma^2}$ value is always larger than the predicted one. This is expected, as the *de facto* R value is calculated from the set of I_o, I_c, $\sigma(I_o)$ with experimental data I_o, $\sigma(I_o)$ and calculated intensities I_c obtained from a model refinement with non-statistical weights $w = 1/\hat\sigma^2$.

The red line in Fig. 5 is a benchmark value for the $wR(F^2)\big|_{w=1/\sigma^2}$ of the respective data set: if weighting was not necessary and all assumptions were met (basically, no systematic errors) a *de facto* $wR(F^2)$ value close to the red line was expected. For some data sets the absolute differences between *de facto* and predicted R values are rather small, *e.g.* sets 1, 2, 3, 4, 22, 23. It is tempting to relate this to the parameters of the weighting scheme. This can be achieved by quantification of how much the weighting affects the s.u.'s, for example by calculating $\langle\sigma^2/\hat\sigma^2\rangle$, with $\hat\sigma$ from equation (14). This ratio can be interpreted as the factor by which the effective variance is on average reduced through the weighting scheme. A graph of this value is shown in Fig. 6 together with a line at the value 0.75. This line was chosen to discriminate between weighting schemes that affect the variances only

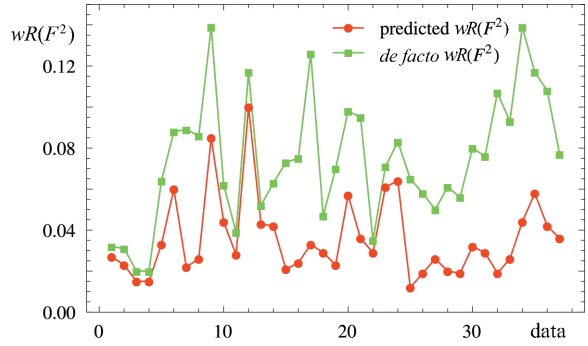

Figure 5
Presentation of the predicted $wR(F^2)\big|_{w=1/\sigma^2}$ (red) and the corresponding *de facto* value as calculated from the diffraction data (green).

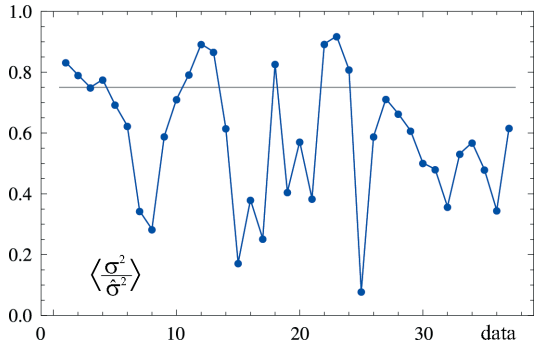

Figure 6
Effective reduction of the variance by employing a weighting scheme: shown is the average value $\langle \sigma^2/\hat{\sigma}^2 \rangle$, with $\hat{\sigma}^2$ from equation (14) where the average is taken over the respective data set. Additionally a line with the value 0.75 is drawn.

slightly (these are above the line) and those that reduce the effective variance to a large extent (under the line).

Comparison of Fig. 6 with Fig. 5 shows that the data sets with a large ratio (above the line) also exhibit a low absolute difference in the actual and predicted R values.

4.3. *De facto* and predicted R_2 values

De facto and predicted R_2 values are shown in Fig. 7.

For the *de facto* R_2 value, which is calculated according to equation (1) from the published data, only I_o and I_c enter the calculation. The I_c were determined by a least-squares refinement employing a weighting scheme, thereby allowing for larger deviations between I_o and I_c in comparison to *e.g.* statistical weights. Therefore, the expected R_2 value is calculated according to equation (7) from α, the set of I_o and $\hat{\sigma}$ values.

From the definitions of $R_2^{de\,facto}$ and $R_2^{\rm pred}$ their ratio follows as

$$\frac{R_2^{de\,facto}}{R_2^{\rm pred}} = \left[\frac{\langle (I_o - I_c)^2 \rangle}{\alpha \langle \hat{\sigma}^2 \rangle} \right]^{1/2}. \qquad (22)$$

When equation (5) is valid, ratios close to 1 are expected. These are displayed in Fig. 8.

The ratio of *de facto* and predicted R_2 values comprises a large range from approximately 0.5 for data sets 19, 32 and 36 to almost 3.5 for data set 25.

The actual R_2 being lower than the predicted one indicates that in these cases the (mean) squared difference between observed and calculated intensities is lower than the mean weighted variance multiplied by α:

$$\langle (I_o - I_c)^2 \rangle < \alpha \langle \hat{\sigma}^2 \rangle. \qquad (23)$$

This might indicate that in these particular cases over-fitting has occurred; however, among these sets only set No. 28 shows a value $S < 1$. Multiplying $R_2^{\rm pred}$ with the respective GoF values does not yield fewer, but even more cases, where the actual R_2 value is smaller than the predicted one. This is in contrast to the case of the wR values, where the deviation of the GoF from 1 was the missing piece of information that established exact agreement between observed and predicted R values.

Explicit calculation of the ratio $\langle (I_o - I_c)^2 \rangle / \alpha \langle \hat{\sigma}^2 \rangle$ shows that among the sets 12, 19, 24, 28, 32, 33 and 36 the largest ratio is 0.53 (for sets No. 12 and 28, see Table 3); the other ratios in this set are all even smaller. For these sets the squared difference between observed and calculated intensity values is on average only half of the respective variance, or even less. For other data sets the ratio is much larger than 1, for example for data sets No. 6, 25 and 31 the mean squared difference between observed and calculated intensities is about 7.3, 11.5 and 5.5 times larger than the mean squared (weighted) variance that has been used to adjust the model parameters and to derive the set of I_c!

How does this conform to the GoF values that are close to 1 for all sets?

The GoF is an *independent* measure of the fit quality only under certain circumstances. One important assumption is that the residuals are normally distributed, *i.e.* Gaussian, with zero mean and unit variance. A normal distribution of residuals indicates that the errors are independent and uncorrelated. The other assumption is that the variances are known.

If they are not known, the sum of squared residuals $\chi^2 = \sum [(I_o - I_c)/\sigma(I_o)]^2$ can be used to guess or adjust the variances. *In this case a good fit is assumed*. The moment this

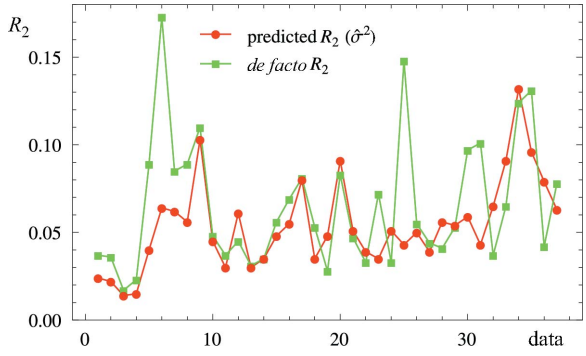

Figure 7
De facto and predicted R values: the predicted R_2 is shown in red and the *de facto* one in green. For the prediction the reduced variance $\hat{\sigma}^2$ was used.

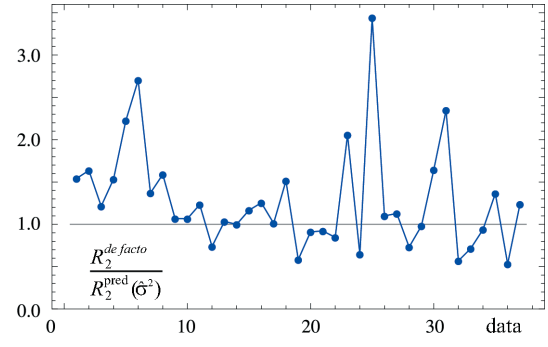

Figure 8
Ratio $R_2^{de\,facto}/R_2^{\rm pred}(\hat{\sigma}^2)$. For the numerical values of $R_2^{de\,facto}$ and $R_2^{\rm pred}(\hat{\sigma}^2)$ see the supplementary material. Additionally a line with value 1 is depicted.

assumption enters the process, the GoF ceases to serve as an independent measure of quality. To cite the numerical recipes: 'Obviously, this approach prohibits an independent assessment of goodness-of-fit, a fact occasionally missed by its adherents' (Press *et al.*, 1992, p. 661). In other words, by applying a weighting scheme to force the GoF value into the neighbourhood of unity, the GoF value becomes meaningless.

The reason is that under the assumptions necessary for a least-squares refinement, the residuals $(I_o - I_c)/\sigma$ are random numbers from a Gaussian ($\mu = 0, \sigma_{Gauss} = 1$) distribution and not correlated. The squared residuals $(I_o - I_c)^2/\sigma^2$ entering the GoF are χ^2-distributed and are also not correlated. For uncorrelated random numbers the distributions can be rewritten

$$\left\langle \frac{(I_o - I_c)^2}{\sigma^2} \right\rangle \simeq \frac{\left\langle (I_o - I_c)^2 \right\rangle}{\langle \sigma^2 \rangle}. \qquad (24)$$

This is used *e.g.* for the prediction of R_2 and leads, under the assumptions made, to GoF $\simeq 1$. The inverse is *not* true. From GoF $\simeq 1$ one cannot conclude that equation (24) is valid. To illustrate this point in more detail, Table 3 shows that the average squared unweighted residuals $\left\langle (I_o - I_c)^2 \right\rangle$ can be much larger or much smaller than the average reduced variance $\left\langle \hat{\sigma}^2 \right\rangle$ (column 2), despite all of the refinements having a GoF close to 1 (column 5). For a linear fit of unweighted squared residuals against α times the reduced variances $\hat{\sigma}^2$ with a linear fit parameter β,

$$(I_o - I_c)^2 = \beta\alpha\hat{\sigma}^2, \qquad (25)$$

values between 0.03 (data set No. 36) and 26.5 (data set No. 6) are obtained for β (column 3). The correlation coefficient *cc* between $\alpha\hat{\sigma}^2$ and $(I_o - I_c)^2$ is generally high with lowest values for data sets No. 19 ($cc = 0.27$), 22 ($cc = 0.21$), 35 ($cc = 0.32$) and 36 ($cc = 0.20$) (column 4). Fit parameters $\beta \simeq 1$ and correlation coefficients close to zero were expected for all data sets in the case of random residuals.[2] So both correlation coefficients and values $\beta \neq 1$ may indicate that the residuals $(I_o - I_c)/\hat{\sigma}$ are not Gaussian-distributed. A further hint is the maximum value of $(I_o - I_c)^2/\alpha\beta\hat{\sigma}^2$ after adjustment of β: for those data sets with $\beta \simeq 1$, Nos. 26, 27 and 35, these maximum values are 50.73, 26.91 and 69.10 (data not shown). These numbers indicate that the weighting scheme was not applied merely in order to 'correct' the s.u.'s, for in this case of an adequate model and s.u.'s in need of correction, a maximum value not much larger than 25 was expected (a value of 25 corresponds to a 5σ event, that happens in one out of 1.74×10^6 cases). The residual distributions are shown in the form of normal probability plots in the supplementary material. All deviate in a systematic rather than random way from the expected behaviour, the *de facto* frequencies of residuals are all above or under the diagonal line, which represents the expected residuals. None of the residual distributions is very close to a normal distribution.

[2] After application of a weighting scheme, the fit residuals cannot be completely uncorrelated any more; however, correlation coefficients as high as 0.98 (data set No. 15), 0.96 (data sets 8, 20, 30), 0.95 (sets 1 and 25) need not appear despite employing a weighting scheme. In these cases of very high correlation it would be appropriate to indicate this, for example by writing $\hat{\sigma}(I_o, I_c)$.

5. Conclusion

The theoretical *R* values give a realistic estimate of the attainable *R* values without the need to explicitly construct a model. The theoretical *R* values assume GoF = 1 and they reproduce the *de facto R* values corresponding to the minimized residuals in this case. The basic assumptions for the application of theoretical *R* values are the same as those for an application of a least-squares fit of model parameters to experimental data, *i.e.* accuracy of the data, adequateness of s.u.'s and the possibility of specifying an adequate model. Accuracy of the data implies that measurement errors are stochastic rather than systematic. The adequateness of the s.u.'s leads to the correct model parameters and GoF = 1. This assumption is less crucial if the s.u.'s are correct relative to each other, but on a wrong scale, as it would not affect the model parameter values but only their s.u.'s and the GoF. The possibility of specifying an adequate model is important too. Consider, for example, the case of a high-resolution data set and an independent-atom model, or the case of a data set from a modulated structure and a not-modulated structure model. These assumptions lead to a normal distribution of residuals. In the case of the s.u.'s being on a wrong scale, the distribution is still of a Gaussian type, but no longer with unit variance. As the requirements for the application of the theoretical *R* values are quite general, they are applicable to any field where model parameters are fitted to experimental data with known precision, or with unit weights. The theoretical *R* values can also be applied to other modelling processes that assume Gaussian errors like applications of the maximum entropy method. We expect the theoretical *R* values to be helpful as experimental data-quality descriptors, as a tool for detection of systematic errors in future applications and as a stopping criterion in crystallographic maximum entropy applications. They may also serve to quantify the costs, in terms of increased *R* values, of changes applied to the standard uncertainties of the experimental data, for example by applying a weighting scheme. The next step of this research is to apply the theoretical *R* values to charge-density studies and to extend the formalism to cases where a normal distribution of residuals does not apply, such as in maximum likelihood refinements.

K. Meindl is acknowledged for providing software code. The authors also thank the referees for their constructive remarks and suggestions.

References

Abrahams, S. C. & Keve, E. T. (1971). *Acta Cryst.* A**27**, 157–165.
Arderne, C. (2013). *Acta Cryst.* C**69**, 526–528.
Arkhipov, S. G., Zakharov, B. A. & Boldyreva, E. V. (2013). *Acta Cryst.* C**69**, 517–521.
Bats, J. W., Schell, P. & Engels, J. W. (2013). *Acta Cryst.* C**69**, 529–533.
Brünger, A. (1992). *Nature (London)*, **355**, 472–475.

Castillo, J. C., Abonía, R., Cobo, J. & Glidewell, C. (2013). *Acta Cryst.* C**69**, 544–548.
Deng, Z.-H., Peng, Z.-J. & Huang, J. (2013). *Acta Cryst.* C**69**, 467–470.
Farrugia, L. J. (2012). *J. Appl. Cryst.* **45**, 849–854.
Gomes, L. R., Low, J. N., Rodrigues, A. S. M. C., Wardell, J. L., Lima, C. H. da S. & Souza, M. V. N. de (2013). *Acta Cryst.* C**69**, 549–555.
Görbitz, C. H. & Yadav, V. N. (2013). *Acta Cryst.* C**69**, 556–559.
He, H., Stoyko, S. S., Mar, A. & Bobev, S. (2013). *Acta Cryst.* C**69**, 455–459.
Henn, J. & Meindl, K. (2010). *Acta Cryst.* A**66**, 676–684.
Hirshfeld, F. L. & Rabinovich, D. (1973). *Acta Cryst.* A**29**, 510–513.
Huang, X.-H. (2013). *Acta Cryst.* C**69**, 483–485.
Jennene Boukharrata, N., Duclère, J.-R., Laval, J.-P. & Thomas, P. (2013). *Acta Cryst.* C**69**, 460–462.
Krause, J. A., Zhao, D., Chatterjee, S., Yeung, B. M., Connick, W. B. & Collins, S. N. (2013). *Acta Cryst.* C**69**, 498–502.
Luo, S.-H. (2013). *Acta Cryst.* C**69**, 494–497.
Mesto, E., Scordari, F., Lacalamita, M., De Cola, L., Ragni, R. & Farinola, G. M. (2013). *Acta Cryst.* C**69**, 480–482.
Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in C: the Art of Scientific Computing.* New York: Cambridge University Press.
Prince, E. (2004). *Mathematical Techniques in Crystallography and Materials Science.* Berlin, Heidelberg, New York: Springer Verlag.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Smith, G. & Wermuth, U. D. (2013*a*). *Acta Cryst.* C**69**, 534–537.
Smith, G. & Wermuth, U. D. (2013*b*). *Acta Cryst.* C**69**, 538–543.
Su, B.-Y., Wang, J.-X., Liu, X. & Li, Q.-D. (2013). *Acta Cryst.* C**69**, 522–525.
Urzhumtsev, A., Afonine, P. V. & Adams, P. D. (2009). *Acta Cryst.* D**65**, 1283–1291.
Volkov, A., Macchi, P., Farrugia, L. J., Gatti, C., Mallinson, P. R., Richter, T. & Koritsanszky, T. (2006). *XD2006. A computer program package for multipole refinement, topological analysis of charge densities and evaluation of intermolecular energies from experimental and theoretical structure factors.*
Wang, Y.-L., Shao, Z.-J. & Qian, S.-S. (2013). *Acta Cryst.* C**69**, 471–475.
Wu, D.-H. & Jin, L. (2013). *Acta Cryst.* C**69**, 491–493.
Zhang, Y.-F., Ma, J.-P., Liu, Q.-K. & Dong, Y.-B. (2013). *Acta Cryst.* C**69**, 476–479.
Zhou, M., Song, L., Niu, F., Shu, K. & Chai, W. (2013). *Acta Cryst.* C**69**, 463–466.
Zhou, W. & Wang, J. (2013). *Acta Cryst.* C**69**, 486–490.