*Acta Cryst.* (1978). B**34**, 1751–1756

# Variation of Mean Si—O Bond Lengths in Silicon–Oxygen Tetrahedra

BY WERNER H. BAUR

*Department of Geological Sciences, University of Illinois, Chicago, Illinois* 60680, *USA*

(*Received* 10 *August* 1977; *accepted* 20 *December* 1977)

Multiple linear regression analysis was applied to a sample of 314 $SiO_4$ tetrahedra. The mean Si—O bond lengths were treated as the dependent variable. The independent variables which explain most of the variation in $(Si—O)_{mean}$ are: NC, the number of bridging O atoms per tetrahedron; CNM, the mean coordination number of all O atoms within the tetrahedron; NSECM, the mean value of the secant of the bridging angles Si—O—*T*, where non-bridging O atoms are assigned an NSEC value of 2·0. The regression equation $(Si—O)_{mean}$ = 1·615 − 0·0047NC + 0·0054CNM explains 57·6% of the variation in $(Si—O)_{mean}$, with a standard deviation about regression of 0·007 Å. This equation can be used for predictive purposes when $(Si—O)_{mean}$ values are needed for computer simulation of crystal structures. The equation $(Si—O)_{mean}$ = 1 560 + 0·032NSECM + 0·0031CNM can be used when the structure has been determined and the Si—O—*T* angles are known. It explains 66% of the variation in $(Si—O)_{mean}$ with a standard deviation of 0·007 Å. The electronegativity of the cations in the structure and the distortion indices of the tetrahedra did not contribute significantly to the regression sum of squares. The dependence of $(Si—O)_{mean}$ on CNM is not as large as has been found previously in univariate linear regression.

## Introduction

In recent years the variation of mean bond lengths in tetrahedral anions has received some attention: Smith & Bailey (1963) investigated the effects of tetrahedral linkage on Al—O and Si—O distances; Brown & Gibbs (1969) showed that the mean bond length Si—O is a function of the mean coordination number of the O atoms; Shannon (1971), Shannon & Calvo (1973) and Shannon (1976*a,b*) studied the dependence of mean tetrahedral bond lengths in $BO_4$, $SiO_4$, $GeO_4$, $PO_4$, $AsO_4$, $VO_4$ and $SeO_4$ on the mean electronegativity of the cations in the structure; Baur (1974) gave a regression equation for $PO_4$ in which the mean coordination number of the O atoms and the distortion index of the tetrahedral group were used to estimate the mean bond length P—O. Nobody seems to have investigated by multiple linear regression analysis the influence of all these factors together on a large sample of mean tetrahedral distances of a given element.

## Data

The mean Si—O bond lengths of 314 silicate tetrahedra observed in 155 crystal structures determined by X-ray or neutron diffraction were collected. The means of the estimated standard deviations of the Si—O bond lengths in any single structure included here are not larger than 0·010 Å. The mean of the standard deviations of all the bond lengths used is 0·005 Å. All bond lengths and estimated standard deviations have been recalculated from the data in the original papers. Discrepancies were checked by contacting the authors concerned. Only bond lengths not corrected for thermal motion were used.

For each silicate tetrahedron the following data were obtained:

(1) NC, the number of bridging O atoms within the tetrahedron. An O atom is considered to be bridging if it is shared with another tetrahedral Al, B, Ga, P or Si atom. If the neighboring tetrahedrally coordinated element is Be, Li, Mg, or Zn it is not considered a bridging atom.

(2) CNM, the mean coordination number of all O atoms in a given tetrahedron. If an O atom is the acceptor of a hydrogen bond this is counted as a coordinating contact (Baur, 1970, 1974).

(3) SECM, the mean value per $SiO_4$ tetrahedron of the negative secant of the angle Si—O—*T* (Gibbs, Hamil, Louisnathan, Bartell & Yow, 1972) where *T* is Al, B, Ga, P or Si. Non-bridging O atoms do not have a defined value of Si—O—*T* and their SECM is not included in calculating the average (or any other statistic).

(4) NSECM, the mean value per $SiO_4$ tetrahedron of the negative secant of the angle Si—O—*T*, where all non-bridging O atoms have been arbitrarily assigned an NSEC value of 2·0. For an orthosilicate the value of NSECM is therefore 2·0. For four-connected tetrahedra NSECM is identical with SECM; for NC = 1, 2, and 3 the average NSECM values are intermediate between 2·0 and the corresponding SECM value (Table 1).

Table 1. *Simple sample statistics of the mean* Si—O *distances used for the regressions, sorted according to the degree of polymerization of the tetrahedra (NC)*

Estimated standard deviations, here as elsewhere in this paper, are in parentheses following the value and are in units of the least significant digits of the value.

| Number of observations | NC | $(Si–O)_{mean}$ | Minimum $(Si–O)_{mean}$ | Maximum $(Si–O)_{mean}$ | $NSECM_{mean}$ | $CNM_{mean}$ | $SECM_{mean}$ | $EN_{mean}$ | $DITO_{mean}$ |
|---|---|---|---|---|---|---|---|---|---|
| 314 | 0–4 | 1·623 (11) Å | 1·584 Å | 1·654 Å | 1·64 (25) | 3·26 (61) | 1·33 (17) | 1·64 (14) | 0·00016 (16) |
| 50 | 0 | 1·636 (8) | 1·622 | 1·654 | 2·00 (0) | 3·97 (62) | – | 1·59 (16) | 0·00006 (11) |
| 37 | 1 | 1·627 (9) | 1·610 | 1·646 | 1·82 (6) | 3·36 (46) | 1·27 (25) | 1·58 (13) | 0·00019 (16) |
| 108 | 2 | 1·626 (7) | 1·607 | 1·649 | 1·68 (7) | 3·34 (32) | 1·37 (14) | 1·67 (11) | 0·00024 (14) |
| 53 | 3 | 1·617 (6) | 1·601 | 1·632 | 1·49 (12) | 3·09 (42) | 1·32 (16) | 1·62 (10) | 0·00018 (19) |
| 66 | 4 | 1·610 (9) | 1·584 | 1·629 | 1·31 (15) | 2·70 (56) | 1·31 (15) | 1·70 (15) | 0·00007 (7) |

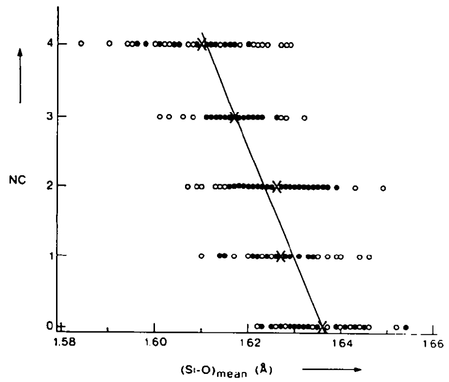


Fig. 1. Scatter plot of observed $(Si—O)_{mean}$ against NC. The line is the simple-linear-regression line of $(Si—O)_{mean}$ on NC [slope = −0·0063 (3)]. Open circles: single points; solid circles: multiple points; crosses: average for NC group.

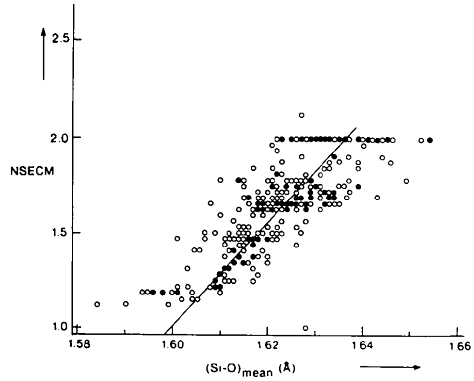


Fig. 3. Scatter plot of observed $(Si—O)_{mean}$ against NSECM. The regression line has a slope of 0·037 (2).

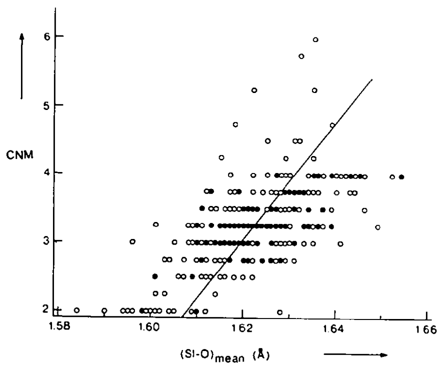


Fig. 2. Scatter plot of observed $(Si—O)_{mean}$ against CNM. The regression line [slope = 0·0119 (8)] results from simple linear regression of $(Si—O)_{mean}$ on CNM.

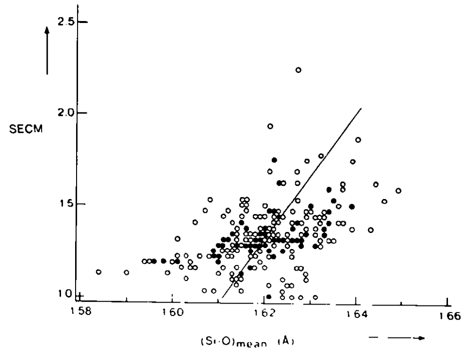


Fig. 4. Scatter plot of observed $(Si—O)_{mean}$ against SECM. The slope of the regression line is 0·027 (3).

(5) DITO, the distortion index of the tetrahedron as defined by Brown & Shannon (1973).

(6) EN, the mean electronegativity of all the cations in a given structure as defined by Shannon (1971). The individual electronegativity values themselves were taken from Allred (1961).

The observed mean Si—O distances in the sample have a range of 0·07 Å (Table 1). If one breaks the sample down according to the linkage of the silicate tetrahedra the range within each group is about half as large. However, all the groups overlap: tetrahedra with mean Si—O values between 1·622 and 1·629 Å could be orthosilicates or could equally well be four-connected tetrahedra (Fig. 1). The data are displayed in Figs. 1 to 4, except for DITO and EN which contribute to the regressions only in a minor way.*

* A list of compounds, minerals and their references has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33323 (6 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

### Calculations

The correlation-coefficient matrix (Table 2) shows that of the 21 non-trivial possible correlation coefficients no less than 17 are significant. The significance probabilities are based on the $t$ test and give us the probability that a value of the correlation coefficient ($r$) as large or larger in absolute value than the one computed would have been observed fortuitously even though the two variables are not in fact correlated. Therefore we can say that we have more than a 99% chance that all but the correlations between the pairs NSECM with DITO, SECM with NC, DITO with NC, and SECM with EN are significant. That a correlation is significant may not necessarily be helpful in establishing a regression equation of practical value, since the per cent variation explained (percentage of the sum of squares due to regression) is proportional to the square of the correlation coefficient. This means that NSECM explains 64·5% of the variation observed in $(Si-O)_{mean}$, NC explains 52%, while DITO explains only 3% of this variation. Further tests showed that multiple regressions in which $(Si-O)_{mean}$ was the dependent variable, while NC and CNM were the independent variables, explained 57·6% of the variation in $(Si-O)_{mean}$. Adding to that a third independent variable explained 58·3% if DITO was added, 58·6% if EN was added and 62·8% if SECM was added (in this latter case for 264 observations only).

By itself NSECM explains more of the variation than NC, CNM and SECM combined. There is no point in adding NC to the multiple-regression model when NSECM is present because the correlation between NC and NSECM is very high (−0·92, see Table 2). Actually, NSECM combines within one variable the attributes of both NC and SECM. It splits the sample population into groups, as does NC, and allows for a gradation within the groups according to SECM. Since NSECM and CNM are less correlated than NSECM and NC, it pays to include CNM with NSECM in a regression model. This raises $r^2$, the per cent variation explained, to 66%.

The four best regression models based on these considerations are listed as models 1 to 4 in Table 3. All independent variables in these models have a very high probability of being statistically significant. However, the goodness of fit for these models is almost constant. The standard deviation between observed and estimated values of $(Si-O)_{mean}$ in all four models is 0·007 Å. The differences between these standard deviations are in the fourth decimal place.

A study of regression models for the individual groups in NC was less fruitful. None of the independent variables considered here contribute significantly to the variation of mean Si—O bond lengths in orthosilicates (NC = 0). The $r^2$ values for the groups with NC = 1, 2 and 3 are all below 0·40. Less than 40% of the variation within these groups can be explained by NSEC, NC, CNM, SECM, EN and DITO. The significance probabilities of CNM in models 6 and 7 (Table 3) are not as high as one would wish. Only for NC = 4 can more than half of the variation be explained by a model including NSECM and CNM, and in this case the CNM variable is acceptably significant. In none of the best models does EN contribute in a significant manner to the regression sum of squares. The correlation coefficient between EN and $(Si-O)_m$ was also found by Shannon (1976*a*) to be small (0·40).

All statistical calculations were performed on HP-65 and IBM 370/158 computers (using SAS76; Barr, Goodnight, Sall & Helwig, 1976).

Table 2. *Correlation coefficients r (in upper row) between the variables considered here*

All coefficients involving SECM are based on 264 observations, since SECM is not defined for orthosilicates. The other coefficients are based on the complete set of 314 observations. In the lower row, the significance probability is listed (see text). A significance probability of 0·0001 means that it is 0·0001 or less.

| | $(Si-O)_{mean}$ | NSECM | NC | CNM | SECM | EN | DITO |
|---|---|---|---|---|---|---|---|
| $(Si-O)_{mean}$ | 1·0 | 0·80 | −0·72 | 0·63 | 0·45 | −0·20 | 0·16 |
| | 0·0 | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0003 | 0·0046 |
| NSECM | | 1·0 | −0·92 | 0·67 | 0·48 | −0·26 | 0·12 |
| | | 0·0 | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·038 |
| NC | | | 1·0 | −0·63 | −0·03 | 0·24 | −0·03 |
| | | | 0·0 | 0·0001 | 0·65 | 0·0001 | 0·58 |
| CNM | | | | 1·0 | 0·32 | −0·54 | 0·21 |
| | | | | 0·0 | 0·0001 | 0·0001 | 0·0002 |
| SECM | | | | | 1·0 | 0·004 | 0·26 |
| | | | | | 0·0 | 0·95 | 0·0001 |
| EN | | | | | | 1·0 | −0·18 |
| | | | | | | 0·0 | 0·0011 |
| DITO | | | | | | | 1·0 |
| | | | | | | | 0·0 |

Table 3. *Multiple-regression equations for* (Si—O)$_{mean}$ *for different samples: correlation coefficient squared* ($r^2$); *significance probability (s.p.), where* 0·0001 *means that it is* 0·0001 *or less; slopes for different independent variables* $b_1$ *(for NSECM)*, $b_2$ *(for NC)*, $b_3$ *(for CNM)*, $b_4$ *(for SECM) and* $b_5$ *(for DITO); standard deviation about regression, which is the square root of the mean square error*

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| NC | 0 to 4 | 1 to 4 | 0 to 4 | 0 to 4 | 1 | 2 | 3 | 4 |
| Sample size | 314 | 264 | 314 | 314 | 37 | 108 | 53 | 66 |
| $r^2$ | 0·576 | 0·628 | 0·660 | 0·645 | 0·383 | 0·342 | 0·317 | 0·516 |
| $a_1$ intercept | 1·615 (4) | 1·591 (4) | 1·560 (3) | 1·562 (3) | 1·474 (33) | 1·545 (14) | 1·571 (10) | 1·554 (7) |
| s.p. ($a$) | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0001 |
| $b_1$ (NSECM) | – | – | 0·032 (2) | 0·037 (2) | 0·084 (18) | 0·038 (9) | 0·026 (6) | 0·034 (6) |
| s.p. ($b_1$) | – | – | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0001 | 0·0001 |
| $b_2$ (NC) | −0·0047 (4) | −0·0053 (4) | – | – | – | – | – | – |
| s.p. ($b_2$) | 0·0001 | 0·0001 | – | – | – | – | – | – |
| $b_3$ (CNM) | 0·0054 (9) | 0·0043 (10) | 0·0031 (8) | – | – | 0·0037 (20) | 0·0033 (19) | 0·0045 (16) |
| s.p. ($b_3$) | 0·0001 | 0·0001 | 0·0002 | – | – | 0·063 | 0·084 | 0·006 |
| $b_4$ (SECM) | – | 0·022 (2) | – | – | – | – | – | – |
| s.p. ($b_4$) | – | 0·001 | – | – | – | – | – | – |
| $b_5$ (DITO) | – | – | – | – | – | 14·2 (4·3) | −12·3 (4·3) | – |
| s.p. ($b_5$) | – | – | – | – | – | 0·0012 | 0·007 | – |
| Standard deviation | 0·007 Å | 0·007 Å | 0·007 Å | 0·007 Å | 0·007 Å | 0·006 Å | 0·005 Å | 0·006 Å |

## Discussion

Each of the columns of Table 3 corresponds to a possible equation for estimating the mean tetrahedral Si—O distance under different conditions and for different types of $SiO_4$ groups. The first of these equations [(Si—O)$_{mean}$ = 1·615 − 0·0047NC + 0·0054CNM] has the advantage of being predictive. For a given crystal structure NC and CNM can be determined even if the structure is only imprecisely known. They can also be determined if one is dealing with a hypothetical structure which one wishes to simulate on a computer (Baur, 1977*a*). In this case a predictive equation is of practical value. The better the available predicted bond-length values are, the more reliable is the computer simulation of crystal structures.

The slope of (Si—O)$_{mean}$ with CNM is 0·0054 (9). This value is statistically identical with the slope of 0·0047 (9) obtained for (P—O)$_{mean}$ with CNM (Baur, 1974). It is significantly different from the slope of 0·015 obtained for the tetrahedral (Si—O)$_{mean}$ with CNM (Brown & Gibbs, 1969), from the slope of 0·013 found for the octahedral (Si—O)$_{mean}$ (Baur, 1977*b*) and from the slope (0·012) which can be read from Fig. 5 of Shannon & Prewitt (1969) and which applies not only to silicates but to a wide variety of cation–oxygen bonds. The univariate slope entered in Fig. 2 also has a value of 0·012. The difference between these situations of course is that regressions on only one variable will result in a steeper slope when compared with multiple regressions involving other independent variables which in fact are not truly independent but instead correlated with each other (the $r$ between NC and CNM is −0·63, Table 2). Since SECM is not strongly correlated with NC and CNM the addition of SECM into the multiple-regression model [equation (2), Table 3] does not affect

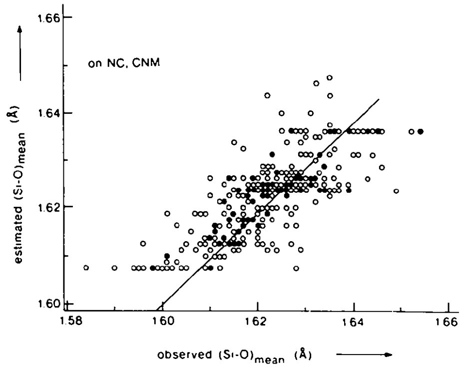


Fig. 5. Scatter plot of observed (Si—O)$_{mean}$ against estimated (Si—O)$_{mean}$ based on equation (1). The line is the regression line with a slope of 1·00 (5).

the slopes of NC and CNM in a significant way. The slope of NC is −0·047 in equation (1), significantly different from the univariate slope [−0·063 (3)] of (Si—O)$_{mean}$ with NC (Fig. 2). Strangely enough, however, it is identical with the slope (−0·048) which can be read from Fig. 1 of Smith & Bailey (1963).

The scatter in a plot of observed (Si—O)$_{mean}$ *versus* estimated (Si—O)$_{mean}$ is appreciable (Fig. 5), but the standard deviation of regression [(*A*), Table 4] is only 0·007 Å. This compares favorably with the mean precision of all bond lengths involved (0·005 Å). The extreme differences between observation and estimate range from −0·02 up to 0·03 Å, as they do for all the other regression equations considered here (Table 4). Some of these deviations might be due to systematic errors in the observations. Errors in unit-cell-length determinations or overlooked partial occupancies of the tetrahedral site by Al could cause this.

Table 4. *Standard deviations about regression and minimum and maximum Δ values between observed* $(Si—O)_{mean}$ *and estimated* $(Si—O)_{mean}$

| | | (*A*) Estimate based on equation (1)* | | | (*B*) Estimate based on equation (3)* | | | (*C*) Estimates based on equations (5), (6), (7) and (8)* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| NC | Sample size | Standard deviation | $\Delta_{min}$ | $\Delta_{max}$ | Standard deviation | $\Delta_{min}$ | $\Delta_{max}$ | Standard deviation | $\Delta_{min}$ | $\Delta_{max}$ |
| 0–4† | 314 | 0·007 Å | −0·023 Å | 0·026 Å | 0·007 Å | −0·018 Å | 0·030 Å | 0·006 Å | −0·017 Å | 0·031 Å |
| 0 | 50 | 0·008 | −0·021 | 0·017 | 0·008 | −0·018 | 0·018 | – | – | – |
| 1 | 37 | 0·008 | −0·018 | 0·017 | 0·007 | −0·016 | 0·015 | 0·007 | −0·015 | 0·014 |
| 2 | 108 | 0·007 | −0·012 | 0·026 | 0·006 | −0·014 | 0·022 | 0·006 | −0·017 | 0·019 |
| 3 | 53 | 0·006 | −0·017 | 0·012 | 0·005 | −0·016 | 0·011 | 0·005 | −0·015 | 0·009 |
| 4 | 66 | 0·008 | −0·023 | 0·021 | 0·006 | −0·018 | 0·530 | 0·006 | −0·017 | 0·031 |

* Table 3.
† The estimate under (*C*) is only for NC 1 to 4; the number of observations is therefore 264.

Equation (3) (Table 3) can be used for 'post'-dictive purposes: $(Si—O)_{mean} = 1·560 + 0·032NSECM + 0·0031CNM$. This means that once a crystal structure is determined and refined and the angles Si—O—*T* are known, one can compare the observed with the estimated mean Si—O bond lengths. The overall agreement between observation and estimation is improved for this regression model as compared with the equation (1) model [Fig. 6, and (*B*) Table 4]. However, there is one datum which lies outside the range of the others. It is the zunyite (Louisnathan & Gibbs, 1972) point with an observed $(Si—O)_{mean}$ of 1·628 Å, and an estimated $(Si—O)_{mean}$ of 1·598 Å. A possible explanation for this is provided by the notion that the lower limit of the size of the Si—O—*T* angles is determined by the non-bonding interactions between the tetrahedral cations (Glidewell, 1975, 1977*a*,*b*; O'Keeffe & Hyde, 1976, 1978; Baur, 1977*c*). This does not preclude that for a given Si—O bond length the Si—O—*T* angle is wider than would be allowed by the non-bonded interactions of the Si atoms with each other. The $SiO_4$ tetrahedron in zunyite seems to represent just such a case. In a regression calculation involving *individual* Si—O bond lengths (Baur, 1977*c*) the same four-connected $SiO_4$ tetrahedron in zunyite joins several other data points in lying off the trend established by the lower limit of the Si—O—*T* angles. This means that the correlation between Si—O distances and Si—O—*T* at the lower limit of these angles is due to the geometrically simple dependence of the angle Si—O—*T* on the Si—O and *T*—O distances and not *vice versa* as required by π-bonding theory (Cruickshank, 1961).

The physical significance of the dependence of $(Si—O)_m$ on NSECM and SECM seems to be due to this geometric relationship between the distances Si—O, Si—Si and the angle Si—O—Si. The non-bonded Si···Si distances are at least as constant as the bonded Si—O distances (O'Keeffe & Hyde, 1978); therefore the angles Si—O—Si have to be correlated with the Si—O distances (and with the Si—Si distances).

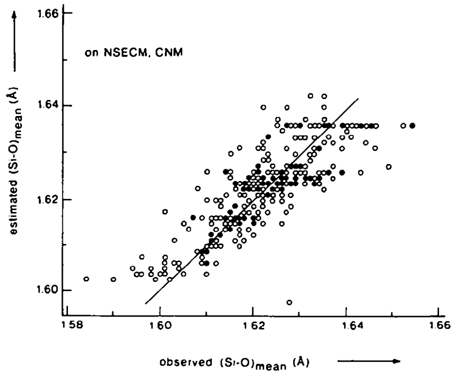


Fig. 6. Scatter plot of observed $(Si—O)_{mean}$ against estimated $(Si—O)_{mean}$ using equation (3). Regression of the estimated on the observed distances results in a line with a slope of 1·00 (4).

Equations (5) to (8) (Table 3) improve somewhat the fit between observations and estimates. However, they are of questionable value since $r^2$ for each of them is smaller than for either equations (1) or (3), the CNM contributions to equations (6) and (7) are not of sufficiently high significance, while the DITO contributions are of opposite sign, and last but not least they require a larger number of fitted parameters. Equations (1) and (3) have only three fitted parameters each. Even with equation (8) the zunyite datum is not satisfactorily accounted for.

The zunyite datum is an example of a poor fit. Ussingite and low albite are examples of a slightly better than average fit between observation and estimate (Table 5). It is remarkable that the three-parameter equation (1), which applies to a wide variety of silicates, gives a better agreement with observation than the four-parameter equation derived from albite and applied to ussingite by Ribbe (1974). Ribbe's equation presupposes a detailed knowledge of Si—O—*T* angles and Na—O distances. Equation (1) relies on NC and CNM only.

Based on the $(Si—O)_m$ *versus* CNM relationship established here the radius of $Si^{4+}$ in tetrahedral

Table 5. *Observed mean* Si—O *bond lengths in ussingite,* $Na_2AlSi_3O_8(OH)$, *and albite,* $NaAlSi_3O_8$, *and estimates based on equation* (1) [$(Si—O)_{e1}$], *equation* (3) [$(Si—O)_{e3}$], *equations* (7) *and* (8) [$(Si—O)_{e78}$] *and on Ribbe's* (1974) *equation* [$(Si—O)_{er}$]

| | | $(Si—O)_{obs}$ | $(Si—O)_{e1}$ | $(Si—O)_{e3}$ | $(Si—O)_{e78}$ | $(Si—O)_{er}$ |
|---|---|---|---|---|---|---|
| Ussingite | $Si_2$ | 1·620 Å | 1·618 Å | 1·619 Å | 1·619 Å | 1·616 Å |
| | $Si_3$ | 1·620 | 1·612 | 1·612 | 1·613 | 1·608 |
| | $Si_4$ | 1·622 | 1·618 | 1·619 | 1·618 | 1·620 |
| Albite | Si1*m* | 1·609 | 1·614 | 1·609 | 1·610 | 1·609 |
| | Si2*o* | 1·613 | 1·612 | 1·612 | 1·613 | 1·614 |
| | Si2*m* | 1·615 | 1·612 | 1·614 | 1·615 | 1·615 |
| | Standard deviation about regression | | 0·006 | 0·004 | 0·004 | 0·006 |

coordination with O is 0·257 Å assuming that the O radii are as determined by Shannon & Prewitt (1969). This is essentially identical with the radius of 0·26 Å reported both by Shannon & Prewitt (1969) and by Shannon (1976*b*). The small difference is within the tolerance assumed by these authors. It is also a small difference compared with the standard deviation between observed and estimated Si—O distances observed in this work.

The arbitrary assumption that NSECM for the orthosilicates has the value of 2·0 can be justified by the success of the model. It is also interesting to note that the mean Si—O distances in orthosilicates have values similar to those found in polymerized silicate tetrahedra, when the Si—O—*T* angles are narrow and consequently SECM has a high value, sometimes ranging up to 2·0, in one case even exceeding this value.

## Conclusion

Equation (1) established here allows the mean Si—O distances of 314 silicate tetrahedra to be estimated within 0·007 Å. This equation can be used for the prediction of mean Si—O distances needed in the computer simulation of crystal structures (see, for example, the simulation of the superstructure of low tridymite; Dollase & Baur, 1976). It is disappointing, however, that the model is not able to estimate in detail the variations in the bond lengths of orthosilicates, even though their observed values range from 1·622 to 1·654 Å. It is also obvious that two of the factors which were isolated in previous work on other tetrahedral groups, namely the distortion index and the mean electronegativity, are not responsible for much of the mean-bond-length variation observed for this sample of 314 $SiO_4$ tetrahedra.

I thank the Computer Center of the University of Illinois for computer time.

## References


ALLRED, A. L. (1961). *J. Inorg. Nucl. Chem.* **17**, 215–221.

BARR, A. J., GOODNIGHT, J. H., SALL, J. P. & HELWIG, J. T. (1976). *A User's Guide to SAS76*. Raleigh: SAS Institute.

BAUR, W. H. (1970). *Trans. Am. Crystallogr. Assoc.* **6**, 129–155.

BAUR, W. H. (1974). *Acta Cryst.* B**30**, 1195–1215.

BAUR, W. H. (1977*a*). *Phys. Chem. Mineral.* **2**, 3–20.

BAUR, W. H. (1977*b*). *J. Solid State Chem.* **22**, 445–446.

BAUR, W. H. (1977*c*). *Acta Cryst.* B**33**, 2615–2619.

BROWN, G. E. & GIBBS, G. V. (1969). *Am. Mineral.* **54**, 1528–1539.

BROWN, I. D. & SHANNON, R. D. (1973). *Acta Cryst.* A**29**, 266–282.

CRUICKSHANK, D. W. J. (1961). *J. Chem. Soc.* pp. 5486–5504.

DOLLASE, W. A. & BAUR, W. H. (1976). *Am. Mineral.* **61**, 971–978.

GIBBS, G. V., HAMIL, M. M., LOUISNATHAN, S. J., BARTELL, L. S. & YOW, H. (1972). *Am. Mineral.* **57**, 1578–1613.

GLIDEWELL, C. (1975). *Inorg. Chim. Acta*, **12**, 219–227.

GLIDEWELL, C. (1977*a*). *Inorg. Nucl. Chem. Lett.* **13**, 65–68.

GLIDEWELL, C. (1977*b*). *Inorg. Chim. Acta*, **25**, 77–90.

LOUISNATHAN, S. J. & GIBBS, G. V. (1972). *Am. Mineral.* **57**, 1089–1108.

O'KEEFFE, M. & HYDE, B. G. (1976). *Acta Cryst.* B**32**, 2923–2936.

O'KEEFFE, M. & HYDE, B. G. (1978). *Acta Cryst.* B**34**, 27–32.

RIBBE, P. H. (1974). *Am. Mineral.* **59**, 341–344.

SHANNON, R. D. (1971). *Chem. Commun.* pp. 881–882.

SHANNON, R. D. (1976*a*). *Physics and Chemistry of Minerals and Rocks*, edited by R. G. J. STRENS, pp. 403–431. London: John Wiley.

SHANNON, R. D. (1976*b*). *Acta Cryst.* A**32**, 751–767.

SHANNON, R. D. & CALVO, C. (1973). *J. Solid State Chem.* **6**, 538–549.

SHANNON, R. D. & PREWITT, C. T. (1969). *Acta Cryst.* B**25**, 925–946.

SMITH, J. V. & BAILEY, S. W. (1963). *Acta Cryst.* **16**, 801–811.