

# A simplified empirical model for approximation of the 'bond valence–bond length' correlation for H—O bonds

Vasyl Sidey

Department of Chemistry and Research Institute  
for Physics and Chemistry of Solids, Uzhgorod  
National University, Pidgirna Street 46,  
Uzhgorod 88000, Ukraine

Correspondence e-mail: vasylsidey@yahoo.com

Received 13 February 2011

Accepted 19 March 2011

The 'bond valence ( $s$ )–bond length ( $r$ )' correlation reported earlier for H—O bonds [Brown (2009). *Chem. Rev.* **109**, 6858–6919] has been closely approximated using the function  $s = (\alpha - \beta r)^{1/\gamma}$ , where  $\alpha = 1.185$  (10),  $\beta = 0.321$  (8) and  $\gamma = 0.36$ .

Based on the crystal structures accurately determined by neutron diffraction, Professor I. D. Brown has recently presented the experimentally determined 'bond valence–bond length' plot for H—O bonds (Brown, 2002, 2009). In the modern bond-valence model, the most commonly adopted empirical expressions for the relationship between the bond valences  $s$  and the bond lengths  $r$  are (1) and (2), where  $r_0$ ,  $n$  and  $b$  are the empirically determined constants (bond-valence parameters) for a given type of  $A-X$  chemical bonds (Brown, 2002, 2009).

$$s = (r_0/r)^n \quad (1)$$

$$s = \exp[(r_0 - r)/b] \quad (2)$$

In most cases the performances of (1) and (2) in approximating the 'bond valence–bond length' correlations are reasonably high; however, the 'bond valence–bond length' plot determined by Brown (2002, 2009) for H—O bonds clearly cannot be closely approximated using (1) or (2). In order to overcome this problem, Brown (2002) has recommended the use of (2) with *three* different sets of the  $r_0$  and  $b$  parameters for different  $r$  ranges:  $r_0 = 0.907$  and  $b = 0.28$  Å for  $r \leq 1.05$  Å;  $r_0 = 0.569$  and  $b = 0.94$  Å for  $1.05 \leq r \leq 1.70$  Å;  $r_0 = 0.990$  and  $b = 0.59$  Å for  $1.70 \leq r$ .

The curve calculated from the above sets of the bond-valence parameters is shown in Fig. 1, along with the experimental datapoints recovered by the author (see below). One can see that the Brown curve has the 'bulge' in the centre and, according to Brown (2002), this 'bulge' is an artifact of anion–anion repulsion and is not intrinsic to the H—O bond itself. Unfortunately, owing to significant scatter of the datapoints observed in Fig. 1, the true shape of the 'bond valence–bond length' correlation for H—O bonds remains unclear; so existence of the aforementioned 'bulge' is, strictly speaking, questionable. Furthermore, in the short-bond range the Brown curve seems to be too steep and does not fit the short bonds very well.

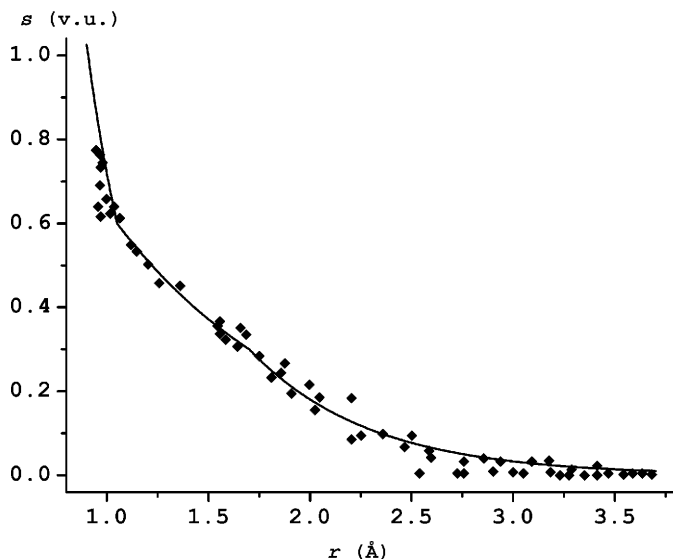
For routine bond-valence analyses of the structures containing H—O bonds, the  $r_0$  and  $b$  parameters recommended by Brown (2002) are undoubtedly sufficient; however, the performance and success of any theory predicting some observed  $y = f(x)$  correlation can be properly examined and illustrated only in comparison with the *reference empirical model* that:

(i) demonstrates reasonably close approximation of the experimental data,

(ii) is uniformly defined over the *whole* observed  $x$  and  $y$  ranges, and

(iii) is as simple as possible.

Thus, the present work is an attempt of the author to find such a simple empirical model for the 'bond valence–bond length' correlation for H—O bonds. In order to obtain unbiased results, in solving



**Figure 1**  
Approximation of the ‘bond valence–bond length’ correlation for H–O bonds, according to Brown (2002).

the problem there were no presumptions about the properties of hydrogen bonding in solids<sup>1</sup> and therefore the experimental data collected and reported by Brown (2002, 2009) were studied exclusively.

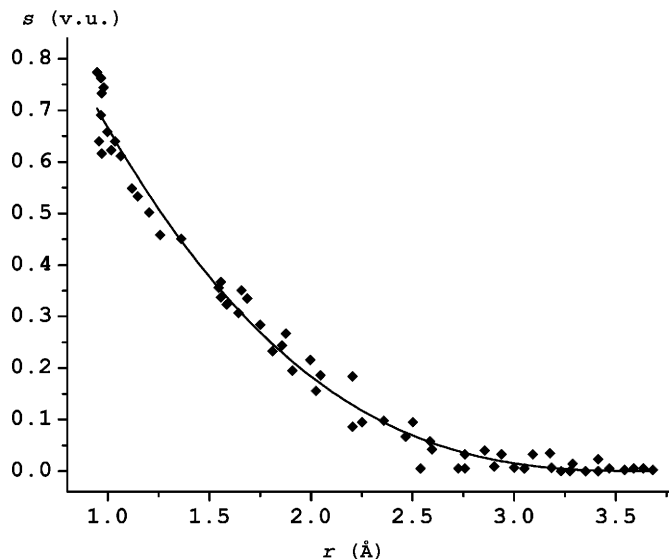
As the original numerical ‘bond valence–bond length’ data for H–O bonds are lost (Brown, 2010), these data were recovered from Fig. 19 of the recent review by Brown (2009) using a graph-digitizing procedure (Mitchell, 2002). The maximum errors that could be introduced by the above graph-digitizing procedure have been estimated to be  $\sim 0.005$  valence units (v.u.) for  $s$  and  $\sim 0.01$  Å for  $r$ ; the total influence of errors on the final results was considered to be statistically insignificant because the number of datapoints is reasonably high ( $N = 62$ ) and the errors introduced are expected to be random.

In order to find a suitable empirical model for approximation of the ‘bond valence–bond length’ correlation for H–O bonds,  $\sim 100$  different functions (exponential, logarithmic, power-mode, polynomial *etc.*) commonly used for nonlinear regression analysis (*e.g.* Seber & Wild, 1989; Rawlings *et al.*, 1998) have been tested by using a common plotting/curve-fitting program. After the preliminary curve-fitting procedures and visual inspections of the curves obtained, most of the above functions had to be rejected owing to the unacceptably poor approximation of the experimental data and/or because of negative (*i.e.* physically impossible)  $s$  values yielded from the refined models for longer H–O bonds.

In further calculations, the *adjusted coefficient of determination*  $R^2_{\text{adjusted}}$  (Rawlings *et al.*, 1998) was considered to be the primary statistical indicator for goodness-of-fit: unlike the ‘unadjusted’ *coefficient of determination*  $R^2$ ,  $R^2_{\text{adjusted}}$  cannot be artificially improved by inclusion of the statistically redundant refinable parameters into the model under consideration. Nevertheless, the curves obtained with the reasonably high  $R^2_{\text{adjusted}}$  values ( $\geq 0.98$ ) were also visually inspected in order to reject the models giving the curves with physically unrealistic shapes.<sup>2</sup>

<sup>1</sup> In other words, all geometrical parameters reported in the literature for oxygen and hydrogen atoms and for nominally ‘single’ H–O bonds were not taken into account.

<sup>2</sup> For example, a number of rejected multi-parameter models gave curves which were wavy in the long-bond range.



**Figure 2**  
Approximation of the ‘bond valence–bond length’ correlation for H–O bonds using the equation  $s = (\alpha - \beta r)^{1/\gamma}$  with  $\alpha = 1.185$ ,  $\beta = 0.321$  and  $\gamma = 0.36$ .

Analysis of the results obtained for the ‘candidate’ functions has revealed the fact that the three-parameter function (3) (where  $\alpha$ ,  $\beta$  and  $\gamma$  are the refinable parameters) is probably the best ‘simple’ model for approximation of the ‘bond valence–bond length’ correlation for H–O bonds (Fig. 2).

$$s = (\alpha - \beta r)^{1/\gamma} \quad (3)$$

In the scientific literature this type of  $y = f(x)$  function is usually referred to as the ‘simplified Bleasdale–Nelder model’ (Bleasdale & Nelder, 1960; Seber & Wild, 1989). In addition to its high performance in approximation of the ‘bond valence–bond length’ correlation for H–O bonds [ $R^2_{\text{adjusted}} = 0.986$ ; maximum discrepancy of  $\sim 0.08$  v.u.;  $\alpha = 1.185$  (10),  $\beta = 0.321$  (8),  $\gamma = 0.36^3$ ], formula (3) has the following advantages:

- (i) the number of refinable parameters is reasonably small;
- (ii) the function is reversible [*i.e.*  $r$  can be directly calculated from  $s$  by using (4)] and easily programmable.

$$r = (\alpha - s^\gamma)/\beta \quad (4)$$

It must be noted that further (though small) improvement of the approximation of the above ‘bond valence–bond length’ correlation is definitely possible, but only at a cost of:

- (i) a significant complication of the empirical model,
- (ii) destabilization of the least-squares refinement due to correlations between the refinable parameters, and sometimes
- (iii) the unrealistic shape of the obtained correlation curve.

Thus, for example, the use of the ten-parameter polynomial function  $s = A_0 + A_1 r + A_2 r^2 + \dots + A_9 r^9$  resulted in  $R^2_{\text{adjusted}} = 0.987$  [which is only marginally better than 0.986 obtained with (3)], but the refinement was highly unstable and the obtained correlation curve was slightly wavy for the long-bond range.

Interestingly, refinements of the ‘bond valence–bond length’ correlation for H–O bonds using polynomial functions with higher numbers of parameters (such functions are widely used

<sup>3</sup> In order to avoid the superfluous and unjustified precision for the  $\gamma$  parameter, in the last refinement cycle the fitted value of this parameter was rounded to and fixed at the nearest two-decimal value.

for approximation of complex correlations) revealed the 'bulge' in the centre of the plot, as reported and theoretically explained by Brown (2002); however, for the long-bond range most of the tested polynomial functions also revealed smaller and clearly artificial 'bulges' (as mentioned in the above example) which cannot be explained in terms of the bond-valence model. Nevertheless, none of the polynomial functions showed the performance that can be regarded as *significantly* better than that demonstrated by (3) in approximation of the above correlation; therefore, the existence of the 'bulge' in the centre of the 'bond valence–bond length' correlation curve for H–O bonds has no clear statistical evidence yet, and the further complication of the empirical model is not really justified.

Taking all the aforementioned facts into account, equation (3) with the refined values of the parameters  $\alpha = 1.185$ ,  $\beta = 0.321$  and  $\gamma = 0.36$  can be recommended for bond-valence analysis as a tentative

empirical model for the description of the 'bond valence–bond length' correlation curve for H–O bonds.

## References

- Bleasdale, J. K. A. & Nelder, J. A. (1960). *Nature*, **188**, 342.  
Brown, I. D. (2002). *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*. IUCr Monograph on Crystallography 12. Oxford University Press.  
Brown, I. D. (2009). *Chem. Rev.* **109**, 6858–6919.  
Brown, I. D. (2010). Personal communication.  
Mitchell, M. (2002). *Engauge Digitizer*, Version 4.1, <http://digitizer.sourceforge.net/>.  
Rawlings, J. O., Pantula, S. G. & Dickey, D. A. (1998). *Applied Regression Analysis: A Research Tool*, 2nd ed. New York: Springer-Verlag.  
Seber, G. A. F. & Wild, C. J. (1989). *Nonlinear Regression*. New York: John Wiley and Sons.