Research paper

Assessment of the Thermal Decomposition Temperature of High-Energy Heterocyclic Aromatic Compounds in Order to Increase Their Safety during Storage, Handling, and Application

Mohammad Hossein Keshavarz*[)], Sajjad Mousavi, Masoumeh Drikvand

Faculty of Applied Sciences, Malek-Ashtar University of Technology, Iran
* *E-mail:* mhkeshavarz@mut-es.ac.ir

ORCID information
Keshavarz M.H.: https://orcid.org/0000-0001-9269-752

Abstract: Heterocyclic aromatic compounds containing various derivatives of five or six-membered nitrogen-containing rings, *viz.* triazoles, tetrazoles, furazans, triazines, and tetrazines, and some of their salts have been studied. High nitrogen content and a large positive heat of formation are two important properties of these compounds. Two new models are introduced for the reliable prediction of the thermal stability of these compounds through the thermal decomposition temperature (onset). The reported data for 181 compounds have been used to derive and test the new models. For a training set containing 132 heterocyclic aromatic compounds, the values of the average absolute deviation (AAD) and the coefficient of determination (R^2) of the improved correlation were 9.72 K and 0.959, respectively. For triazoles, tetrazoles, furazans, triazines, and tetrazines, the predicted results of ADD and R^2 for the external test data set for this model containing 41 compounds were 23.03 K and 0.664, respectively, which are closer to experimental data than those obtained by the core correlation, *i.e.* 26.49 K and 0.653, respectively. The correlation coefficients of cross-validation for leave-one-out (Q^2_{LOO}) and 5-fold (Q^2_{5CV}) of the improved correlation were 0.955 and 0.951, respectively, which confirm that it is not an over-fitted model, robust and well-behaved.

Keywords: thermal decomposition temperature, heterocyclic aromatic compound, molecular structure, correlation, safety aspect

Supplementary Information (SI):

SI contains the Tables S1 and S2 (cited in this paper), and informative only Table S3 (not cited), which contains an index of the names used in Tables S1 and S2. The molecular structures of the compounds mentioned in this paper, heterocyclic aromatic compounds and some of their salts, are given in Tables S1 and S2, as well as the results of calculations for the training and test sets, according to equations presented in this paper and taken from literature.

Nomenclature

AAD	Average absolute deviation
AARD%	Average absolute relative deviation
AD	Average deviation
AD_{max}	Absolute maximum deviation
C_{Azide}	The presence of azide groups
$C_{>C=NNO_2 \text{ or } -C(=NH)-NH(NO_2)}$	The presence of $>C=NNO_2$ or $-C(=NH)-NH(NO_2)$ groups
$C_{N-N(NH)-C(NO_2) \text{ or } N=N-NH-C(NO_2)}$	The presence of $N-N(NH)-C(NO_2)$ or $N=N-NH-C(NO_2)$ groups
$C_{ONO_2 \text{ or } CH(NO_2)_2 \text{ or dinitroamide ion}}$	The presence of ONO_2 or $CH(NO_2)_2$ or dinitroamide ion groups
CI	Confidence interval
CV	Cross-validation
DSC	Differential Scanning Calorimetry
LOO	Leave-one-out
LOO CV	Leave-one-out cross-validation
MLR	Multiple linear regression
n_C	The number of carbon atoms
$n_{furazan}$	The number of furazan rings
n_H	The number of hydrogen atoms
n_N	The number of nitrogen atoms
$n_{NH_2+Cl-C(NO_2)-C(NO_2)+2(CuN+1.5NONH_3^+)+3\times\text{itrotetrazolate salts of sodium or silver}}$	The sum of NH_2, $Cl-C(NO_2)-C(NO_2)$ and two times the number of CuN bound and three times the number of hydroxylammonium ions, as well as three times the number of nitrotetrazolate salts of sodium or silver ions

$n_{\text{NH or NHNH}_2}$	The number of NH or NHNH$_2$ groups
n_O	The number of oxygen atoms
Q^2	The correlation coefficient of cross-validation
Q$^2_{\text{LOO}}$	The correlation coefficient of leave-one-out cross-validation
Q$^2_{\text{5CV}}$	The correlation coefficient of 5-fold cross-validation
QSPR	Quantitative Structure-Property Relationship
R^2	The coefficient of determination
RMSE	Root-mean-square error
SE	Standard error
T_d	Thermal decomposition temperature (onset) [K]
$T_{d,\text{Add}}$	Additive contribution of some polar groups and furazan rings
$T_{d,\text{core}}$	The core thermal decomposition temperature [K]
$T_{d,\text{Elem}}$	The contribution of elemental composition
$T_{d,\text{Non–add}}$	The non-additive contribution of thermally active groups including –ONO$_2$, –N$_3$, –CH(NO$_2$)$_2$, and >C=NNO$_2$
T_d^{in}	Correcting parameter for a large negative deviation of the core correlation from the experimental data
T_d^{dec}	Correcting parameter for a large positive deviation of the core correlation from the experimental data

1 Introduction

Organic compounds containing energetic groups, and their salts, with a high-energy content can ignite, decompose and detonate when exposed to various stimuli such as heat, impact, friction, and electric spark [1-5]. Heat sensitivity of these compounds shows the temperature at which their thermal decomposition occurs, which can be used to estimate their thermal stability [2]. It can also be used to assess the shelf life of these compounds because it is important in the safety aspects of high energy compounds in storage, handling, and application [6-8]. Since the activation energy is an important parameter for the assessment of the shelf life of energetic materials, some studies have been made to relate some types of sensitivity to their activation energies [9, 10]. Different types of thermal analysis methods including Differential Scanning Calorimetry (DSC) have been widely used to determine thermal decomposition temperatures [2]. The thermal decomposition temperature (onset) can be determined from the first deflection from the baseline. It usually differs from the temperature at which the maximum mass loss occurs [11].

Different methods of Quantitative Structure-Property Relationship (QSPR) have been widely used in recent years to predict the heat of thermal decomposition and temperature of organic compounds, ionic liquids, and salts containing energetic groups [12-21]. Most of these approaches used effective elemental composition, functional groups and molecular moieties, rather than using complex descriptors and computer codes. The ability to predict thermal properties, in particular, the thermal decomposition temperature(onset) (T_d) through the molecular structure, is of ultimate importance because the reported values of T_d are usually higher than the other types of decomposition temperatures, *e.g.* temperature of maximum mass loss [11]. Thus, it is valuable to have reliable methods that can be used for the prediction of T_d for a wide range of energetic compounds.

Heterocyclic aromatic compounds and some of their salts, such as high nitrogen containing materials, have been widely introduced in recent years as high energy content organic compounds [22]. They usually have large positive heat of formation, which can provide a high heat of combustion/detonation [23, 24]. Furthermore, they generate nitrogen gas as a major product with a relatively large bond energy, *i.e.* 941 kJ/mol [25]. If heterocyclic aromatic compounds with high density are used as high explosives, they can provide high detonation velocities, pressures, temperatures, and Gurney energies [26]. The use of heterocyclic aromatic compounds in the composition of propellants can also provide high combustion performance [27].

The design and discovery of new energetic compounds with improved thermodynamic stability, performance, and safety, before they are synthesized is highly desirable for scientists and the chemical industry because reliable predictive methods prevent investing enormous resources for compounds with low decomposition temperatures [28]. Moreover, the development of these compounds requires expensive and time-consuming synthesis and characterization loops [29]. Since many synthesis experiments yield little useful information, it is important to assess the critical properties of energetic materials [30]. Due to the high-energy content of heterocyclic aromatic compounds and their salts, it is important to estimate their thermal decomposition temperatures as an important starting point for evaluating their thermal stability [31].

The purpose of the present work was to introduce a general correlation for predicting the thermal decomposition temperature (onset) of a wide range of heterocyclic aromatic compounds and some of their salts. These compounds included various derivatives of five or six-membered nitrogen-containing rings, such as triazoles, tetrazoles, furazans, triazines, and tetrazines. A core correlation is introduced based on a combination of the number of some

elemental compositions and some specific functional groups. An improved correlation is developed by considering some structural moieties through two improving functions. For some of the azole derivatives, the predicted data of two correlations are compared with the method of Zohari *et al.* [18], as the best available model where it can be applied. The improved correlation may be used as a simple predictive model with high reliability for a quick estimate of the decomposition temperature (onset) of any desired heterocyclic aromatic compound and some of its salts.

2   Development of the New Models

2.1 Experimental data for building and testing

The reported data of 181 heterocyclic aromatic compounds and some of their salts, including sodium (or potassium), silver, hydroxylammonium, and dinitroamide were used for the development of the new models, where these data were collected by Zohari *et al.* [18] and Klapötke [4] from different sources. For azole-based energetic compounds, Zohari *et al.* [18] used 115 and 30 compounds with different molecular structures as the training and test sets, respectively. These data were also used here to build and test the new models. The remaining data from other classes of energetic compounds were added to the mentioned data. Thus, the training and test data sets of the current study contain 132 and 49 heterocyclic aromatic compounds, respectively. The measured data, as well as the molecular structures of these heterocyclic aromatic compounds and some of their salts, in the training and test sets, are given in Tables S1 and S2 (see SI), respectively.

2.2 Core correlation based on elemental composition and specific polar groups

Predicting the thermal decomposition temperature (onset) of heterocyclic aromatic compounds and their salts is very complex because it is important to specify active factors for the initiation of decomposition. Previous studies of the decomposition temperature of some classes of high-energy organic and metal-organic frameworks have indicated that the contribution of some elemental compositions is one of the important factors [16, 17, 19, 20, 32]. The presence of some functional groups can also increase or decrease the temperature sensitivity of an energetic material. It was found that the presence of $-NH_2$ and $-NH-$ groups can decrease the sensitivity of some classes of organic energetic compounds, *e.g.* nitroaromatic compounds, to the stimulus of impact, electric spark, shock, and heat [2, 33-35]. The presence of amino groups exhibits

an additive contribution to various stimuli. For example, 1,3,5-triamino-2,4,6-trinitrobenzene (TATB) is more insensitive than 1,3-diamino-2,4,6-trinitrobenzene (DATB) and 2,4,6-trinitroaniline (TNA) [36-38]. This situation may be confirmed for the thermal decomposition temperatures of heterocyclic aromatic compounds, *e.g.* the reported thermal decomposition temperature of 1-(2-(5-amino-3-nitro-1*H*-1,2,4-triazol-1-yl)ethyl)-1*H*-tetrazol-5-amine, with two amino groups, is higher (550 K) [39] than that for 1-(2-(3-nitro-1*H*-1,2,4-triazol-1-yl)ethyl)-1*H*-tetrazol-5-amine, with one amino group (528 K) [39]. Detailed studies of various substituents on heterocyclic aromatic compounds have shown that the number of $-NNH_2NO_2$, 5-chloro-3,4-dinitropyrazole, CuN (copper nitrogen bond), hydroxylammonium ion, sodium ion, silver ion, and $-NHNH_2$, in addition to $-NH_2$ and $-NH-$ groups, can increase the values of the thermal decomposition temperatures in these compounds. It was found that the number of furazan rings could reduce the thermal decomposition temperature. For example, the value of the thermal decomposition temperature of 3,4-diaminofurazan is 518 K [4], which is close to that of 3-(4-amino-1,2,5-oxadiazol-3-yl)-4-(4-nitro-1,2,5-oxadiazol-3-yl)-1,2,5-oxadiazole (508 K) [4]. The presence of three furazan rings can prevent the increasing insensitivity effect of the amino group in 3-(4-amino-1,2,5-oxadiazol-3-yl)-4-(4-nitro-1,2,5-oxadiazol-3-yl)-1,2,5-oxadiazole. It was found that the presence of $-N_3$, $-ONO_2$, $-NHNO_2$,$-CH(NO_2)_2$, $=NNO_2$, and 1,1-dinitroethane groups can decrease the value of the thermal decomposition temperature. The influence of elemental composition, and the additive and non-additive contributions of the mentioned groups, can be regressed for different quantities using the multiple linear regression (MLR) method [40] to provide the largest and the lowest values of the coefficient of determination (R^2) and the root-mean-square error (RMSE) as follows:

$$T_{d,core} = 490.2 + 33.21 T_{d,Elem} + 24.63 T_{d,Add} - 52.82 T_{d,Non-add} \tag{1}$$

where $T_{d,core}$ is the core thermal decomposition temperature (in K). There are three descriptors in Equation 1, including $T_{d,Elem}$, $T_{d,Add}$, and $T_{d,Non-add}$. The parameter $T_{d,Elem}$ shows the contribution of the elemental composition as:

$$T_{d,Elem} = n_C - 0.3339 n_H - 0.2878 n_N - 0.2637 n_O \tag{2}$$

where n_C, n_H, n_N and n_O are the number of carbon, hydrogen, nitrogen, and oxygen atoms. As seen in Equation 2, the coefficient of n_C is positive, and is about three times the coefficients of n_H, n_N, and n_O. Since the contributions of these other

elements cannot improve the values of R^2 and RMSE, their contributions have been neglected. Thus, it can be expected that a compound with a high content of carbon atoms will show a higher decomposition temperature. For example, 4,8-bis(2,4,6-trinitrophenyl)-4H,8H-bis([1,2,5]oxadiazolo)[3,4-b:3′,4′-e] pyrazine, with molecular formula $C_{16}H_4N_{12}O_{14}$, has the highest number of carbon atoms and experimental decomposition temperature (688 K [4]) as compared to the other heterocyclic aromatic compounds given in Tables S1 and S2 (see SI). The parameter $T_{d,Add}$ is the additive contribution of some polar groups and furazan rings as:

$$T_{d,Add} = n_{NH_2+Cl–C(NO_2)–C(NO_2)+2(CuN+1.5NONH_3^+)+3×itrotetrazolate\ salts\ of\ sodium\ or\ silver} + 0.9779n_{NH\ or\ NHNH_2} - 0.9122n_{furazan} \quad (3)$$

where $n_{NH_2+Cl–C(NO_2)–C(NO_2)+2(CuN+1.5NONH_3^+)+3×itrotetrazolate\ salts\ of\ sodium\ or\ silver}$ is the sum of NH_2, $Cl–C(NO_2)–C(NO_2)$ and two times the number of CuN bonds and three times the number of hydroxylammonium ions, as well as three times the number of nitrotetrazolate salts of sodium or silver ion; $n_{NH\ or\ NHNH_2}$ is the number of NH or $NHNH_2$ groups; $n_{furazan}$ is the number of furazan rings. As indicated in Equation 3, the coefficients of $n_{NH_2+Cl–C(NO_2)–C(NO_2)+2(CuN+1.5NONH_3^+)+3×itrotetrazolate\ salts\ of\ sodium\ or\ silver}$ and $n_{NH\ or\ NHNH_2}$ have positive values but $n_{furazan}$ has a negative coefficient. Thus, increasing the number of NH_2, $Cl–C(NO_2)–C(NO_2)$, CuN, $NONH_3^+$, nitrotetrazolate salts of sodium or silver ion, NH and $NHNH_2$ groups can enhance the thermal stability of the compounds. In contrast, an increase in the number of furazan rings may generate thermal instability. The parameter $T_{d,Non–add}$ provides the non-additive contribution of thermally active groups, including $–ONO_2$, $–N_3$, $–CH(NO_2)_2$, and $>C=NNO_2$ as:

$$T_{d,Non–add} = C_{>C=NNO_2\ or\ –C(=NH)–NH(NO_2)} + 2.20C_{N–N(NH)–C(NO_2)\ or\ N=N–NH–C(NO_2)} + 1.171C_{Azide} + 1.037C_{ONO_2\ or\ CH(NO_2)_2\ or\ dinitroamide\ ion} \quad (4)$$

where $C_{>C=NNO_2\ or\ –C(=NH)–NH(NO_2)}$, $C_{N–N(NH)–C(NO_2)\ or\ N=N–NH–C(NO_2)}$, C_{Azide} and $C_{ONO_2\ or\ CH(NO_2)_2\ or\ dinitroamide\ ion}$ show the presence of the specified groups or molecular fragments.
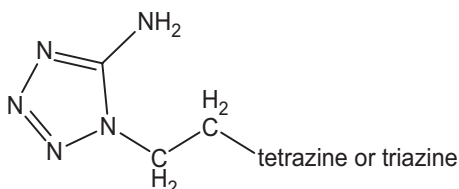
2.2 Improvement of the core correlation

A comparison of the predicted results from the training set by Equation 1 with experimental data of the decomposition temperatures shows that the presence of some structural moieties and specific isomers may give large deviations. The presence of these parameters can provide large positive

or negative deviations. In contrast to melting point and the entropy of fusion of heterocyclic aromatic compounds, there is no relation between these phase change properties and decomposition temperatures because molecular symmetry cannot affect the active site for decomposition [41, 42]. The effects of molecular moieties for large positive or negative deviations by molecular moieties can be explained by one example. The uses of Equations 1-4 for three compounds:

– 1-(2-(3-nitro-1H-1,2,4-triazol-1-yl)ethyl)-1H-tetrazol-5-amine (Exp. = 528 K [39]),
– 1-(2-(5-amino-3-nitro-1H-1,2,4-triazol-1-yl)ethyl)-1H-tetrazol-5-amine (Exp. = 550 K [39]), and
– 1-(2-(5-nitro-2H-tetrazol-2-yl)ethyl)-1H-tetrazol-5-amine (Exp. = 507 K [39])
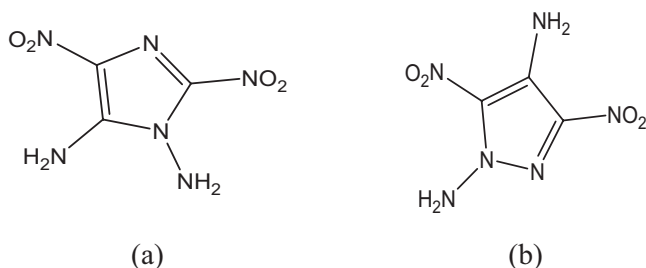
predicted decomposition temperatures of 500 (Dev = –28 K), 504 (Dev = –46 K), and 468 K (Dev = –39 K), respectively. The predicted results of these compounds showed negative large deviations as compared to the experimental data. Thus, it is possible to adjust the estimated results of the core model by a suitable correcting function because they have the following similar molecular fragments (Scheme 1).
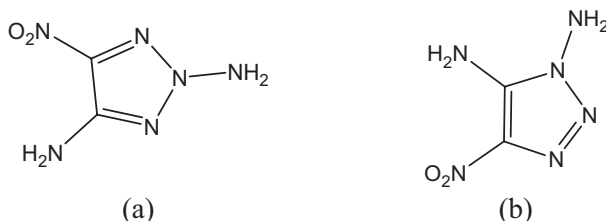

Scheme 1.

For azole derivatives, the positions of the nitrogen atoms with similar substituents are important because the core correlation can predict the same result. For example, 2,4-dinitro-1H-imidazole-1,5-diamine (Scheme 2, (a)) and 3,5-dinitro-1H-pyrazole-1,4-diamine (Scheme 2, (b)) show different decomposition temperatures, *i.e.* Exp. = 463 K [43] and Exp. = 514 K [44], respectively.

(a)   (b)

Scheme 2.

The position of the substituents may give different decomposition temperatures. For instance, the change of position of the amino groups in 5-nitro-2H-1,2,3-triazole-2,4-diamine (Scheme 3, (a)) and 4-nitro-1H-1,2,3-triazole-1,5-diamine (Scheme 3, (b)) give different experimental values, *i.e.* Exp. = 485 K [43] and Exp. = 467 K [43], respectively.

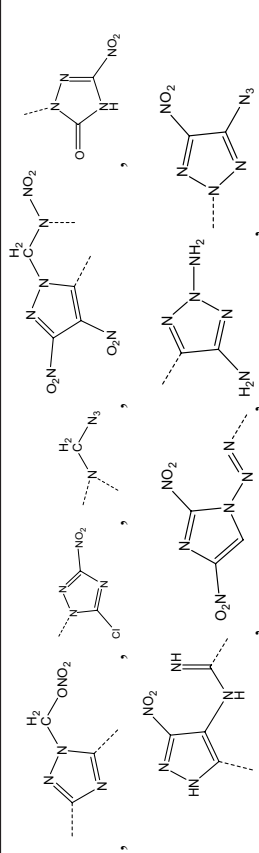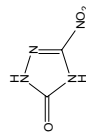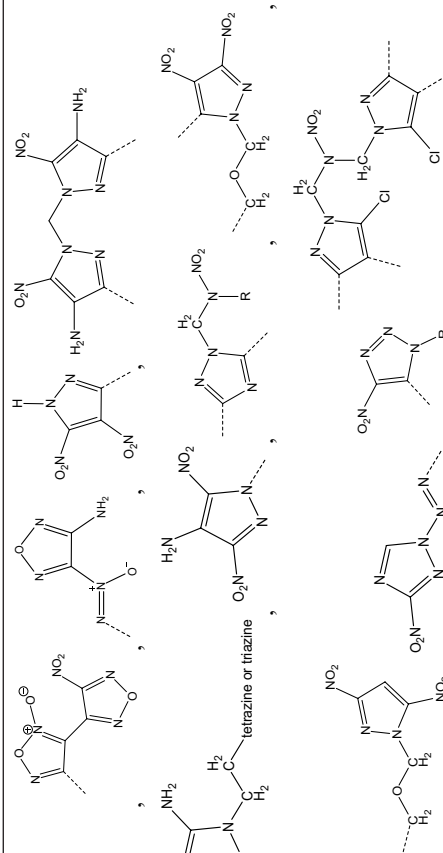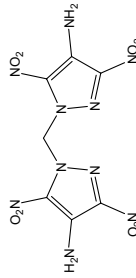(a)   (b)

Scheme 3.

Thus, it is essential to use correcting functions, which are related to non-additive molecular moieties. The study of molecular structures of various heterocyclic aromatic compounds with large deviations shows that it is possible to consider two non-additive descriptors for improvement of the reliability of the initial model. The reported data given by the training set (in SI, see Table S1) were used to derive the improved model using the MLR method [40] as follows:

$$T_d = 485.3 + 34.32 T_{d,Elem} + 25.41 T_{d,Add} - 45.80 T_{d,Non-add} + 62.14 T_d^{in} - 62.91 T_d^{dec} \quad (5)$$

where T_d is in K; T_d^{in} and T_d^{dec} are two correcting parameters. These two new terms in Equation 5 can adjust the large negative and positive deviations of the core correlation from the experimental data, respectively. Two steps should be done to specify the molecular fragments responsible for these deviations and the best values for T_d^{in} and T_d^{dec}. They are demonstrated as follows:

a) *Specification of molecular moieties*: Large deviations of the core correlation from experimental data may be due to different structural parameters. They include the type and position of substituents, the type of heteroaromatic rings, the number of substituents, the attachment of specific substituents on the carbon or nitrogen atoms of the ring, and the type of attachment of the heteroaromatic rings to each other through specific groups. These structural parameters have been specified for different compounds in the training set. Table 1 shows six classes of molecular moieties for T_d^{in} and T_d^{dec} that can be used to adjust moderate, large, and very large negative and positive deviations, respectively. There are also six examples for six classes in Table 1. The coefficients of T_d^{in} and T_d^{dec} have positive and negative signs, respectively. Thus, the presence of three classes of molecular moieties corresponding to T_d^{in} and T_d^{dec} can increase and decrease the underestimated and overestimated results of the core correlation, respectively.

b) *Optimization of T_d^{in} and T_d^{dec}*: The second step is to find the optimum values for T_d^{in} and T_d^{dec}, which are given in Table 1. They were obtained by the regression method to give the lowest and the highest values of RMSE and R^2, respectively.

Table 1. The optimized values for T_d^{in} and T_d^{dec}

Type of compound	T_d^{in}	T_d^{dec}	Example
	0.5	0	
	0.7	0	

Table 1.　Continuation

Type of compound	T_d^{in}	T_d^{dec}	Example
	1.0	0	
	0	0.5	

Table 1. Continuation

Type of compound	T_d^{in}	T_d^{dec}	Example
	0	0.7	
	0	1.1	

3   Results and Discussion

3.1 Statistical evaluation of the core and the improved correlations

Table 2 lists the statistical evaluations of the three and five variables given in Equations 1 and 5, with several statistical parameters including standard error (SE), P-value, and confidence interval (CI). As indicated in Table 2, the coefficients of the variables in both equations had low values of SE as compared to corresponding coefficients, which confirmed that the variables in these correlations are significant [45]. Moreover, the P-values of all coefficients were less than 0.05. Thus, there are no random variations for the effects of the variables, which provided a highly significant impact. The values of SE for Equations 1 and 5 were 36.3 and 12.9 K, respectively, which confirmed the higher precision of Equation 5. The values of F statistic and significance F of Equation 5 were 588.5 and 1.6E–85, respectively. These are lower than the calculated values for Equation 1, *i.e.* 87.0 and 9.4E–31, respectively. These data confirmed the higher reliability of Equation 5 as compared to the core correlation. Table 3 provides the relationships among the five descriptors of Equation 5 using the inter-correlation coefficients. The correlation coefficients provide information about the interrelationship between the descriptors. As seen in Table 3, there is no significant correlation between these descriptors ($P < 0.05$).

Table 2.   Regression coefficients and their SDs, P-values, and CIs of Equations 1 and 5

Model	Descriptor	Coeffi-cients	SE	t-statistic	P-value[a]	Lower bound[b] (95%)	Upper bound[c] (95%)
Eq. 1	Intercept	490.2	4.3	114.6	7.0E–131	481.8	498.7
	$T_{d,Elem}$	33.21	2.65	12.55	4.8--24	27.97	38.45
	$T_{d,Add}$	24.63	2.64	9.33	4.1E–16	19.41	29.85
	$T_{d,Non-add}$	−52.82	5.94	−8.89	4.8E–15	−64.58	−41.07
Eq. 5	Intercept	485.3	2.1	236.6	9.8E–169	481.2	489.3
	$T_{d,Elem}$	34.32	0.95	36.16	2.2E–68	32.44	36.20
	$T_{d,Add}$	25.41	0.95	26.64	1.3E–53	23.52	27.30
	$T_{d,Non-add}$	−45.80	2.15	−21.34	1.1E–43	−50.04	−41.55
	T_d^{in}	62.14	3.53	17.59	9.7E–36	55.15	69.13
	T_d^{dec}	−62.91	3.71	−16.96	2.6E–34	−70.26	−55.57

[a] The probability of rejecting the null hypothesis; [b] Lower limit of a 95% CI; [c] Upper limit of a 95% CI

Table 3. Correlation coefficients matrix of the coefficients of Equation 5

Coefficient	$T_{d,Elem}$	$T_{d,Add}$	$T_{d,Non-add}$	T_d^{in}	T_d^{dec}
$T_{d,Elem}$	1	–	–	–	–
$T_{d,Add}$	0.4345	1	–	–	–
$T_{d,Non-add}$	0.2015	−0.5510	1	–	–
T_d^{in}	−0.4947	0.0306	−0.1124	1	–
T_d^{dec}	0.4973	0.1104	−0.1443	−0.1442	1

Table 4 provides further statistical parameters containing:
– AD (average deviation),
– AAD (average absolute deviation),
– AD_{max} (absolute maximum deviation),
– AARD% (average absolute relative deviation), and
– R^2

for the predicted results of 132 data of the training set (in SI, Table S1). For azole derivatives, the outputs of the core and the improved correlations have been compared with the outputs of the model of Zohari *et al.* [18], the best available model for comparison. The calculated values of AD, AD_{max}, AAD, and AARD% show bias and systematic errors, the worst error, the average size of the errors, and judgment of the extent or importance of errors, respectively. As seen in Table 4, the calculated values of AD, AD_{max}, AAD, and AARD% of Equation 5 were lower than the corresponding values of Equation 1, which confirmed the higher reliability of Equation 5. In contrast, the value of R^2 of Equation 5 was higher than that of Equation 1. The reliability of the core and the improved models were also compared with the model of Zohari *et al.* [18] for 116 azole compounds of the training set, which are also listed in Table 4. As seen, the mentioned statistical parameters of Equation 5 are better than those of the method of Zohari *et al.* [18]. Since the predicted values of AD for Equation 5 and the model of Zohari *et al.* [18] are 0.03 and 2.38 K, respectively, Equation 5 is unbiased without systematic errors. The ratio of the worst error by the model of Zohari *et al.* [18] to that of Equation 5 was 2.59, because their AD_{max} values were 75.00 and 25.00 K, respectively. The ratio of the average size of the errors of the model of Zohari *et al.* [18] to that of Equation 5 was 1.97, because their values of AAD were 18.70 and 9.48 K, respectively. Since the values of AARD% of Equation 5 and the model of Zohari *et al.* [18] were 2.01 and 3.93 K, the extent of the error of Equation 5 is less than that of the model of Zohari *et al.* [18]. The foundations of the mentioned statistical parameters are consistent with the higher value of R^2 of Equation 2 (0.962) as compared to that of the model of Zohari *et al.* [18] (0.821). For the azole

derivatives, it was confirmed that Equation 5 is unbiased with sufficient accuracy and precision, as compared to the model of Zohari *et al.* [18]. Moreover, the outputs of the statistical parameters of the training set in Table 4 have confirmed that Equation 5 provides a good prediction, not only for azole compounds but also for heterocyclic aromatic compounds in general, and some of their salts.

Table 4. Comparison of AD, AD_{max}, AAD, AARD% and R^2 values of Equations 1 and 5 for all data of the training set (132 heterocyclic aromatic compounds), as well as for 116 data of the training set corresponding to azole derivatives, with those of the model of Zohari *et al.* [18]

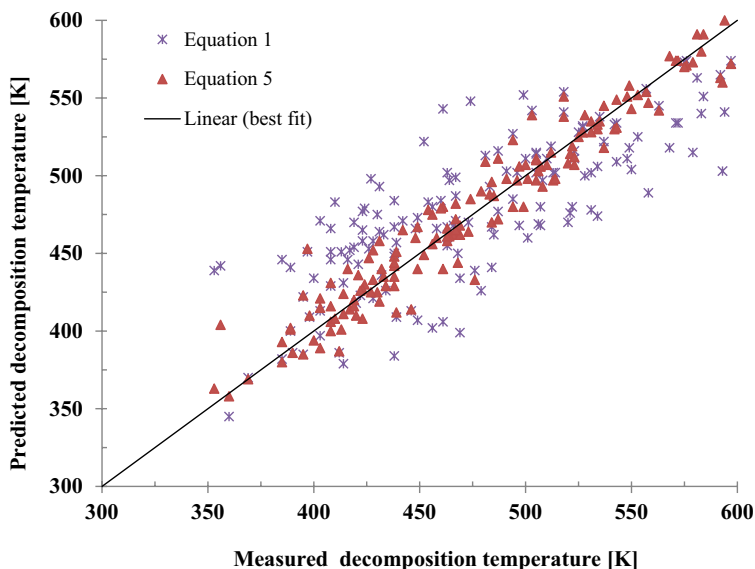| Compounds	Model	AD [K]	AD_{max} [K]	AAD [K]	AARD%	R^2
Heterocyclic aromatic	Equation 1	−0.01	86.00	29.10	6.17	0.671
	Equation 5	−0.01	33.00	9.72	2.05	0.959
Azole	Equation 1	−0.92	86.00	29.60	6.31	0.688
	Equation 5	0.03	29.00	9.48	2.01	0.962
	Zohari *et al.* [18]	2.38	75.00	18.70	3.93	0.821

The outputs of Equations 1 and 5 for heterocyclic aromatic compounds and some of their salts, as well as of the model of Zohari *et al.* [18], for azole derivatives are listed in Table S2 (in SI). Table 5 compares the calculated results of AD, AD_{max}, AAD, AARD%, and R^2 of Equations 1 and 5 for all data of the test set (49 heterocyclic aromatic compounds) as well as for 41 data corresponding to azole derivatives, with the model of Zohari *et al.* [18]. As shown in Table 5, the mentioned statistical parameters have confirmed that Equation 5 has high reliability for heterocyclic aromatic compounds and some of their salts, as well as the model of Zohari *et al.* [18], for azole compounds. In contrast to the training set, it was shown that the calculated data for AD and AD_{max} of Equation 1 were lower than those predicted by the model of Zohari *et al.* [18]. Moreover, the calculated R^2 of Equation 1 was close to the model of Zohari *et al.* [18]. Thus, Equation 5 gives reliable results for both the training and test sets of heterocyclic aromatic compounds and some of their salts, as well as the model of Zohari *et al.* [18] for azole derivatives.
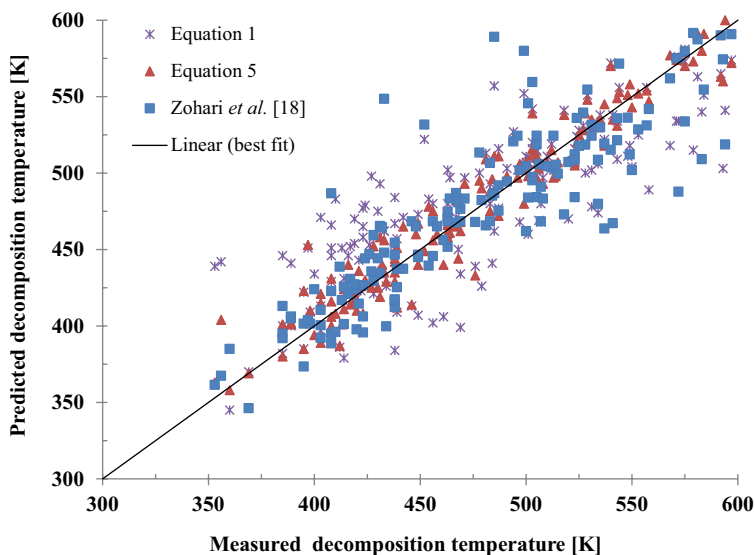
Table 5. Comparison of AD, AD_{max}, AAD, AARD% and R^2 of Equations 1 and 5 for all data of the test set (49 heterocyclic aromatic compounds) as well as 41 data of the test set corresponding to azole derivatives, with the model of Zohari *et al.* [18]

Compounds	Model	AD [K]	AD_{max} [K]	AAD [K]	AARD% [%]	R^2
Heterocyclic aromatic	Equation 1	9.27	86.00	24.29	5.30	0.729
	Equation 5	5.67	56.00	15.96	3.46	0.886
Azole	Equation 1	2.97	86.00	26.49	5.88	0.653
	Equation 5	2.36	56.00	16.10	3.55	0.879
	Zohari *et al.* [18]	19.78	104.1	23.03	4.67	0.664

Figure 1 shows a visual comparison of the predicted data of Equations 1 and 5 for all data of heterocyclic aromatic compounds and some of their salts of the training and test sets. As indicated in Figure 1, the outputs of Equation 5 are closer to the solid line than Equations 1. For the 165 azole compounds given in the training and test sets, Figure 2 compares the predicted results of these compounds by Equations 1 and 5, as well as by the model of Zohari *et al.* [18]. As seen in Figure 2, the predicted results of Equation 5 are closer to the line than the model of Zohari *et al.* [18] and Equation 1. As seen in Figures 1 and 2, the calculated results of Equation 5 are closer to the solid line than the model of Zohari *et al.* [18] and Equation 1 because the solid line shows an exact agreement between the outputs and experiment. Thus, the outputs of Equation 5 for both the training set and test set are given in Tables 4 and 5, which confirm that it may be presented as a reliable method for estimation of the decomposition temperature of different classes of heterocyclic aromatic compounds and some of their salts. Moreover, Equation 5 can be implemented into a suitable computer code such as EMDB [46].

Figure 1. The predicted decomposition temperature using Equations 1 and 5 *vs.* experimental values for 181 data of heterocyclic aromatic compounds and some of their salts in the training and test sets

Figure 2. The predicted decomposition temperature using Equations 1 and 5, as well as using the method of Zohari *et al.* [18] *vs.* experimental values for 157 data of azole derivatives in the training and test sets

3.2 Leave-one-out and 5-fold internal cross-validation

Cross-validation (CV), such as leave-one-out (LOO) and 5-fold is a further validation technique because the result of R^2 alone can be misleading. It can be used to verify internal predictivity, in which different ratios of compounds are iteratively held-out from the training set [47]. A single data point is removed each time in LOO CV. Meanwhile, the whole training set is divided in a 5-fold CV, randomly into five groups, as well as each group being left out in turn. The correlation coefficient of CV (Q^2), Q^2_{LOO} for LOO CV, and Q^2_{5CV} for the 5-fold CV, is executed a large number of times (of the order of 100) by choosing dissimilar grouping arrangements where $Q^2 < R^2$ [48]. It is important to have robust models in which the predicted values of Q^2 are close to R^2 [49]. Moreover, the proposed model is predictive if $Q^2 > 0.5$ and $R^2 > 0.6$ [50]. Table 6 compares the predicted values of Q^2_{LOO} and Q^2_{5CV}, as well as their RMSE values, with the corresponding values of the training set. As seen in Table 6, the values of Q^2_{LOO} and Q^2_{5CV} of Equation 1 were 0.651 and 0.620 (the average of 100 runs), respectively. For Equation 5, the values of Q^2_{LOO} and Q^2_{5CV} were improved to 0.955 and 0.951, respectively. Hence, the calculated Q^2 is smaller than R^2 for both Equations 1 and 5. Equation 5 gives a reliable predictive model because its values of Q^2_{LOO}, Q^2_{5CV}, and R^2 are significantly greater than those of Equation 1. Since the values of Q^2_{LOO} and Q^2_{5CV} of Equation 5 are near to the value of R^2, Equation 5 can be introduced as a well-behaved, robust, and not over-fitted model. Since the ratio of the RMSE values of Equation 5 to Equation 1 is less than one-half for Q^2_{LOO}, and Q^2_{5CV}, Equation 5 has higher reliability as compared to Equation 1.

Table 6. The values of Q^2_{LOO}, Q^2_{5CV} and R^2 as well as RMSE for Equations 1 and 5

Parameter	Number of compounds	R^2 or Q^2 for Equation		RMSE for Equation	
		1	5	1	5
Training	132	0.671	0.959	35.76	12.63
LOO CV[a]	131	0.651	0.955	36.87	13.25
5-fold CV[a]	127	0.620	0.951	36.11	13.26

[a] For 100 runs

4　Conclusions

♦　Two new correlations, as core and improved models, have been presented to estimate the thermal decomposition temperature (onset) of heterocyclic aromatic compounds and some of their salts. The core correlation depends on $T_{d,Elem}$, $T_{d,Add}$, and $T_{d,Non-add}$. Meanwhile, the improved correlation includes T_d^{in} and T_d^{dec} as two correcting parameters, which can be used to adjust large deviations of the core thermal decomposition temperatures from the experimental data.

♦　The values of SE, P-value, and CI have been calculated to assess the statistical evaluations for the variables in the two correlations. Further statistical parameters including AD, AAD, AD_{max}, AARD%, and R^2 for the predicted results of the training and test sets have been used to confirm the higher reliability of Equation 5. For azole derivatives, these statistical parameters were applied to confirm the higher reliability of Equation 5 for the corresponding compounds in both the training and test sets, as compared to the best available method, *i.e.* Zohari *et al.* [18].

♦　The values of Q^2_{LOO}, Q^2_{5CV}, and R^2 as well as RMSE for Equations 1 and 5 were used to demonstrate that Equation 5 is a robust, well-behaved, and not over-fitted correlation.

♦　Thus, Equation 5 can provide the best prediction of the decomposition temperature of heterocyclic aromatic compounds containing various derivatives of five or six-membered nitrogen-containing rings, triazoles, tetrazoles, furazans, triazines, and tetrazines, and some of their salts.

References

[1]　Agrawal, J.P. *High Energy Materials: Propellants, Explosives and Pyrotechnics*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, **2010**; ISBN: 978-3-527-32610-5.

[2]　Keshavarz, M.H.; Klapötke, T.M. *The Properties of Energetic Materials: Sensitivity, Physical and Thermodynamic Properties*. Walter de Gruyter GmbH, Berlin/Boston, **2017**.

[3]　Keshavarz, M.H.; Klapötke, T.M. *Energetic Compounds: Methods for Prediction of Their Performance.* 2nd ed., Walter de Gruyter GmbH, Berlin/Boston, **2020**; ISBN 978-3-11-067764-5.

[4]　Klapötke, T.M. *Energetic Materials Encyclopedia.* 2nd ed., Walter de Gruyter GmbH, Berlin/Boston, **2021**; ISBN 978-3-11-067245-9.

[5]　Klapötke, T.M. *Chemistry of High-Energy Materials.* 5th ed., Walter de Gruyter GmbH, Berlin/Boston, **2019**; ISBN 978-3-11-062438-0.

[6]   Farhadian, A.H.; Tehrani, M.K.; Keshavarz, M.H.; Karimi, M.; Darbani, S.M.R.; Rezayi, A.H. A Novel Approach for Investigation of Chemical Aging in Composite Propellants Through Laser-Induced Breakdown Spectroscopy (LIBS). *J. Therm. Anal. Calorim.* **2016**, *124*(1): 279-286.

[7]   Keshavarz, M.H.; Karimi, M.; Goodarzi, E.; Hosseini, S.H. The Use of the Change of Elongation for Comparison of the Shelf Life of Composite Solid Propellants in the Air and Nitrogen Atmospheres. *Z. Anorg. Allg. Chem.* **2021**, *647*(7): 696-703.

[8]   Ahmadi, S.H.; Keshavarz, M.H.; Atabak, H.R.H. Correlations between Laser Induced Breakdown Spectroscopy (LIBS) and Dynamical Mechanical Analysis (DMA) for Assessment of Aging Effect on Plastic Bonded Explosives (PBX). *Z. Anorg. Allg. Chem.* **2019**, *645*(2): 120-125.

[9]   Zohari, N.; Keshavarz, M.H.; Seyedsadjadi, S.A. A Novel Method for Risk Assessment of Electrostatic Sensitivity of Nitroaromatics Through Their Activation Energies of Thermal Decomposition. *J. Therm. Anal. Calorim.* **2014**, *115*(1): 93-100.

[10] Zohari, N.; Keshavarz, M.H.; Seyedsadjadi, S.A. A Link between Impact Sensitivity of Energetic Compounds and Their Activation Energies of Thermal Decomposition. *J. Therm. Anal. Calorim.* **2014**, *117*(1): 423-432.

[11] Keshavarz, M.H.; Mousaviazar, A.; Hayaty, M. A Novel Approach for Assessment of Thermal Stability of Organic Azides through Prediction of Their Temperature of Maximum Mass Loss. *J. Therm. Anal. Calorim.* **2017**, *129*(3): 1659-1665.

[12] Fathollahi, M.; Sajady, H. QSPR Modeling of Decomposition Temperature of Energetic Cocrystals Using Artificial Neural Network. *J. Therm. Anal. Calorim.* **2018**, *133*(3): 1663-1672.

[13] Prana, V.; Rotureau, P.; André, D.; Fayet, G.; Adamo, C. Development of Simple QSPR Models for the Prediction of the Heat of Decomposition of Organic Peroxides. *Mol. Inf.* **2017**, *36*(10) paper 1700024: 1-8.

[14] Mathieu, D.; Alaime, T.; Beaufrez, J. From Theoretical Energy Barriers to Decomposition Temperatures of Organic Peroxides. *J. Therm. Anal. Calorim.* **2017**, *129*(1): 323-337.

[15] Wang, B.; Yi, H.; Xu, K.; Wang, Q. Prediction of the Self-Accelerating Decomposition Temperature of Organic Peroxides Using QSPR Models. *J. Therm. Anal. Calorim.* **2017**, *128*(1): 399-406.

[16] Zohari, N.; Keshavarz, M.H.; Dalaei, Z. Prediction of Decomposition Onset Temperature and Heat of Decomposition of Organic Peroxides Using Simple Approaches. *J. Therm. Anal. Calorim.* **2016**, *125*(2): 887-896.

[17] Ghani, K.; Keshavarz, M.H.; Jafari, M.; Khademian, F. A Novel Method for Predicting Decomposition Onset Temperature of Cubic Polyhedral Oligomeric Silsesquioxane Derivatives. *J. Therm. Anal. Calorim.* **2018**, *132*(1): 761-770.

[18] Zohari, N.; Abrishami, F.; Zeynali, V. Prediction of Decomposition Temperature of Azole-based Energetic Compounds in Order to Assess of Their Thermal Stability. *J. Therm. Anal. Calorim.* **2020**, *141*: 1453-1463.

[19] Keshavarz, M.H.; Pouretedal, H.R.; Saberi, E. A New Method for Predicting Decomposition Temperature of Imidazolium-based Energetic Ionic Liquids. *Z. Anorg. Allg. Chem.* **2017**, *643*(2): 171-179.

[20] Rajaei, A.; Jafari, M.; Ghani, K. A Novel Method for Predicting Decomposition Onset Temperature of High-Energy Metal-Organic Frameworks. *J. Therm. Anal. Calorim.* **2020**, *142*: 1295-1302.

[21] Keshavarz, M.H. *Combustible Organic Materials. Determination and Prediction of Combustion Properties*. Walter de Gruyter GmbH, Berlin/Boston, **2018**; ISBN 978-3-11-057220-9.

[22] Badgujar, D.; Talawar, M.; Zarko, V.E.; Mahulikar, P.P. Recent Advances in Safe Synthesis of Energetic Materials: An Overview. *Combust. Explos. Shock Waves* **2019**, *55*(3): 245-257.

[23] Keshavarz, M.H.; Jafari, M.; Ebadpour, R. Recent Advances for Assessment of the Condensed Phase Heat of Formation of High-Energy Content Organic Compounds and Ionic Liquids (or Salts) to Introduce a New Computer Code for Design of Desirable Compounds. *Fluid Phase Equilib.* **2021**, *533*: 112913.

[24] Manafi Moghadam, M.; Zamani, M. Performance of NO_2-rich Multifunctionalized C_{60} Derivatives as New High-Energy-Density Nanomaterials. *Int. J. Quantum Chem.* **2021**, *121*(5) paper e26504.

[25] *NIST Chemistry WebBook*. National Institute of Standards and Technology, **2021**, http://webbook.nist.gov.

[26] Mathieu, D. Sensitivity of Energetic Materials: Theoretical Relationships to Detonation Performance and Molecular Structure. *Ind. Eng. Chem. Res.* **2017**, *56*(29): 8191-8201.

[27] Keshavarz, M.H.; Abadi, Y.H.; Esmaeilpour, K.; Damiri, S.; Oftadeh, M. A Novel Class of Nitrogen-rich Explosives Containing High Oxygen Balance to Use as High Performance Oxidizers in Solid Propellants. *Propellants Explos. Pyrotech.* **2017**, *42*(10): 1155-1160.

[28] Tang, J.; Cheng, G.; Feng, S.; Zhao, X.; Zhang, Z.; Ju, X.; Yang, H. Boosting Performance and Safety of Energetic Materials by Polymorphic Transition. *Cryst. Growth Des.* **2019**, *19*(8): 4822-4828.

[29] Badgujar, K.C.; Badgujar, V.C.; Bhanage, B.M. A Review on Catalytic Synthesis of Energy Rich Fuel Additive Levulinate Compounds from Biomass Derived Levulinic Acid. *Fuel Process. Technol.* **2020**, *197* paper 106213: 1-19.

[30] Zhang, W.; Zhang, J.; Deng, M.; Qi, X.; Nie, F.; Zhang, Q. A Promising High-Energy-Density Material. *Nat. Commun.* **2017**, *8*(1): 181-188.

[31] Singh, A.; Sharma, T.C.; Kumar, M.; Narang, J.K.; Kishore, P.; Srivastava, A. Thermal Decomposition and Kinetics of Plastic Bonded Explosives Based on Mixture of HMX and TATB with Polymer Matrices. *Def. Technol.* **2017**, *13*(1): 22-32.

[32] Keshavarz, M.H.; Moradi, S.; Saatluo, B.E.; Rahimi, H.; Madram, A.R. A Simple Accurate Model for Prediction of Deflagration Temperature of Energetic Compounds. *J. Therm. Anal. Calorim.* **2013**, *112*(3): 1453-1463.

[33] Keshavarz, M.H.; Abadi, Y.H.; Esmaeilpour, K.; Damiri, S.; Oftadeh, M. Novel High-Nitrogen Content Energetic Compounds with High Detonation and Combustion Performance for Use in Plastic Bonded Explosives (PBXs) and Composite Solid Propellants. *Cent. Eur. J. Energ. Mater.* **2018**, *15*(2): 364-375.

[34] Gutowski, Ł.; Cudziło, S. Synthesis and Properties of Novel Nitro-based Thermally Stable Energetic Compounds. *Def. Technol.* **2020**, *17*(3): 775-784.

[35] Oliveira, M.A.S.; Oliveira, R.S.S.; Borges, I. Quantifying Bond Strengths via a Coulombic Force Model: Application to the Impact Sensitivity of Nitrobenzene, Nitrogen-rich Nitroazole, and non-Aromatic Nitramine Molecules. *J. Mol. Model.* **2021**, *27*(3): 1-17.

[36] Jiao, F.; Xiong, Y.; Li, H.; Zhang, C. Alleviating the Energy & Safety Contradiction to Construct New Low Sensitivity and Highly Energetic Materials through Crystal Engineering. *CrystEngComm* **2018**, *20*(13): 1757-1768.

[37] Bu, R.; Xiong, Y.; Zhang, C. π–π Stacking Contributing to the Low or Reduced Impact Sensitivity of Energetic Materials. *Cryst. Growth Des.* **2020**, *20*(5): 2824-2841.

[38] Liu, W.-H.; Zeng, W.; Qin, H.; Jiang, C.-L.; Liu, F.-S.; Tang, B.; Lei, Y.-X.; Liu, Q.-J. First-principle Calculations of Electronic, Vibrational, and Thermodynamic Properties of 1,3-Diamino-2,4,6-trinitrobenzene. *J. Mol. Model.* **2019**, *25*(12): 1-14.

[39] Yin, P.; Zhang, Q.; Shreeve, J.M. Dancing with Energetic Nitrogen Atoms: Versatile N-Functionalization Strategies for N-Heterocyclic Frameworks in High Energy Density Materials. *Acc. Chem. Res.* **2016**, *49*(1): 4-16.

[40] He, P.; Pan, Y.; Jiang, J.-C. Prediction of the Self-Accelerating Decomposition Temperature of Organic Peroxides Based on Support Vector Machine. *Procedia Eng.* **2018**, *211*: 215-225.

[41] Keshavarz, M.H.; Maghsoodi, N.K.; Shokrollahi, A. A Reliable Model for Assessment of Melting Points of Cyclic Hydrocarbons Containing Complex Molecular Structures, Isomers and Stereoisomers. *Fluid Phase Equilib.* **2020**, *521*: 112692.

[42] Keshavarz, M.H.; Zakinejad, S.; Esmailpour, K. An Improved Simple Method for Prediction of Entropy of Fusion of Energetic Compounds. *Fluid Phase Equilib.* **2013**, *340*: 52-62.

[43] He, C.; Yin, P.; Mitchell, L.A.; Parrish, D.A.; Shreeve, J.M. Energetic Aminated-Azole Assemblies from Intramolecular and Intermolecular N–H···O and N–H···N Hydrogen Bonds. *Chem. Commun.* **2016**, *52*: 8123-8126.

[44] He, C.; Zhang, J.; Parrish, D.A.; Shreeve, J.M. 4-Chloro-3,5-dinitropyrazole: A Precursor for Promising Insensitive Energetic Compounds. *J. Mater. Chem. A* **2013**, *1*(8): 2863-2868.

[45] Schober, P.; Boer, C.; Schwarte, L.A. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth. Analg.* **2018**, *126*(5): 1763-1768.

[46] Keshavarz, M.H.; Klapötke, T.M.; Sućeska, M. Energetic Materials Designing Bench (EMDB), Version 1.0. *Propellants Explos. Pyrotech.* **2017**, *42*(8): 854-856.

[47] Gramatica, P. Principles of QSAR Modeling: Comments and Suggestions from Personal Experience. *Int. J. Quant. Struct.-Prop. Relat.* **2020**, *5*(3): 1-37.

[48] Achary, P.G.R.; Toropova, A.P.; Toropov, A.A. Prediction of the Self-Accelerating Decomposition Temperature of Organic Peroxides. *Process Saf. Prog.* **2020**, paper 12189.

[49] Fayet, G.; Rotureau, P. Development of Simple QSPR Models for the Impact Sensitivity of Nitramines. *J. Loss Prev. Process Ind.* **2014**, *30*: 1-8.

[50] Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inf.* **2010**, *29*: 476-488.