Pramāṇa, Vol. 17, No. 6, December 1981, pp. 499–508. © Printed in India.

Theoretical evaluation of the overall values of unnormalized discrepancy indices for truncated data in crystals obeying Wilson distributions*

D VELMURUGAN and S PARTHASARATHY
Department of Crystallography and Biophysics, University of Madras, Guindy Campus, Madras 600 025, India

MS received 3 September 1981

Abstract. Theoretical expressions for the local values of six types of unnormalized R-indices are derived for an imperfectly related incomplete model of a crystal (centrosymmetric and non-centrosymmetric) with truncated data which is characterized by a truncation limit y_t. These indices depend on the parameters σ_1, D and y_t. In the situations of practical interest (*i.e.*, $\sigma_1^2 > 0{\cdot}3$ and $y_t < 0{\cdot}2$) R-indices for the centrosymmetric case decrease as y_t increases while these for the non-centro-symmetric case remain more or less constant.

Keywords. Theoretical evaluation; unnormalized R-indices; truncated data.

1. Introduction

The theory and method of obtaining the overall values of six types of normalized R-indices for truncated data corresponding to the imperfectly related incomplete models have been dealt with in two recent papers (Parthasarathy and Velmurugan 1981; Velmurugan and Parthasarathy 1981—hereafter PV 1981; VP 1981). In this paper we shall deal with the corresponding problem for the six unnormalized R-indices. We shall follow the notation in the earlier papers (PV 1981; VP 1981).

2. Derivation of the theoretical expressions for the overall values of the unnormalized R-indices

2.1 *General results*

The notation and definition of the six unnormalized R-indices are given in table 1. Following the procedure used in Parthasarathy and Ponnuswamy (1979) (see also PV 1981) the theoretical expressions for the overall values of these indices can be obtained in terms of their local values in the various ranges of sin θ/λ (S, hereafter) and these are also given in table 1. It is seen from table 1 that the overall values of the various R-indices can be obtained from a knowledge of the theoretical expectation values of simple functions of the normalized variabes y_N and y_P^c and the required expectation values can be obtained by making use of the joint *pdf*

*Contribution No. 569

Table 1. Definition of the overall values of six unnormalized *R*-indices for truncated data and related results.

No.	Name of the *R*-index	Notation for the overall value of the index	Overall value of the Index: Definition*	Overall value of the Index: Equivalent expression	Definition of the local value of the unnormalized indices based on the normalized variables y_N and y_P^c
1.	Conventional index based on *F*	$[\bar{R}(F)]_t$	$\frac{\Sigma^t \lvert \lvert F_N \rvert - \lvert F_P^c \rvert \rvert}{\Sigma^t \lvert F_N \rvert}$	$\frac{\sum_s f_s n_s [R(y)]_t}{\sum_s f_s n_s}$	$[R(y)]_t = \frac{\langle \lvert y_N - \sigma_1 y_P^c \rvert \rangle_t}{\langle y_N \rangle_t}$
2.	Conventional index based on *I*	$[\bar{R}(I)]_t$	$\frac{\Sigma^t \lvert I_N - I_P^c \rvert}{\Sigma^t I_N}$	$\frac{\sum_s f_s^2 n_s [R(z)]_t}{\sum_s f_s^2 n_s}$	$[R(z)]_t = \frac{\langle \lvert y_N^2 - \sigma_1^2 y_P^{c^2} \rvert \rangle_t}{\langle y_N^2 \rangle_t}$
3.	Booth type index based on *F*	$[_B\bar{R}(F)]_t$	$\frac{\Sigma^t (\lvert F_N \rvert - \lvert F_P^c \rvert)^2}{\Sigma^t \lvert F_N \rvert^2}$	$\frac{\sum_s f_s^2 n_s [_BR(y)]_t}{\sum_s f_s^2 n_s}$	$[_BR(y)]_t = \frac{\langle (y_N - \sigma_1 y_P^c)^2 \rangle_t}{\langle y_N^2 \rangle_t}$
4.	Booth type index based on *I*	$[_B\bar{R}(I)]_t$	$\frac{\Sigma^t (I_N - I_P^c)^2}{\Sigma^t I_N^2}$	$\frac{\sum_s f_s^4 n_s [_BR(z)]_t}{\sum_s f_s^4 n_s}$	$[_BR(z)]_t = \frac{\langle (y_N^2 - \sigma_1^2 y_P^{c2})^2 \rangle_t}{\langle y_N^4 \rangle_t}$
5.	Fractional type index based on *F*	$[\bar{R}^f(F)]_t$	$\frac{2}{N_0} \sum^t \left\lvert \frac{\lvert F_N \rvert - \lvert F_P^c \rvert}{\lvert F_N \rvert + \lvert F_P^c \rvert} \right\rvert$	$\frac{\sum_s n_s [R^f(y)]_t}{\sum_s n_s}$	$[R^f(y)]_t = 2 \left\langle \left\lvert \frac{y_N - \sigma_1 y_P^c}{y_N + \sigma_1 y_P^c} \right\rvert \right\rangle_t$
6.	Fractional type index based on *I*	$[\bar{R}^f(I)]_t$	$\frac{2}{N_0} \sum^t \left\lvert \frac{I_N - I_P^c}{I_N + I_P^c} \right\rvert$	$\frac{\sum_s n_s [R^f(z)]_t}{\sum_s n_s}$	$[R^f(z)]_t = 2 \left\langle \left\lvert \frac{y_N^2 - \sigma_1^2 y_P^{c^2}}{y_N^2 + \sigma_1^2 y_P^{c^2}} \right\rvert \right\rangle_t$

*The summation Σ^t is over all the independent reflections of the truncated data. The summation Σ_s is over the different narrow ranges of $\sin\theta/\lambda$. N_0 is the number of reflections in the truncated data.

$\sigma_1^2 = \langle \lvert F_P^c \rvert^2 \rangle / \langle \lvert F_N \rvert^2 \rangle$, where $\lvert F_N \rvert$ is the magnitude of the structure factor of a reflection **H** for the crystal and $\lvert F_P^c \rvert$ is the corresponding calculated value for a model consisting of *P* atoms out of a total of N atoms in the unit cell.

$P_t(y_N, y_P^c)$ of y_N and y_P^c applicable to the truncated data. This joint *pdf* has been shown to be (PV 1981; VP 1981)

$$P_t(y_N, y_P^c) = \frac{1}{\beta_C} \phi_C (y_N, y_P^c) \text{ for the } C \text{ case}$$

$$= \frac{1}{\beta_{NC}} \phi_{NC} (y_N, y_P^c) \text{ for the } NC \text{ case}$$

$$y_t \leqslant y_N < \infty, \; 0 \leqslant y_P^c < \infty, \tag{1}$$

where we define ϕ_C and ϕ_{NC} to be

$$\phi_C = \frac{2}{\pi \, \sigma_B} \exp \left[- \frac{(y_N^2 + y_P^{c^2})}{2 \, \sigma_B^2} \right] \cosh \left(\frac{\sigma_A \, y_N \, y_P^c}{\sigma_B^2} \right), \tag{2}$$

and

$$\phi_{NC} = \frac{4 \, y_N \, y_P^c}{\sigma_B^2} \exp \left[- \frac{(y_N^2 + y_P^{c^2})}{\sigma_B^2} \right] I_0 \left(\frac{2 \, \sigma_A \, y_N \, y_P^c}{\sigma_B^2} \right). \tag{3}$$

If $g(y_N, y_P^c)$ is a function of y_N and y_P^c then its expectation value for the truncated data will be given by

$$\langle g(y_N, y_P^c) \rangle_t = \int_{y_N = y_t}^{\infty} \int_{y_P^c = 0}^{\infty} g(y_N, y_P^c) \, P_t(y_N, y_P^c) \, dy_N \, dy_P^c. \tag{4}$$

Substituting the appropriate function $P_t(y_N, y_P^c)$ from (1) in (4) and the relevant expression for g (see table 1) and carrying out resulting integrations the expectation value of any of the functions required for the evaluation of local values of the six R-indices can be calculated. To facilitate numerical evaluation of (4) it is convenient to transform to new variables (u, v) by using the transformation

$$u = y_N/(1 + y_N), \qquad v = y_P^c/(1 + y_P^c), \tag{5}$$

so that (4) can be rewritten as

$$\langle g(y_N, y_P^c) \rangle_t = \int_r^1 \int_0^1 [g(y_N, y_P^c) \, P_t(y_N, y_P^c)]_{u, v} \frac{du \, dv}{(1-u)^2 \, (1-v)^2}, \tag{6}$$

where r is defined to be

$$r = y_t/(1 + y_t) \tag{7}$$

The notation $[g\,(y_N, y_P^c)\,P_t\,(y_N, y_P^c)]_{u,v}$ in (6) implies that the y_N and y_P^c are to be replaced by $u/(1-u)$ and $v/(1-v)$ during the evaluation of the quantity within the parantheses.

2.2 *Derivation of the theoretical expressions for the local values of the R-indices applicable to truncated data*

Some of the results required in this paper have already been derived in PV (1981) and VP (1981) in connection with the normalized R-indices and the essential results required in our derivations are summarized in table 2.

2.2a *Results for the C-case:*

Index $[R(y)]_t$: From the definition of $[R(y)]_t$ in table 1 and the expression for $\langle y_N \rangle_t$ from table 2 we obtain

$$[R(y)]_t = \left(\frac{\pi}{2}\right)^{1/2} \exp\,(y_t^2/2) \int_r^1 \int_0^1 [\,|y_N - \sigma_1 y_P^c|\,\phi_C\,(y_N, y_P^c)\,]_{u,v} \frac{du\,dv}{(1-u)^2\,(1-v)^2} \tag{8}$$

Index $[R(z)]_t$: From the definition of $[R(z)]_t$ in table 1 and the expression for $\langle y_N^2 \rangle^t$ from table 2 we obtain

$$[R(z)]_t = \left[\beta_C + \left(\frac{2}{\pi}\right)^{1/2} y_t \exp\,(-y_t^2/2)\right]^{-1} \int_r^1 \int_0^1 [\,|y_N^2 - \sigma_1^2\,(y_P^{c2})|\,\phi_C\,(y_N, y_P^c)\,]_{u,v} \frac{du\,dv}{(1-u)^2\,(1-v)^2}. \tag{9}$$

Index $[_BR(y)]_t$: From the definition of $[_BR(y)]_t$ in table 1 and the result

$$\langle (y_N - \sigma_1 y_P^c)^2 \rangle_t = \langle y_N^2 \rangle_t - 2\,\sigma_1 \langle y_N y_P^c \rangle_t + \sigma_1^2 \langle (y_P^c)^2 \rangle_t, \tag{10}$$

we obtain

$$[_BR(y)]_t = [\,\langle y_N^2 \rangle_t - 2\,\sigma_1 \langle y_N y_P^c \rangle_t + \sigma_1^2 \langle (y_P^c)^2 \rangle_t]/\langle y_N^2 \rangle_t. \tag{11}$$

Substituting for $\langle y_N^2 \rangle_t$, $\langle y_N y_P^c \rangle_t$ and $\langle (y_P^c)^2 \rangle_t$ from table 2 in (11) and simplifying we obtain

$$[_BR(y)]_t = (\beta_C + A_t)^{-1} \Bigg((1 + \sigma_1^2)\,\beta_C + (1 + \sigma_1^2 \sigma_A^2)\,A_t - \frac{4\sigma_1}{\pi}\Big[\sigma_A \sin^{-1}(\sigma_A) + \sigma_B \{\sigma_A^2 + \sigma_B^2 \exp\,(-y_t^2/2\,\sigma_B^2)\} - (\pi)^{1/2}\,\sigma_A \int_0^{\epsilon} u^{1/2} \exp\,(-u)\,\mathrm{erf}\left(\frac{\sigma_A}{\sigma_B}\,u^{1/2}\right) du\Big]\Bigg), \tag{12}$$

Table 2. Theoretical expressions for the various joint moments of y_N and y_P^c for the C and NC cases.

No.	Moments	C Case	NC Case
1.	$\langle y_N \rangle_t$	G	$y_t + \frac{\pi^{1/2}}{2\beta_{NC}} \operatorname{erfc}(y_t)$
2.	$\langle y_N^2 \rangle_t$	$1 + G\, y_t$	$1 + y_t^2$
3.	$\langle y_N^4 \rangle_t$	$3 + G\, y_t\,(3 + y_t^2)$	$2 + 2y_t^2 + y_t^4$
4.	$\langle (y_P^c)^2 \rangle_t$	$1 + \sigma_A^2\, G\, y_t$	$1 + \sigma_A^2\, y_t^2$
5.	$\langle y_N\, y_P^c \rangle_t$	$\frac{2}{\pi\beta_c}\Big[\sigma_A \sin^{-1}(\sigma_A) + \sigma_A^2\, \sigma_B + \sigma_B^3 \exp(-y_t^2/2\sigma_B^2) - \sqrt{\pi}\,\sigma_A \int_0^{\epsilon} u^{1/2} \exp(-u) \operatorname{erf}\left(\frac{\sigma_A}{\sigma_B} u^{1/2}\right) du\Big]$	$\frac{1}{\beta_{NC}}\Big[E(\sigma_A) - \frac{\sigma_B^2}{2} K(\sigma_A)\Big] - \frac{\sqrt{\pi}}{2\beta_{NC}\,\sigma_B} \int_0^{\delta} u^{1/2} \exp\left\{-\frac{(1+\sigma_B^2)}{2\sigma_B^2} u\right\} \left\{(\sigma_B^2 + \sigma_A^2\, u)\, I_0\left(\frac{\sigma_A^2\, u}{2\sigma_B^2}\right) + \sigma_A^2\, u\, I_1\left(\frac{\sigma_A^2\, u}{2\sigma_B^2}\right)\right\} du$
6.	$\langle (y_P^c)^4 \rangle_t$	$3 + \sigma_A^2\,\{3(1 + \sigma_B^2) + \sigma_A^2\, y_t^2\}\, G\, y_t$	$2 + \sigma_A^2\, y_t^2\,\{2\,(1 + \sigma_B^2) + \sigma_A^2\, y_t^2\}$
7.	$\langle (y_N\, y_P^c)^2 \rangle_t$	$1 + 2\sigma_A^2 + G\, y_t\,(1 + 2\sigma_A^2 + \sigma_A^2\, y_t^2)$	$(1 + \sigma_A^2)\,(1 + y_t^2) + \sigma_A^2\, y_t^4$

Note: For β_C and the expressions for $\langle y_N \rangle_t$, $\langle y_N^2 \rangle_t$, etc. for the C case see (A−10), (A−11), (A−13), (A−15), (A−18), (A−22), (A−23) and (A−24) of PV (1981). For β_{NC} and the expressions for $\langle y_N \rangle_t$; $\langle y_N^2 \rangle_t$, etc. for the NC case see (A−5), (A−7), (A−11), (A−13), (A−14), (A−15), (A−16) and (A−17) of VP (1981). In this table we use the following symbolic notation: $G = \sqrt{2/(\pi\beta_C^2)} \exp(-y_t^2/2)$, $\beta_C = \operatorname{erfc}(y_t/\sqrt{2})$, $\beta_{NC} = \exp(-y_t^2)$ $\epsilon = y_t^2/2$, $\delta = y_t^2$, $\sigma_A = \sigma_1 D$, $\sigma_B = (1 - \sigma_A^2)^{1/2}$, $D = \exp(-\pi^3 \langle |\Delta \mathbf{r}| \rangle^2 H^2/4)$, $H = 2 \sin \theta/\lambda$ and y_t is the truncation limit (*i.e.* the reflections for which $y_N < y_t$ are omitted for evaluating the R-indices).

where we have defined A_t to be

$$A_t = (2/\pi)^{1/2}\, y_t \exp\,(-\, y_t^2/2). \tag{13}$$

Index $[_BR\,(z)]_t$: From the definition of $[_BR\,(z)]_t$ in table 1 and the result that

$$\langle [y_N^2 - (\sigma_1\, y_P^c)^2]^2\rangle_t = \langle y_N^4\rangle_t - 2\,\sigma_1^2\,\langle (y_N\, y_P^c)^2\rangle_t + \sigma_1^4\,\langle (y_P^{c4})\rangle_t, \tag{14}$$

we obtain

$$[_BR\,(z)]_t = [\langle y_N^4\rangle_t - 2\,\sigma_1^2\,\langle (y_N\, y_P^c)^2\rangle_t + \sigma_1^4\,\langle (y_P^{c4})\rangle_t]\,/\,\langle y_N^4\rangle_t. \tag{15}$$

Substituting for $\langle y_N^4\rangle_t$, $\langle (y_N\, y_P^c)^2\rangle_t$ and $\langle (y_P^{c4})\rangle_t$ from table 2 and simplifying we obtain

$$\begin{aligned}[_BR(z)]_t = {} & [3\,\beta_C + A_t\,(3 + y_t^2)]^{-1}\,\big(3\,\beta_C + A_t\,(3 + y_t^2) \\ & - 2\,\sigma_1^2\,\{(1 + 2\,\sigma_A^2)\,(\beta_C + A_t) + \sigma_A^2\, A_t\, y_t^2\} \\ & + \sigma_1^4\,[3\,\beta_C + A_t\,\{3\,\sigma_A^2\,(1 + \sigma_B^2) + \sigma_A^4\, y_t^2\}]\big)\end{aligned} \tag{16}$$

Index $[R^f(y)]_t$: From the definition of $[R^f(y)]_t$ in table 1 we obtain

$$[R^f(y)]_t = \frac{2}{\beta_C}\int_r^1\int_0^1 \left[\left|\frac{y_N - \sigma_1\, y_P^c}{y_N + \sigma_1\, y_P^c}\right|\phi_C\,(y_N, y_P^c)\right]_{u,\,v} \frac{du\, dv}{(1-u)^2\,(1-v)^2}. \tag{17}$$

Index $[R^f(z)]_t$: From the definition of $[R^f(z)]_t$ in table 1 we obtain

$$[R^f\,(z)]_t = \frac{2}{\beta_C}\int_r^1\int_0^1 \left[\left|\frac{y_N^2 - \sigma_1^2\, y_P^{c2}}{y_N^2 + \sigma_1^2\, y_P^{c2}}\right|\phi_C\,(y_N, y_P^c)\right]_{u,\,v} \frac{du\, dv}{(1-u)^2\,(1-v)^2}. \tag{18}$$

2.2b *Results for the NC case:* Since the results for the *NC* case can be obtained by following an analogous procedure we shall give only the final steps. By making use of the theoretical expressions for the local values of *R*-indices given in table 1 and appropriate results in table 2 we can arrive at the following expressions:

$$\begin{aligned}[R\,(y)]_t = {} & \left[y_t\,\beta_{NC} + \frac{\sqrt{\pi}}{2}\,\mathrm{erf}\,(y_t)\right]^{-1} \\ & \int_r^1\int_0^1 [\,|\,y_N - \sigma_1\, y_P^c\,|\,\phi_{NC}\,(y_N, y_P^c)]_{u,\,v}\, \frac{du\, dv}{(1-u)^2\,(1-v)^2},\end{aligned} \tag{19}$$

$$[R(z)]_t = \left[\frac{\exp(y_t^2)}{1+y_t^2}\right] \int_r^1 \int_0^1 [|y_N^2 - \sigma_1^2 (y_P^c)^2| \phi_{NC}(y_N, y_P^c)]_{u,v} \frac{du\, dv}{(1-u)^2 (1-v)^2}, \tag{20}$$

$$[{}_BR(y)]_t = [\beta_{NC}(1+y_t^2)]^{-1} \Big[(1+\sigma_1^2)\beta_{NC} + y_t^2 \beta_{NC}(1+\sigma_1^2 \sigma_A^2)$$

$$-2\sigma_1 \left\{E(\sigma_A) - \frac{\sigma_B^2}{2} K(\sigma_A)\right\} + \sqrt{\pi} \frac{\sigma_1}{\sigma_B} \int_0^\delta u^{1/2}$$

$$\exp\{-u(1+\sigma_B^2)/2\sigma_B^2\} \left\{(\sigma_B^2 + \sigma_A^2 u) I_0\left(\frac{\sigma_A^2 u}{2\sigma_B^2}\right)\right.$$

$$\left. + \sigma_A^2 u I_1 \left(\frac{\sigma_A^2 u}{2\sigma_B^2}\right)\right\} du\Big], \tag{21}$$

$$[{}_BR(z)]_t = (y_t^4 + 2y_t^2 + 2)^{-1} [y_t^4 (\sigma_A^2 \sigma_1^2 - 1)^2 + 2y_t^2 \{\sigma_1^4 (1 - \sigma_B^4)$$

$$- \sigma_1^2 (1+\sigma_A^2) + 1\} + 2\{\sigma_1^4 - \sigma_1^2 (1+\sigma_A^2) + 1\}], \tag{22}$$

$$[R^f(y)]_t = \frac{2}{\beta_{NC}} \int_r^1 \int_0^1 \left[\left|\frac{y_N - \sigma_1 y_P^c}{y_N + \sigma_1 y_P^c}\right| \phi_{NC}(y_N, y_P^c)\right]_{u,v} \frac{du\, dv}{(1-u)^2 (1-v)^2}, \tag{23}$$

$$[R^f(z)]_t = \frac{2}{\beta_{NC}} \int_r^1 \int_0^1 \left[\left|\frac{y_N^2 - \sigma_1^2 (y_P^c)^2}{y_N^2 + \sigma_1^2 (y_P^c)^2}\right| \phi_{NC}(y_N, y_P^c)\right]_{u,v} \frac{du\, dv}{(1-u)^2 (1-v)^2}. \tag{24}$$

3. Discussion of the theoretical results

The theoretical expressions for the local values of the six unnormalized R-indices applicable for truncated data derived in §§2·2a and 2·2b reduce to the results expected for the case of a complete untruncated data in the limit $y_t \to 0$ as expected. A study of these expressions shows that unlike the normalized indices* the unnormalized indices depend on both σ_1 and D explicitly (besides y_t). The overall values of these indices can be calculated from their local values by following a procedure analogous to the one described in PV (1981). Tables of local values of the R-indices as functions of σ_1^2 and D required for this purpose have been obtained for $y_t = 0$, 0·15 and 0·3 by numerical integration from their expressions derived in §§2·2a and 2·2b.**

*These are explicit functions of σ_A and y_t (see PV 1981 and VP 1981).

**These results are, however, not given here to conserve space and may be obtained from the authors on request. It may be noted that these results may be used to test the correctness of trial structures of actual crystals whose data always suffer a truncation due to the unobserved reflections.

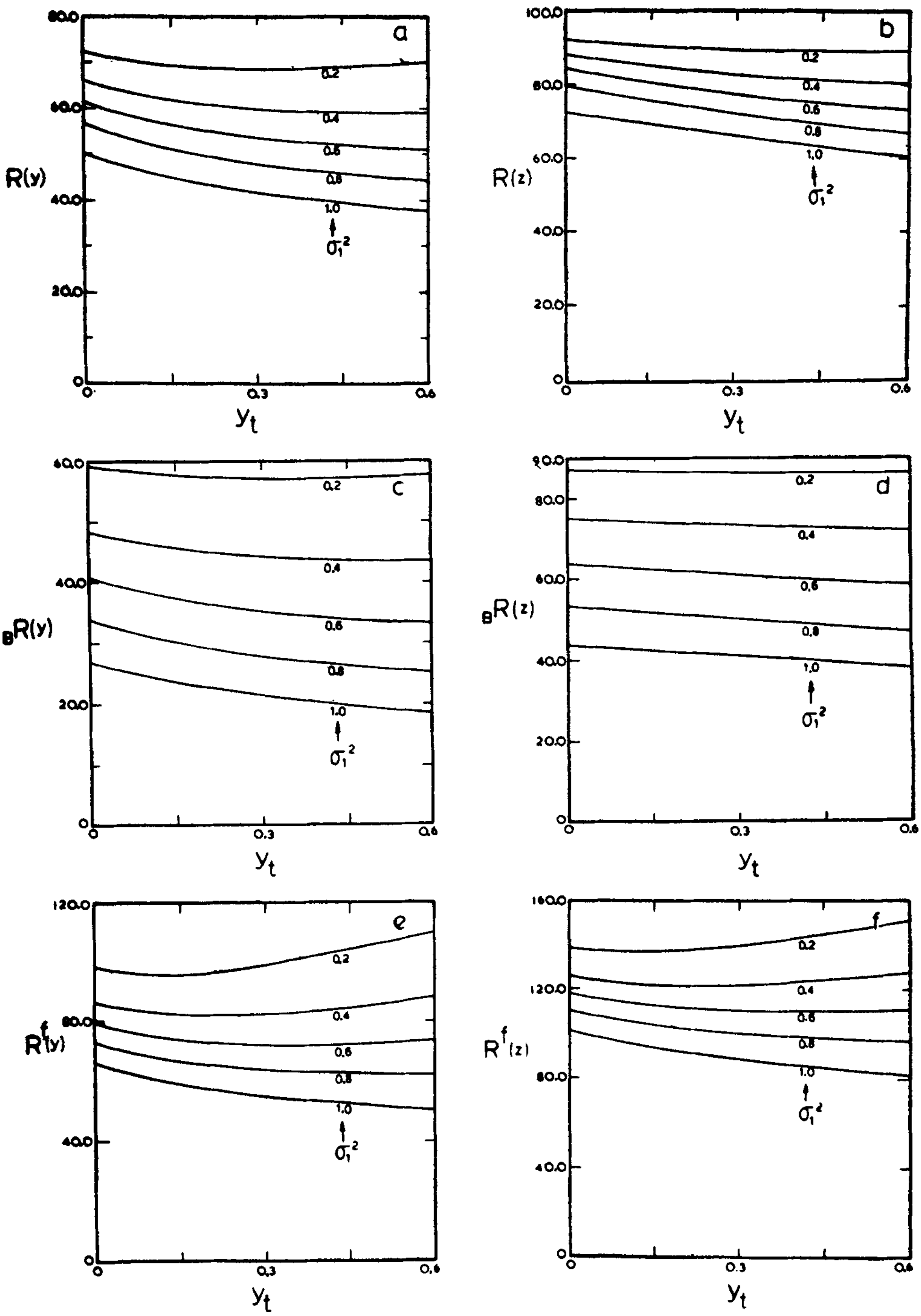

Figure 1. Variation of the local values of the six unnormalized R-indices as a function of y_t corresponding to different fixed values of σ_1^2 for the C case. These curves are drawn for $\langle | \Delta \mathbf{r} | \rangle = 0{\cdot}2Å$ and $\sin\theta/\lambda = 0{\cdot}4Å^{-1}$.

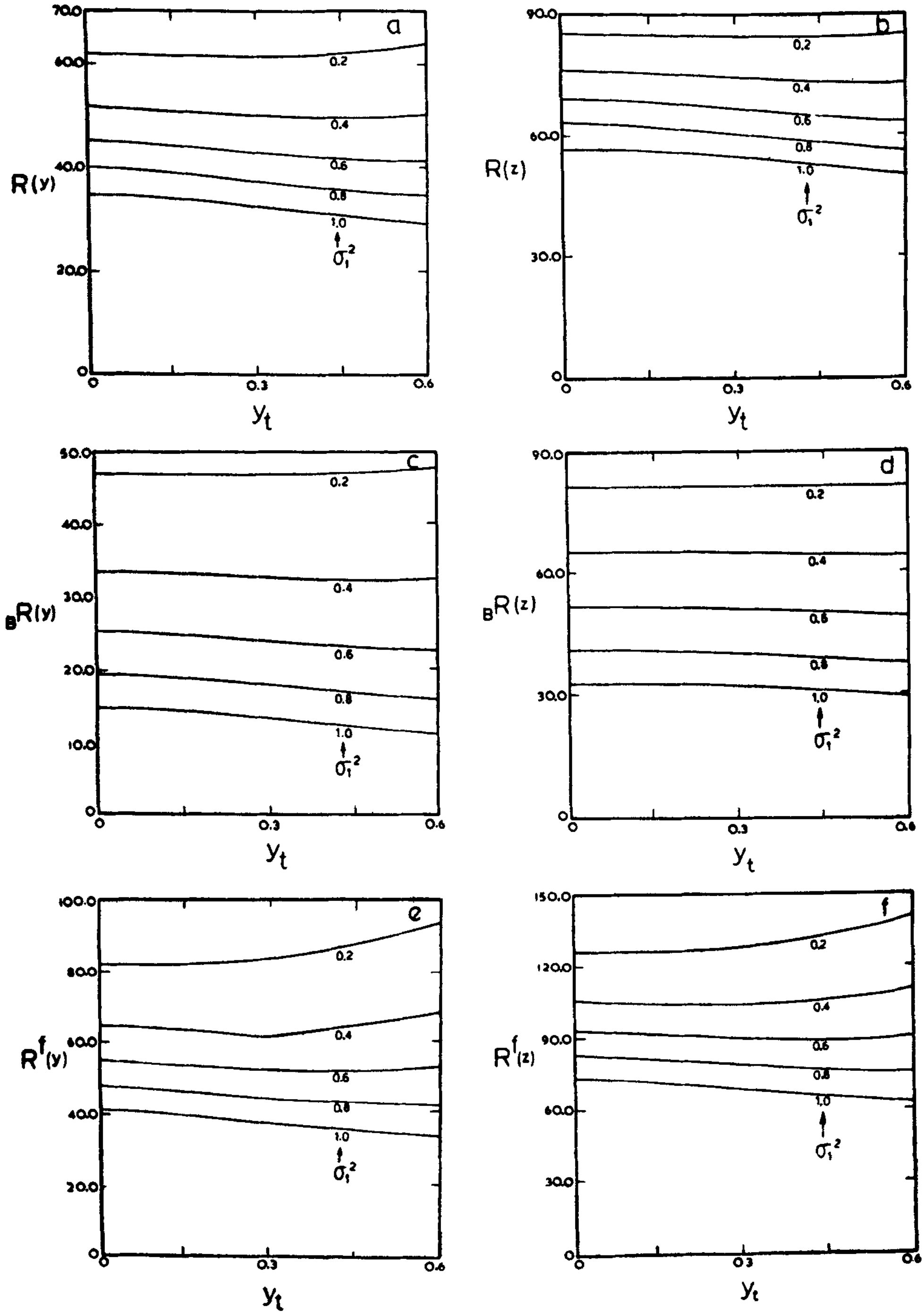

Figure 2. Variation of the local values of the six unnormalized *R*-indices as a function of y_t corresponding to different fixed values of σ_1^2 for the *NC* case. These curves are drawn for $\langle|\Delta \mathbf{r}|\rangle = 0{\cdot}2$Å and $\sin\theta/\lambda = 0{\cdot}4$Å^{-1}.

The variation of the local values of *R*-indices as functions of y_t for different fixed values of σ_1^2 are shown in figures 1 and 2 for the *C* and *NC* cases respectively for the typical values $\langle|\Delta \mathbf{r}|\rangle = 0{\cdot}2$ Å and $\sin\theta/\lambda = 0{\cdot}4$ Å^{-1}. In the *C* case, for the situations of general practical interest (*i.e.* $\sigma_1^2 > 0{\cdot}3$ and $y_t < 0{\cdot}2$), the *R*-indices decrease as y_t increases while these for the *NC* case remain more or less constant (see figures 1 and 2). The fractional type indices increase (for both *C* and *NC* cases) as y_t increases in the region $y_t > 0{\cdot}2$ particularly when σ_1^2 is small (*i.e.* $\sigma_1^2 > 0{\cdot}3$).

References

Parthasarathy S and Ponnuswamy M N 1979 *Acta Crystallogr.* **A35** 672
Parthasarathy S and Velmurugan D 1981 *Acta Crystallogr.* **A37** 472
Velmurugan D and Parthasarathy S 1981 *Acta Crystallogr.* **A37** 480