

Scale factor database for the vibration frequencies calculated in M06-2X, one of the DFT methods

Yener Ünal^a, Wassim Nassif^b, Burak Can Özeydin^b, Koray Sayin^{b,c,*}

^a Department of Statistics and Computer Science, Faculty of Science, Sivas Cumhuriyet University, 58140 Sivas, Turkey

^b Department of Chemistry, Faculty of Science, Sivas Cumhuriyet University, 58140 Sivas, Turkey

^c Advanced Technology Research and Application Center, Sivas Cumhuriyet University, 58140 Sivas, Turkey

ARTICLE INFO

Keywords:

Vibrational scale factor
M06-2X method
Concordance correlation coefficient
Correlation
Regression

ABSTRACT

Vibrational scale factors were determined for the M06-2X method. The calculated scale factors were supported statistically. The main objective is to determine the concordance between frequencies. Descriptive statistics of the experimental and corrected frequencies are given. The relationship between corrected and experimental frequencies is formed by a simple linear regression method. The Concordance Correlation Coefficient determines the consistency between the two variables. It evaluates the direct relationship between X and Y variables measured under constraints. Also, the difference between the mean of the two variables was tested. All results are interpreted and presented with their graphs.

1. Introduction

Vibrational spectroscopy is a general technique dealing with the vibrational energy of molecules. This energy is considerably related to chemical bonds of molecules. Michelson interferometer and the Fourier algorithms played an essential role in the development of vibrational spectroscopy [1]. Three types of vibrational spectroscopy techniques are infrared (IR), Raman, and Terahertz spectroscopy. This spectroscopic technique is usually used in the broad spectrum, for instance, cosmetics, the pharmaceutical industry, food industry, biological, microbiology, and chemistry [2–4]. Infrared spectroscopy is based on infrared radiation interaction with a sample, leading to some bond vibrations, stretching, bending, rocking, scissoring, and twisting [5,6]. IR and Raman spectroscopy can be obtained by experimental methods as well as by computational chemistry methods.

In quantum chemical calculations, density functional theory (DFT) and ab-initio are the most used two methods. Both methods have advantages and disadvantages compared to each other. DFT methods have the ability about the more accurate calculations in the observable structural and electronic properties at the ground state when compared to ab-initio methods. In this stage, the main problems are the description of electron correlation, anharmonicity, and the solvation of the Schrödinger equation. To correct this error and obtain more accurate results, the calculation, tabulation, and use of the multiplicative linear

scaling factors proposed by Pople et al. have been popular in both literature and practice. Many published vibrational scale factors belong to the related calculation level in literature and the National Institute of Standards and Technology (NIST) computational chemistry comparison and benchmark database (CCCBDB) [7]. The M06-2X has been popular recently, but there is a little vibrational frequency scale factor that belongs to M06-2X (Table 1).

In this paper, 191 molecules in which their formula, name, and point group are given in Supp. Table S1 in Supplementary Information is taken into account. These molecules are separated into two groups, which are labeled as analysis molecules and test molecules. The analysis group contains 157 molecules, while test one contains 34 molecules. Mentioned molecules are optimized at M06-2X method with twenty five basis sets which are STO-3G, 3-21G, 3-21G*, 3-21+G, 3-21+G*, 6-31G, 6-31G(d), 6-31G(d,p), 6-31+G, 6-31+G(d), 6-31+G(d,p), 6-31++G, 6-31++G(d), 6-31++G(d,p), 6-311G, 6-311G(d), 6-311G(d,p), 6-311+G, 6-311+G(d), 6-311+G(d,p), 6-311++G, 6-311++G(d), 6-311++G(d,p), cc-pVDZ and aug-cc-pVDZ. The vibrational frequency corresponding to the symmetry type was recorded for each molecule in the analysis group and compared with the experimental ones. The vibrational scale factor belongs to each calculation level was derived. Then, the vibrational frequency corresponding to the symmetry type was recorded for each molecule in the test group. These frequencies are multiplied with derived scale factor. Corrected vibrational frequencies

* Corresponding author at: Department of Chemistry, Faculty of Science, Sivas Cumhuriyet University, 58140 Sivas, Turkey.

E-mail addresses: krysayin@gmail.com, ksayin@cumhuriyet.edu.tr (K. Sayin).

Table 1
Scale for values of the concordance correlation coefficient.

Value of ρ_c	Strength of agreement
> 0.99	Almost perfect
0.95 – 0.98	Substantial
0.90 – 0.94	Moderate
< 0.89	Poor

are compared with experimental results. In this study, descriptive statistics (mean, min, max, Std. Dev.) of corrected frequencies are presented in Table 4.

Since the sample size is quite large, independent sample *t*-test was used to determine differences means of corrected and experimental variables. These tests were tested by taking $\alpha = 0.05$ or 0.01 (margin of error), and the test results are given in Table 4. Correlation means the relationship between variables. In statistics, the coefficient showing the degree and direction of the relationship between variables is called the correlation coefficient, and the equation that determines the functional form of the relationship between variables is called the regression equation [8]. The correlation coefficient is denoted by the letter "r". If one of the variables increases, the other increases or one decreases while the other decreases, there is a positive (same direction) relationship. If one of the variables increases or decreases while the other changes in the other direction, there is a negative relation. It shows a strong correlation between variables when the coefficient *r* is close to 1 or -1; being close to 0 indicates no relationship. It is expected to have a high and positive relationship between experimental and corrected frequencies. The correlation coefficient *r* for two variables of unit *n*, such as X and Y, is calculated by the following formula.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

By considering corrected frequencies as a function of experimental frequencies, the relationship between them is formed by a simple linear regression method. As a dependent variable, Y shows experimental frequencies, independent variable X shows adjusted frequencies. The simple linear equation is expressed as follows;

$$Y = \beta_0 + \beta_1 X \quad (2)$$

Here β_0 is the ordinate of the point at which the regression line intersects the Y-axis. β_1 is the slope of the regression line or the effect of 1-unit change in the independent variable on the dependent variable. The rate of explanation of the dependent variable of the independent variable, i.e., the percentage of change in the dependent variable, is calculated by the coefficient of certainty. In the simple linear regression equation, the ratio of certainty is the square of the correlation coefficient and is denoted by R^2 . This value is between 0 and 1. The regression equation will be obtained by standardizing the experimental and correctable frequencies. In the model obtained, the coefficient β_0 will be 0 or very close to 0. Therefore, we can express the model. The coefficient of determination is expected to be very close to 1 [9].

While the experimental and corrected frequencies need to be in a high relationship, it is insufficient to say the correctness factor. They must also be highly compatible. The consistency between the two variables is determined by the Lin's Concordance Correlation Coefficient (CCC). CCC evaluates the linear relationship between X and Y variables measured under constraints where the cut-off point is zero and the slope is 1; that is, it evaluates the distance of each pair of measurements to the line drawn at an angle of 45°. Accordingly, CCC (ρ_c) is formulated as

$$\rho_c = \frac{2\sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 + (\mu_X - \mu_Y)^2} \quad (3)$$

Here, σ_{XY} denotes the covariance between the variables, denotes the mean of variables, and denotes the variance of variables [10].

The correlation coefficient and CCC together with C_b coefficient is a measure showing the relationship between variables. It is called the "bias correction factor" and is defined as the distance between the equality line and the line obtained from the regression equation. C_b can also be called a measure of accuracy and takes a value between 0 and 1 ($0 < C_b < 1$). $C_b = 1$ means that there is no separation from equality. The C_b side correction factor for the two variables is calculated as follows [11].

$$C_b = \left[\frac{(v+1/v+u^2)/2}{v} \right]^{-1} \quad (4)$$

$$v = \frac{\sigma_1}{\sigma_2}$$

$$u = (\mu_1 - \mu_2) / \sqrt{\sigma_1 \sigma_2}$$

Here, $0 < C_b \leq 1$ is a bias correction factor that measures how far the best-fit line deviates from the 45° line (measure of accuracy). No deviation from the 45° line occurs when $C_b = 1$. The further C_b is from 1, the greater the deviation is from the 45° line. The Pearson correlation coefficient *p* measures how far each observation deviated from the best-fit line (measure of precision).

2. Methodology

2.1. Computational method

In the first stage, the studied compounds were prepared in the appropriate geometry corresponding to the related point group and pre-optimization calculations were performed in the UFF method [12], which is one of the molecular mechanics methods by using personal computer. Then, the related compounds are optimized at M06-2X method [13] with STO-3 G [14,15], 3-21 G [16–21], 3-21G*, 3-21+G, 3-21+G*, 6-31 G [22–31], 6-31 G(d), 6-31 G(d,p), 6-31+G, 6-31+G(d), 6-31+G(d,p), 6-31++G, 6-31++G(d), 6-31++G(d,p), 6-311 G [28,29,32–38], 6-311 G(d), 6-311 G(d,p), 6-311+G, 6-311+G(d), 6-311+G(d,p), 6-311++G, 6-311++G(d), 6-311++G(d,p), cc-pVDZ [39–44] and AUG-cc-pVDZ [40,41] in gas phase by using TUBITAK TR-Grid server system. The whole structures were prepared by using GaussView 5.0.8 software program [45]. In the calculations, Gaussian IA32W-G09RevA.02 and Gaussian AS64L-G09RevD.01 programs were used [46,47].

Vibrational frequencies corresponding to the type of symmetry were obtained. For the data in the analysis group, the vibration scale factor (*c*) in each calculation level was calculated by using Eq. (5):

$$c = \frac{\sum \nu_i \omega_i}{\sum \omega_i^2} \quad (5)$$

where ν_i and ω_i are experimentally observed vibrational frequencies and theoretical vibrational frequencies, respectively. The relative uncertainty (u_r) of the vibration scale factor is calculated by using Eq. (6):

$$u_r = \frac{\sum (\omega_i^2 \cdot (c - \frac{\nu_i}{\omega_i})^2)}{\sum \omega_i^2} \quad (6)$$

2.2. Statistical method

In this study, it is aimed to determine the compatibility of experimental frequencies with corrected frequencies. For this purpose, all statistical analyzes and graphs were generated by using SPSS 23, Minitab 15, and Medcalc package programs.

Table 2

The compound names, formulas and point groups in the test group.

Name	Formula	Point Group	Name	Formula	Point Group
1,1,1-trichloroethane	CH ₃ CCl ₃	C _{3v}	Pyrazine	C ₄ H ₄ N ₂	D _{2h}
Cyclopropane	C ₃ H ₆	D _{3h}	Acetone	CH ₃ COCH ₃	C _{2v}
1H-pyrazole	C ₃ H ₄ N ₂	C _s	2-propane-1-ol	C ₃ H ₈ O	C ₁
1H-imidazole	C ₃ H ₄ N ₂	C _s	Oxetane	C ₃ H ₆ O	C _{2v}
1,3,5-triazine	C ₃ H ₃ N ₃	D _{3h}	Methoxyacetoneitrile	CH ₃ OCH ₂ CN	C ₁
Dimethyl ether	CH ₃ OCH ₃	C _{2v}	Propane	C ₃ H ₈	C _{2v}
Furan	C ₄ H ₄ O	C _{2v}	1,3-cyclopentadiene	C ₅ H ₆	C _{2v}
β-propiolactone	C ₃ H ₄ O ₂	C _s	Cyclopropylamine	C ₃ H ₇ N	C ₁
Dimethyl sulfite	CH ₃ SCH ₃	C _{2v}	Pyridine	C ₅ H ₅ N	C _{2v}
thiophene	C ₄ H ₄ S	C _{2v}	2-chloro-propane	CH ₂ CHClCH ₃	C _s
1,3-butadiene	CH ₂ CHCHCH ₂	C _{2h}	1,2-dichloro-propane	CH ₂ ClCHClCH ₃	C ₁
Cyclobutane	C ₄ H ₆	C _{2v}	Benzene	C ₆ H ₆	D _{6h}
Methylene Cyclopropane	C ₄ H ₆	C _{2v}	3-chloro-1-butene	CH ₂ CHCHClCH ₃	C ₁
Pyrole	C ₄ H ₅ N	C _{2v}	Spiropentane	C ₅ H ₈	D _{2d}
Cyclopropanecarbonitrile	C ₄ H ₅ N	C _s	Trimethylamine	N(CH ₃) ₃	C _{3v}
Pyridazine	C ₄ H ₄ N ₂	C _{2v}	2-methyl-propanal	CHOCH(CH ₃)CH ₃	C _s
1,3-diazine	C ₄ H ₄ N ₂	C _{2v}	Butane	CH ₃ CH ₂ CH ₂ CH ₃	C _{2h}

The following steps have been taken to accomplish the purpose of the project:

The compounds in the analysis group are optimized by considering their point group, and frequency assignments are performed.

The scale factor and relative uncertainty are derived.

The mentioned molecules in the test group are optimized, and their calculated frequencies are revised using the derived scale factor.

The reliability of the obtained results is compared statistically, and it is decided whether the derived scale factor is suitable.

Table 3

The calculated scale factor (c) and relative uncertainty (ur) of the related basis set at M06-2X method.

Basis Set	c	u _r	Basis Set	c	u _r
3-21G	0.9664	0.0616	6-31++G(d,p)	0.9509	0.0512
3-21G*	0.9601	0.0525	6-311G	0.9685	0.0593
3-21+G	0.9710	0.0687	6-311 G(d)	0.9540	0.0372
3-21+G(d)	0.9666	0.0602	6-311 G(d,p)	0.9553	0.0401
6-31G	0.9600	0.0606	6-311+G	0.9707	0.0620
6-31 G(d)	0.9496	0.0397	6-311+G(d)	0.9567	0.0421
6-31 G(d,p)	0.9509	0.0382	6-311+G(d,p)	0.9569	0.0369
6-31+G	0.9651	0.0620	6-311++G	0.9703	0.0626
6-31+G(d)	0.9529	0.0360	6-311++G(d)	0.9559	0.0363
6-31+G(d,p)	0.9522	0.0439	6-311++G(d,p)	0.9567	0.0371
6-31++G	0.9656	0.0616	cc-pVDZ	0.9575	0.0463
6-31++G(d)	0.9502	0.0489	aug-cc-pVDZ	0.9594	0.0701

3. Results and discussion

The studied compounds and their experimentally vibrational frequencies are taken from the Computational Chemistry Comparison and Benchmark Database and the NIST Chemistry Webbook [7]. One hundred ninety-one compounds are taken into consideration in this project. Additionally, 824 frequency values are used in the calculation of scale

Table 4

Descriptive statistics of corrected values obtained for all basis sets and independent sample t-test results.

Basis Set	Variable	Mean	St. Dev	Minimum	Maximum	P (Sig.)	Basis Set	Variable	Mean	St. Dev	Minimum	Maximum	P (Sig.)
3-21 g	Corrected	1487.7	937.8	85.3	3535.7	0.702	6-31++g(d,p)	Corrected	1451	925	40.2	3725.2	0.742
3-21 g(d)	Corrected	1477.8	931.8	84.7	3512.7	0.833	6-311(g)	Corrected	1486.2	938.8	38.3	3642.8	0.666
3-21+g	Corrected	1487.9	940	55.4	3544.4	0.660	6-311 g(d)	Corrected	1458	924.7	47.6	3704.5	0.859
3-21+gd	Corrected	1482.2	934.2	55.2	3528.3	0.729	6-311 g(d,p)	Corrected	1455.5	923.6	60.8	3752.5	0.816
6-31 g	Corrected	1482.3	939.4	36.1	3600.3	0.759	6-311+g	Corrected	1486.2	939.1	47.1	3653.4	0.667
6-31 g(d)	Corrected	1457.8	926.6	34.1	3642.2	0.825	6-311+g(d)	Corrected	1460.8	927.4	50.1	3710	0.908
6-31 g(d,p)	Corrected	1455	925.5	44.2	3717.4	0.778	6-311+g(d,p)	Corrected	1456.4	925.8	58.7	3760.7	0.832
6-31+g	Corrected	1487.1	946.3	25.7	3856.3	0.683	6-311++g	Corrected	1486.2	940.1	48.9	3651.6	0.666
6-31+g(d)	Corrected	1457.7	930.3	40.7	3658.4	0.854	6-311++g(d)	Corrected	1459.3	926.5	59	3706.8	0.882
6-31+g(d,p)	Corrected	1452.7	926.7	35.1	3730.6	0.740	6-311++g(d,p)	Corrected	1455.7	925.6	60.9	3759.6	0.820
6-31++g	Corrected	1486.3	944.2	34.4	3617.9	0.665	cc-pVDZ	Corrected	1453.4	929.5	70	3706.8	0.782
6-31++g(d)	Corrected	14,547	9265	441	3648	0.804	aug-cc-pVDZ	Corrected	1455	933.2	52.9	3741.1	0.808

factors, while 799 frequencies are used in determining the reliability of the derived scale factors' results. The analysis and test compounds name and point group of them are given in Supp. Table S1 and Table 2, respectively.

3.1. Computational investigations in analyses and test groups

The related molecules in the analysis group are optimized at the M06-2X method with the mentioned basis set. The point group is taken into consideration in the optimization step, and no imaginary frequencies are observed in the IR spectrum. The harmonic frequencies are calculated and compared with their experimental ones. The scale factor (c) and relative uncertainty (u_r) are calculated by using Eq. (1) and (2), respectively. The obtained results are given in Table 3. The harmonic frequencies in the analysis group are given in Supp. Table S2 – S25.

The frequencies of compounds in the test group are scale by using a derived scale factor. The related equation is given as follow:

$$\text{Corrected Frequency} = \text{Calculated Frequency} \times c \quad (7)$$

The experimental frequency and corrected frequency of related molecules in the test group is given in Supp. Table S26 – S49. According to obtained results, experimental and corrected frequencies are considerably agreement with each other. For all that, the compromise between

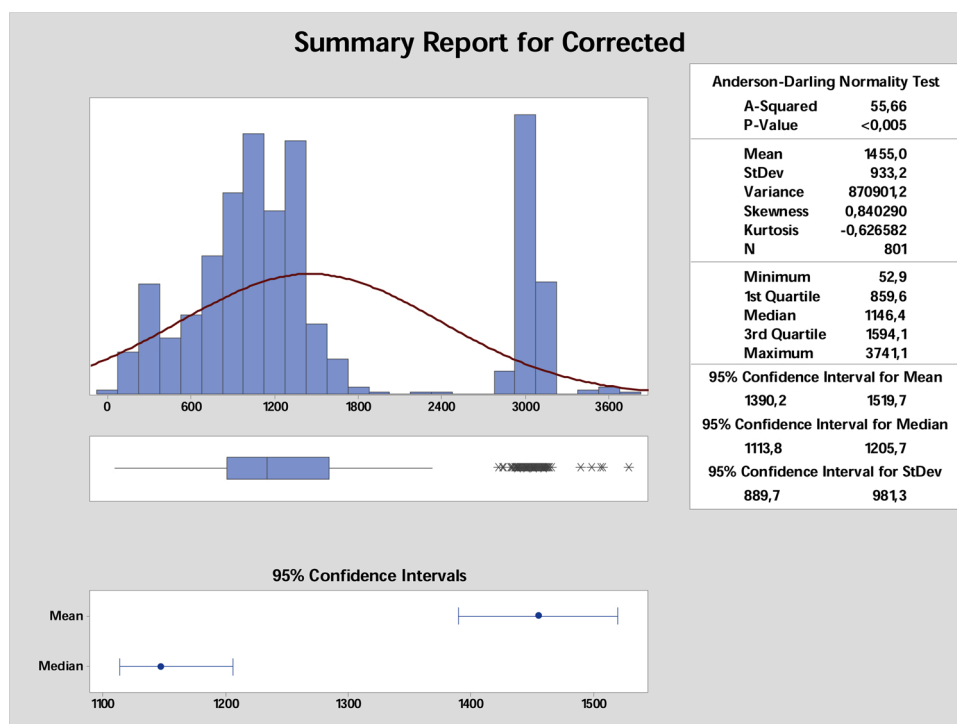


Fig. 1. Summary Report for Corrected Frequencies at M062-2X/ aug-cc-pVDZ level.

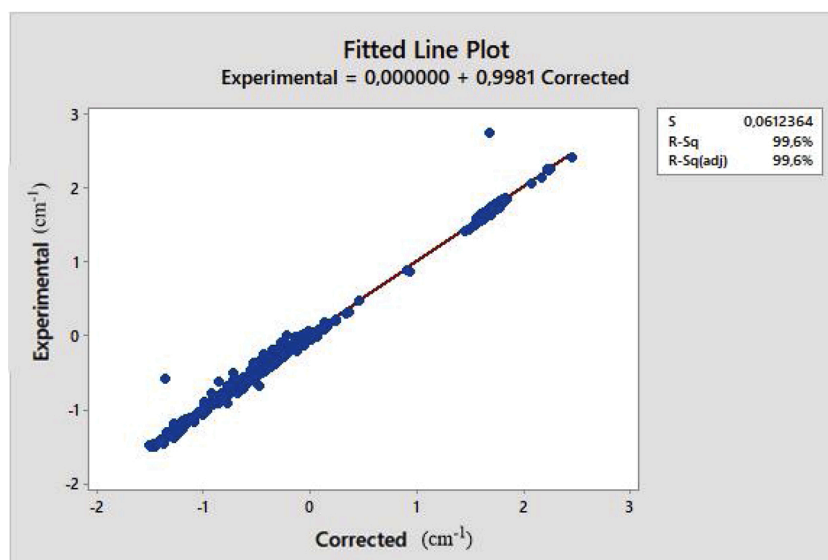


Fig. 2. Fitted line plot corrected and experimental values for the aug-cc-pVDZ basis set.

experimental and calculated results and reliability of the derived scale factor is statistically evaluated in the following section.

3.2. Statistical analysis of the results and determining the accuracy and precision of derived scale factors

Descriptive statistics of experimental and corrected values obtained for all basis sets and *t*-test results are given in Table 4. The mean of Experimental values is found 1466.2 (± 918.6). The minimum value is 75, and the maximum value is 3974.

The basic statistics and histogram graph of the corrected values obtained for the aug-cc-pVDZ basis set are given in Fig. 1. According to the Anderson-Darling normality test, it is seen that fixed values are not

suitable for normal distribution. Skewness and Kurtosis are calculated as 0.841 and -0.626, respectively. The deviation of the skewness value from 0 and the kurtosis value from 3 means that it differs from the normal distribution.

According to Table 4, there was no statistically significant difference between the median values of experimental and corrected values for all basis sets ($p(\text{sig.}) > 0.05$).

For a aug-cc-pVDZ basis set, the simple linear equation formed between experimental frequencies and corrected frequencies is as follows.

$$Y = 0.998 * X$$

Here, Y denotes the experimental values, and X indicates the corrected values. In the equation, the coefficient of the fixed values is calculated as

Table 5
Results of analysis of all basis sets.

Basis Set	<i>r</i>	$Y = \beta_1 X$	R ²	CCK	<i>C_b</i>
STO-3g	0.992	$Y = 0.997 * X$	0.995	0.9964	0.9990
3-21 g	0.995	$Y = 0.998 * X$	% 99.57	0.9974	0.9995
3-21 g(d)	0.997	$Y = 1.0087 * X$	% 99.48	0.9972	0.9998
3-21+g	0.998	$Y = 1.0166 * X$	% 99.55	0.9972	0.9995
321+g(d)	0.998	$Y = 1.012 * X$	% 99.57	0.9976	0.9997
6-31 g	0.998	$Y = 1.0137 * X$	% 99.61	0.9977	0.9996
6-31 g(d)	0.998	$Y = 0.9979 * X$	% 99.65	0.9982	0.9999
6-31 g(d,p)	0.998	$Y = 0.9961 * X$	% 99.64	0.9982	0.9999
6-31+g	0.9973	$Y = 1.018 * X$	% 99.45	0.9966	0.9993
6-31+g(d)	0.998	$Y = 0.9987 * X$	% 99.50	0.9974	0.9999
6-31+g(d,p)	0.998	$Y = 0.9954 * X$	% 99.64	0.9981	0.9999
6-31++g	0.998	$Y = 1.0172 * X$	% 99.61	0.9975	0.9994
6-31++g(d)	0.998	$Y = 0.9963 * X$	% 99.65	0.9982	0.9999
6-31++g(d,p)	0.998	$Y = 0.994 * X$	% 99.63	0.9981	0.9998
6-311 g	0.998	$Y = 1.0155 * X$	% 99.63	0.9977	0.9995
6-311 g(d)	0.998	$Y = 0.9974 * X$	% 99.66	0.9983	0.9999
6-311 g(d,p)	0.998	$Y = 0.9958 * X$	% 99.65	0.9982	0.9999
6-311+g	0.997	$Y = 1.0154 * X$	% 99.50	0.9970	0.9995
6-311+g(d)	0.998	$Y = 0.9996 * X$	% 99.66	0.9983	0.9999
6-311+g(d,p)	0.998	$Y = 0.9969 * X$	% 99.65	0.9982	0.9999
6-311++g	0.998	$Y = 1.0159 * X$	% 99.62	0.9976	0.9995
6-311++g(d)	0.998	$Y = 0.9986 * X$	% 99.66	0.9983	0.9999
6-311++g(d,p)	0.998	$Y = 0.9965 * X$	% 99.65	0.9982	0.9999
cc-pVDZ	0.998	$Y = 0.9965 * X$	% 99.59	0.9979	0.9998
aug-cc-pVDZ	0.998	$Y = 0.9985 * X$	% 99.59	0.9979	0.9998

0.998. The fact that this value is approximately equal to 1 indicates that the corrected values are highly compatible with the experimental values. The ratio of fixed values to experimental values was calculated. This value is a reasonably high explanation rate. Therefore, 99.59 % of the variability in the experimental data is explained by the regression model, indicating a perfect fit of the model.

Used fitted line Plot to display the relationship between one continuous predictor and response. A fitted line plot shows a scatterplot of the data with a regression line representing the regression equation. The distribution graph in Fig. 2 shows a strong correlation between the experimental and corrected values and the appropriate linear equation.

CCC was obtained as 0.997 for experimental and corrected frequencies. This value shows excellent compatibility between experimental and corrected frequencies. It *C_b* is called the "side correction factor" and is defined as the measure of the distance of the equality line to the line obtained from the regression equation. *C_b* can also be called a measure of accuracy and takes a value between 0 and 1 ($0 < C_b < 1$). *C_b* = 1 means that there is no separation from equality. The value for experimental and corrected frequencies was calculated as 0.9998. It was determined that the obtained regression line did not deviate from the equation line. As a result, the generated graphs and analysis showed that the experimental and corrected frequencies were statistically perfectly compatible. Thus, the accuracy of the calculated correction factor was determined. All the results of the other basis sets are given Table 5 and interpreted the same as aug-cc-pVDZ basis set.

It was determined that the experimental and corrected frequencies were not compatible with normal distribution according to the Anderson Darling normality test. The magnitude and direction of the relationship between the two variables were examined by Spearman's rho test, a non-parametric test. According to the test results, a high positive correlation was calculated between the two variables at the rate of 0.998.

4. Conclusion

One hundred ninety-one small molecules are selected for the

derivation of the scale factor. Firstly, all compounds are optimized at the related calculation level. In the calculation, point groups are taken into account. The scale factors are calculated by using the result of 157 compounds. Then, the frequencies of the rest compounds are scaled by using derived scale factors and compare with experimental values. All results are interpreted as statistics. The main objective of the statistical studies is to determine the concordance between the frequencies corrected with experimental frequencies. For this purpose, these two variables were compared with various statistical methods. First of all, whether there is the independent sample *t*-test tested a difference between them. It was determined that there was no difference between the median values of the experimental and corrected values for all basis settlements. With the correlation test, a very high level of positive relationship was found between the two variables. By considering corrected frequencies as a function of experimental frequencies, the relationship between them is formed by a simple linear regression method. *Y* shows experimental frequencies; the independent variable shows adjusted frequencies. From the equations obtained, the coefficient of the corrected values (adj. R) was calculated for all basis sets. All values approximately equal to 1 indicate that the corrected values are highly compatible with the experimental values. The consistency between experimental and corrected frequencies was determined by the Concordance Correlation Coefficient (CCC). It was seen that the obtained coefficient of compatibility was almost 1 for all basis sets. This value shows excellent compatibility between the two variables. All of the results show that frequencies corrected with experimental frequencies are exceptionally compatible and highly related.

Declaration of Competing Interest

The authors report no declarations of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.vibspec.2020.103189>.

References

- [1] I.M. Alecu, J. Zheng, Y. Zhao, D.G. Truhlar, Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries, *J. Chem. Theory Comput.* 6 (2010) 2872–2887.
- [2] D.O. Kashinski, G.M. Chase, R.G. Nelson, O.E. Di Nallo, A.N. Scales, D. L. VanderLey, E.F.C. Byrd, Harmonic vibrational frequencies: approximate global scaling factors for TPSS, M06, and M11 functional families using several common basis sets, *J. Phys. Chem. A* 121 (2017) 2265–2273.
- [3] A.M. Teixeira, C. Sousa, A review on the application of vibrational spectroscopy to the chemistry of nuts, *Food Chem.* 277 (2019) 713–724.
- [4] S. Bureau, D. Cozzolino, C.J. Clark, Contributions of Fourier-transform mid-infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: a review, *Postharvest Biol. Technol.* 148 (2019) 1–14.
- [5] B. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*, Lta: John Miley & Sons, New York, 2004.
- [6] B.C. Smith, *Fundamentals of Fourier Transform Infrared Spectroscopy*, 2nd edition, CRC Press, Taylor & Francis Group, 2011.
- [7] NIST Computational Chemistry Comparison and Benchmark Database, 2013. WWW Page, <http://cccbdb.nist.gov>.
- [8] G. Akkoca, Statistical Methods Used for Measuring Agreement Between Continuous Outcome Measurement Methods In Clinical Researches, Ankara University Graduate School of Health Sciences, Ankara, 2012. Biostatistics Master Thesis.
- [9] H. Gamgam, B. Altunkaynak, Regresyon Analize, Seçkin Yayıncılık, Ankara, 2015.
- [10] L. Lin, A concordance correlation coefficient to evaluate reproducibility, *Biometrics* 45 (1989) 255–268.
- [11] V. Chinchilli, J. Martel, S. Kumanyika, L. Lloyd, A weighted concordance correlation coefficient for repeated measurement designs, *Biometrics* 52 (1996) 341–353.
- [12] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard III, W.M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *J. Am. Chem. Soc.* 114 (1992) 10024–10035.
- [13] Y. Zhao, D.G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theor. Chem. Acc.* 120 (2008) 215–241.

- [14] W.J. Hehre, R.F. Stewart, J.A. Pople, Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals, *J. Chem. Phys.* 1 (1969) 2657–2664.
- [15] J.B. Collins, P.V.R. Schleyer, J.S. Binkley, J.A. Pople, Self-consistent molecular orbital methods. XVII. Geometries and binding energies of second-row molecules. A comparison of three basis sets, *J. Chem. Phys.* 64 (1976) 5142–5511.
- [16] J.S. Binkley, J.A. Pople, W.J. Hehre, Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements, *J. Am. Chem. Soc.* 102 (1980) 939–947.
- [17] M.S. Gordon, J.S. Binkley, J.A. Pople, W.J. Pietro, W.J. Hehre, Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements, *J. Am. Chem. Soc.* 104 (1982) 2797–2803.
- [18] W.J. Pietro, M.M. Francl, W.J. Hehre, D.J. Defrees, J.A. Pople, J.S. Binkley, Self-consistent molecular orbital methods. 24. Supplemented small split-valence basis sets for second-row elements, *J. Am. Chem. Soc.* 104 (1982) 5039–5048.
- [19] K.D. Dobbs, W.J. Hehre, Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third- and fourth-row, main-group elements, *J. Comp. Chem.* 7 (1986) 359–378.
- [20] K.D. Dobbs, W.J. Hehre, Molecular orbital theory of the properties of inorganic and organometallic compounds 5. Extended basis sets for first-row transition metals, *J. Comp. Chem.* 8 (1987) 861–879.
- [21] K.D. Dobbs, W.J. Hehre, Molecular orbital theory of the properties of inorganic and organometallic compounds. 6. Extended basis sets for second-row transition metals, *J. Comp. Chem.* 8 (1987) 880–893.
- [22] R. Ditchfield, W.J. Hehre, J.A. Pople, Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules, *J. Chem. Phys.* 54 (1971) 724–728.
- [23] W.J. Hehre, R. Ditchfield, J.A. Pople, Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules, *J. Chem. Phys.* 56 (1972) 2257–2261.
- [24] P.C. Hariharan, J.A. Pople, The influence of polarization functions on molecular orbital hydrogenation energies, *Theor. Chem. Acc.* 28 (1973) 213–222.
- [25] P.C. Hariharan, J.A. Pople, Accuracy of AH_n equilibrium geometries by single determinant molecular orbital theory, *Mol. Phys.* 27 (1974) 209–214.
- [26] M.S. Gordon, The isomers of silacyclopentane, *Chem. Phys. Lett.* 76 (1980) 163–168.
- [27] M.M. Francl, W.J. Pietro, W.J. Hehre, J.S. Binkley, D.J. Defrees, J.A. Pople, M. S. Gordon, Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements, *J. Chem. Phys.* 77 (1982) 3654–3665.
- [28] R.C. Binning Jr., L.A. Curtiss, Compact contracted basis sets for third-row atoms: Ga–Kr, *J. Comp. Chem.* 11 (1990) 1206–1216.
- [29] J. –P. Blaudeau, M.P. McGrath, L.A. Curtiss, L. Radom, Extension of Gaussian-2 (G2) theory to molecules containing third-row atoms K and Ca, *J. Chem. Phys.* 107 (1997) 5016–5021.
- [30] V.A. Rassolov, J.A. Pople, M.A. Ratner, T.L. Windus, 6-31G* basis set for atoms K through Zn, *J. Chem. Phys.* 109 (1998) 1223–1229.
- [31] V.A. Rassolov, M.A. Ratner, J.A. Pople, P.C. Redfern, L.A. Curtiss, 6-31G* basis set for third-row atoms, *J. Comp. Chem.* 22 (2001) 976–984.
- [32] K. Raghavachari, J.S. Binkley, R. Seeger, J.A. Pople, Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, *J. Chem. Phys.* 72 (1980) 650–654.
- [33] A.D. McLean, G.S. Chandler, Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11–18, *J. Chem. Phys.* 72 (1980) 5639–5648.
- [34] A.J.H. Wachters, Gaussian basis set for molecular wavefunctions containing third-row atoms, *J. Chem. Phys.* 52 (1970) 1033–1036.
- [35] P.J. Hay, Gaussian basis sets for molecular calculations. The representation of 3d orbitals in transition-metal atoms, *J. Chem. Phys.* 66 (1977) 4377–4384.
- [36] K. Raghavachari, G.W. Trucks, Highly correlated systems. Excitation energies of first row transition metals Sc–Cu, *J. Chem. Phys.* 91 (1989) 1062–1065.
- [37] M.P. McGrath, L. Radom, Extension of Gaussian-1 (G1) theory to bromine-containing molecules, *J. Chem. Phys.* 94 (1991) 511–516.
- [38] L.A. Curtiss, M.P. McGrath, J. –P. Blaudeau, N.E. Davis, R.C. Binning Jr., L. Radom, Extension of Gaussian-2 theory to molecules containing third-row atoms Ga–Kr, *J. Chem. Phys.* 103 (1995) 6104–6113.
- [39] T.H. Dunning Jr., Rates of convergence and error estimation formulas for the Rayleigh–Ritz variational method, *J. Chem. Phys.* 83 (1985) 1173–1196.
- [40] R.A. Kendall, T.H. Dunning Jr., R.J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, *J. Chem. Phys.* 96 (1992) 6796–6806.
- [41] D.E. Woon, T.H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon, *J. Chem. Phys.* 98 (1993) 1358–1371.
- [42] K.A. Peterson, D.E. Woon, T.H. Dunning Jr., Benchmark calculations with correlated molecular wave functions. IV. The classical barrier height of the H+H₂→H₂+H reaction, *J. Chem. Phys.* 100 (1994) 7410–7415.
- [43] A.K. Wilson, T. van Mourik, T.H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon, *J. Mol. Struct. (Theochem.)* 338 (1996) 339–349.
- [44] E.R. Davidson, Comment on “Comment on Dunning’s correlation-consistent basis sets”, *Chem. Phys. Lett.* 260 (1996) 514–518.
- [45] GaussView 5.0, Dennington II, Roy Dennington, Todd A. Keith, and John M. Millam., Wallingford, CT, 2009.
- [46] Gaussian IA32W-G09RevA.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.
- [47] Gaussian AS64L-G09RevD.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.