	QMRF identifier (JRC Inventory): **To be entered by JRC**
	QMRF Title: **QSPR model for the impact sensitivity of nitramines using constitutional descriptors**
	Printing Date: **25 mars 2020**

1.QSAR identifier

1.1.QSAR identifier (title):

QSPR model for the impact sensitivity of nitramines using constitutional descriptors

1.2.Other related models:

1.3.Software coding the model:

2.General information

2.1.Date of QMRF:

25/03/2020

2.2.QMRF author(s) and contact details:

[1]Guillaume FAYET INERIS guillaume.fayet@ineris.fr

[2]Patricia ROTUREAU INERIS patricia.rotureau@ineris.fr

2.3.Date of QMRF update(s):

18/10/2017 - Version 1

25/03/2020 - Version 2

2.4.QMRF update(s):

25/03/2020 - Guillaume FAYET - corrections of the software name in sections 4.4, 4.5 and 4.6.

2.5.Model developer(s) and contact details:

[1]Guillaume FAYET INERIS guillaume.fayet@ineris.fr

[2]Patricia ROTUREAU INERIS patricia.rotureau@ineris.fr

2.6.Date of model development and/or publication:

2014

2.7.Reference(s) to main scientific papers and/or software package:

Guillaume Fayet, Patricia Rotureau, Development of simple QSPR models for the impact sensitivity of nitramines, Journal of Loss Prevention in the Process Industries, 30 (2014) 1-8 http://www.sciencedirect.com/science/article/pii/S0950423014000576

2.8.Availability of information about the model:

This model is non proprietary. Training and validation sets are available.

2.9.Availability of another QMRF for exactly the same model:

3.Defining the endpoint - OECD Principle 1

3.1.Species:

3.2.Endpoint:

QMRF 1. Physical Chemical Properties QMRF 1.21. Explosive Properties

3.3.Comment on endpoint:

The endpoint is related to explosive properties, namely the impact sensitivity. This property, characterizing the tendency of a material to react under the effect of a mechanical impact, is one of the most

important to classify explosible substances according to the Classification, Labelling and Packaging of substances and mixtures (CLP regulation) and according to the Transport of Dangerous Goods (International regulation).

3.4.Endpoint units:

cm - logarithmic unit

3.5.Dependent variable:

log h50% - logarithm of the impact sensitivity (in cm)

3.6.Experimental protocol:

The impact sensitivity, obtained by drop weight impact test, measures the height (in cm) from which the fall of a weight of a given mass on a sample causes a reaction with 50% probability level.

3.7.Endpoint data quality and variability:

Data are taken from a single reference to ensure consistency of experiments [ref 2]. This reference collects data which were obtained from the works of Kamlet and Adolph [ref. 3-4].

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSPR - Multilinear regression

4.2.Explicit algorithm:

log h50% = 0.94 + 86.3 nC=O/Mw – 0.017 OB + 0.14 nC-O-C – 0.21 nC=O

4.3.Descriptors in the model:

[1]nC=O/Mw Number of C=O fragments on molecular weight

[2]OB Oxygen Balance according to TDG definition: OB = -1600(2 nC + 0.5 nH - nO)/Mw

[3]nC-O-C Number of C-O-C fragments

[4]nC=O Number of C=O fragments

4.4.Descriptor selection:

A total of 64 constitutional descriptors were calculated using Codessa software and including some descriptors from examination of chemical structures. Then, the best multi-linear regression algorithm (BMLR) implemented in the CodessaPro software was used to select descriptors. The first step in the BMLR analysis is to reduce the initial set of descriptors by rejecting all descriptors with insignificant variance.

At the same time, if two descriptors are highly intercorrelated, only the one that has the strongest correlation with the property is retained. After that, starting from pairs of orthogonal descriptors, higher rank models are computed by successively including orthogonal descriptors as soon as an increase in correlation (based on Fisher coefficient F) is observed. Then, the BMLR analysis selects the best models at each rank.The final model is chosen by expert knowledge considering the statistics of the model, the descriptors included in the model (chemical interpretation) and the quality of experimental data.

For more details about descriptors and the BMLR method, see ref 1.

4.5.Algorithm and descriptor generation:

Multiple linear regression was computed by Codessa software to generate the model. Descriptors were calculated from examination of chemical structures.

4.6.Software name and version for descriptor generation:

Codessa 2.7.16

selection of descriptors and multilinear regression calculation

http://www.semichem.com/codessa/default.php

4.7.Chemicals/Descriptors ratio:

60 chemicals / 4 descriptors = 15

The ratio for the training set is 40/4 = 10

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The applicability domain of the model is defined in terms of:

- chemical family: nitramines compounds, i.e. presenting a N-NO2 bond;
- chemical structure space of the training set based on the descriptors of the model using the method described in section 5.2;
- range of property values in the training set.

Only one molecule of the validation set is out of the applicability domain: bis-(2,2,2-trinitroethyl)-nitramine.

5.2.Method used to assess the applicability domain:

The applicability domain of the model was built, for each descriptor and for the property, by the range of values represented among the molecules of the training set. The AD ranges for each descriptor and for the property are available in section 5.4.

5.3.Software name and version for applicability domain assessment:

5.4.Limits of applicability:

Ranges of values defining the applicability domain:

- nC=O: [0 ; 0.01]
- OB: [-59.1 ; 10.4]
- nC-O-C: [0 ; 2]
- nC=O: [0 ; 2]

-log h50%: [0.70 ; 2.51]

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

6.2.Available information for the training set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: Yes

MOL file: Yes

6.3.Data for each descriptor variable for the training set:

All

6.4.Data for the dependent variable for the training set:

All

6.5.Other information about the training set:

The training set contains 2/3 of the data, i.e. 40 molecules. Selection of molecules in both set was performed by random selection. No bias of representativity was identified in the property distribution or in terms of chemical diversity (assessed by a principal component analysis and an inspection of chemical structures).

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

n=40; R^2=0.876; MAE=0.11(log)=8.4%; RMSE=0.15(log)

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

Q^2LOO=0.825

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

$Q^2$10CV=0.822; $Q^2$5CV=0.802

6.10.Robustness - Statistics obtained by Y-scrambling:

500 iterations; Average R^2(YS)=0.103 ; Standard Deviation of R^2(YS)=0.069

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

Yes

7.2.Available information for the external validation set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: Yes

MOL file: Yes

7.3.Data for each descriptor variable for the external validation set:

All

7.4.Data for the dependent variable for the external validation set:

All

7.5.Other information about the external validation set:

The training set contains 1/3 of the data, i.e. 20 molecules.

7.6.Experimental design of test set:

Selection of molecules in both set was performed by random selection.

7.7.Predictivity - Statistics obtained by external validation:

n=20; R^2=0.903; MAE=0.08(log)=6.4%; RMSE=0.14(log); Q^2F1=0.857; Q^2F2=0.857; Q^2F3=0.928; CCC=0.934

In AD: n=19; R^2=0.880; MAE=0.08(log)=6.4%; RMSE=0.14(log); Q^2F1=0.823; Q^2F2=0.823; Q^2F3=0.916; CCC=0.917

7.8.Predictivity - Assessment of the external validation set:

No bias of representativity was identified in the property distribution or in terms of chemical diversity (assessed by a principal component analysis and an inspection of chemical structures).

7.9.Comments on the external validation of the model:

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

The model is focused on nitramines that are known to decompose following a same decomposition mechanism involving the breaking of the N-NO2 bond [refs 5-6].

In the equation, the most significant descriptor (as evidenced in terms of t-test) is the Oxygen Balance which is a preselection criterion in the classification of explosives. So its occurence into the model is relevant.

8.2.A priori or a posteriori mechanistic interpretation:

Constitutional descriptors were calculated using the Codessa software on the one hand and including additional descriptors from examination of chemical structures on the other hand (like Oxygen Balance).

Besides, all the descriptors included in the model, developed by a statistical approach, were additional descriptors. This enforces a posteriori the relevance of this approach.

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

[1]G. Fayet, P. Rotureau, Development of simple QSPR models for the impact sensitivity of nitramines, Journal of Loss Prevention in the Process Industries, 30 (2014) 1-8 http://www.sciencedirect.com/science/article/pii/S0950423014000576

[2]C.B. Storm, J.R. Stine, J.F. Kramer, Sensitivity relationships in energetic materials. Chemistry and physics of energetic materials, Kluwer Academic Publishers, 1990.

[3]M.J. Kamlet, The relationship of impact sensitivity with structure of organic high explosives. I. Polynitroaliphatic explosives, In Proceedings of the Sixth Symposium (International) on Detonation, Coronado, California, (1976) 312-322.

[4]H.G. Adolph, J.R. Holden and D.A Chicra, Relationships between the impact sensitivity of high energy compounds and some molecular properties which determine their performance, NSWC TR 80-495, White Oak, MD, (1981).

[5]D. Chakraborty, R.P. Muller, S. Dasgupta, A detailed model for the decomposition of nitramines: RDX and HMX. Journal of Computer-Aided Materials Design, 8 (2001) 203-212. http://link.springer.com/article/10.1023/A:1020074113000

[6]N.E. Ermolin,V.E. Zarko, Mechanism and kinetics of the thermal decomposition of cyclic nitramines, Combustion, Explosion and Shock Waves, 33 (1997) 251-269 http://link.springer.com/article/10.1007/BF02671865

9.3.Supporting information:

Training set(s)

train-mol.zip	file:///C:\Users\Fayet\Documents\WORK\Modèles QSPR\Proposition Modèles Toolbox\QMRF nitramines\fichiers out\train-mol.zip
train-data.pdf	file:///C:\Users\Fayet\Documents\WORK\Modèles QSPR\Proposition Modèles Toolbox\QMRF nitramines\train-data.pdf

Test set(s)

valid-mol.zip	file:///C:\Users\Fayet\Documents\WORK\Modèles QSPR\Proposition Modèles Toolbox\QMRF nitramines\fichiers out\valid-mol.zip
valid-data.pdf	file:///C:\Users\Fayet\Documents\WORK\Modèles QSPR\Proposition Modèles Toolbox\QMRF nitramines\valid-data.pdf

Supporting information

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

To be entered by JRC

10.2.Publication date:

To be entered by JRC

10.3.Keywords:

To be entered by JRC

10.4.Comments:

To be entered by JRC