Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information

Which molecular properties determine the impact sensitivity of an explosive? A machine learning quantitative investigation of nitroaromatic explosives†

Julio Cesar Duarte,[a,b] Romulo Dias da Rocha,[b] Itamar Borges Jr., [a,c,*]

[a] *Departamento de Engenharia de Computação, Instituto Militar de Engenharia.*
[b] *Programa de Pós-Graduação em Engenharia de Defesa, Militar de Engenharia.*
[c] *Departamento de Química, Militar de Engenharia.*

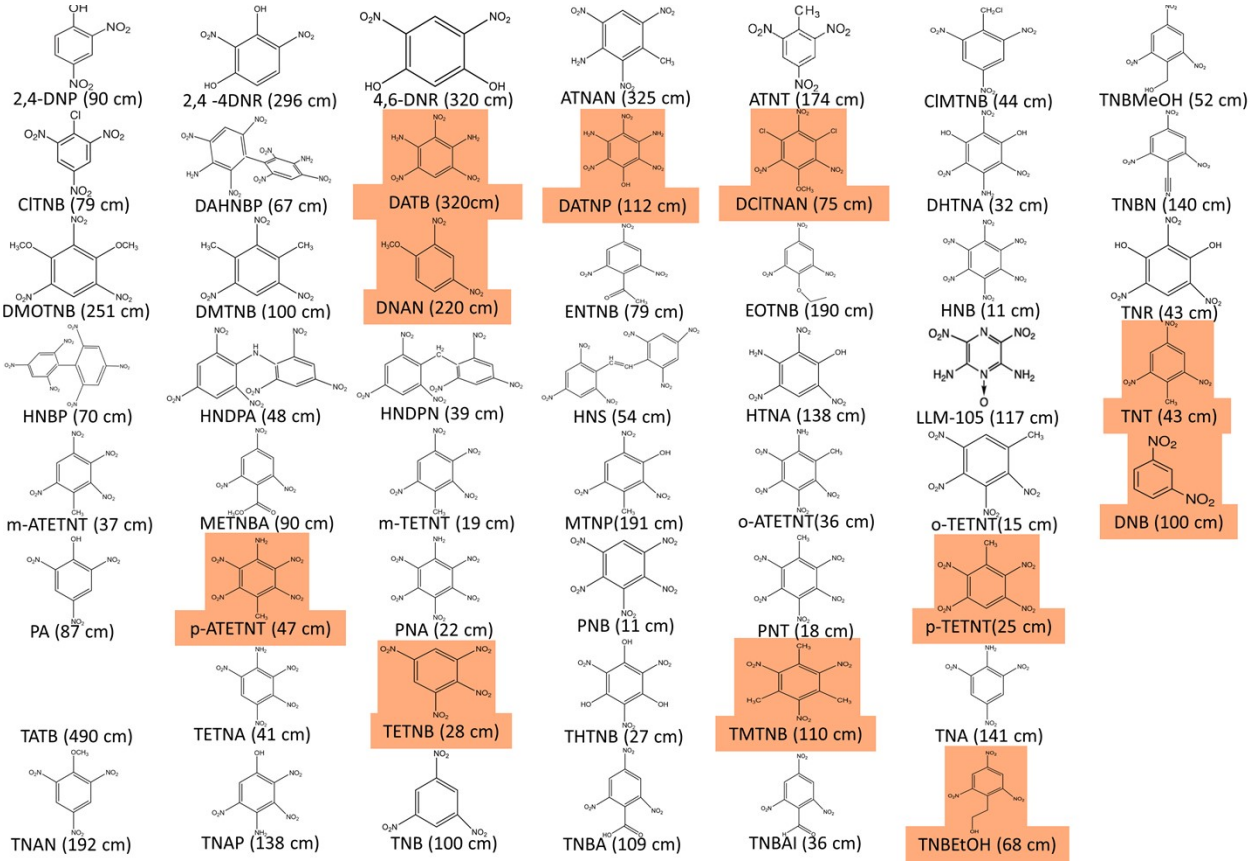
Instituto Militar de Engenharia, Praça General Tibúrcio, 80, Urca, Rio de Janeiro (RJ), 22290-270, Brazil: E-mail: itamar@ime.eb.br

Scheme 1S – The 53 investigated nitroaromatic molecules. The acronyms of the molecules are identified in Table 1S of this Electronic Supplementary Information. The colored molecules comprise the test set.

Table 1S – Acronyms and names of the 53 molecules.

Table 2S – Input data for the algorithms: the four features (molecular properties) and the target-features (h_{50} values).

Table 3S – The 42 different algorithms tested employing the Lazy Predict Tool, and the respective root mean square errors (RMSE).



Scheme 1S – The 53 investigated nitroaromatic molecules. The acronyms of the molecules are identified in Table 1S of this Electronic Supplementary Information. The colored molecules comprise the test set.

Table 1S – Acronyms and names of the 53 molecules.

2,4-DNP - 2,4-dinitrophenol
2,4-DNR – 2,4-dinitroresorcinol
4,6-DNR – 4,6-dinitroresorcinol
ATNAN – 3-amino-2,4,6-trinitroanisole
ATNT – 3-amino-2,4,6-trinitrotoluene
ClMTNB – Chloromethyl-2,4,6-trinitrobenzene
ClTNB – Chloro-2,4,6-trinitrobenzene
DAHNBP –3,3'-diamino-2,2',4,4',6,6'-hexanitro-biphenyl
DATB – 1,3-diamino-2,4,6-trinitrobenzene
DATNP – 3,5 diamino- 2,4,6-trinitro-phenol
DClTNAN – 3,5-dichloro-2,4,6-trinitroanisole
DHTNA – 3,5-dihydroxy-2,4,6-trinitroaniline
DMOTNB – 1,3-dimethoxy-2,4,6-trinitrobenzene
DMTNB – 1,3-dimethyl-2,4,6-trinitrobenzene
DNAN – 2,4-dinitroanisole
DNB – 1,3-dinitro-benzene
ENTNB – 2,4,6-trinitrophenylethanone
EOTNB – Ethoxy-2,4,6-trinitrobenzene
HNB – Hexanitrobenzene
HNBP – 2,2',4,4',6,6'-hexanitrobiphenyl
HNDPA – 2,2',4,4'6,6'- hexanitrodiphenylamine (dipicrylamine)
HNDPM – 2,2',4,4',6,6'- hexanitrodiphenylmethane
HNS – 2,2',4,4',6,6'- hexanitrostilbene
HTNA – 3-hydroxy-2,4,6-trinitroaniline
LLM-105 – 2,6-diamino-3,5-dinitropyrazine-1-oxide
m-ATETNT – 3-aminotetranitrotoluene
METNBA – 3-methylester-2,4,6-trinitrobenzoic acid
m-TETNT – 2,3,4,6-tetranitrotoluene
MTNP – 3-methyl-2,4,6-trinitrophenol
o-ATETNT – 2-aminotetranitrotoluene
o-TETNT – 2,3,4,5-tetranitrotoluene
PA – 2,4,6-trinitrophenol (picric acid)
p-ATETNT – 4-aminotetranitrotoluene
PNA – Pentanitroaniline
PNB – Pentanitrobenzene
PNT – Pentanitrotoluene
p-TETNT – 4-aminotetranitrotoluene
TATB – 1,3,5-triamino-2,4,6-trinitrobenzene
TETNA – 2,3,4,6-tetranitroaniline
TETNB – 1,2,3,5-tetranitrobenzene

THTNB – 1,3,5-trihydroxy-2,4,6-trinitrobenzene
TMTNB – 1,3,5-trimethyl-2,4,6-trinitrobenzene
TNA – 2,4,6-trinitroaniline
TNAN – 2,4,6-trinitroanisole
TNAP – 4-amino-2,3,5-trinitrophenol
TNB – 1,3,5-trinitrobenzene
TNBA – 2,4,6-trinitrobenzoic acid
TNBAl – 2,4,6-trinitrobenzaldehyde
TNBEtOH – 2,4,6-trinitrobenzeneethanol
TNBMeOH – 2,4,6-trinitrobenzenemethanol
TNBN – 2,4,6-trinitrobenzonitrile
TNR – 2,4,6-trinitroresorcinol
TNT – 2,4,6-trinitrotoluene

Table 2S – Input data for the algorithms: the four features (molecular properties) and the target-features (h_{50} values).

Molecule	$\sum Q_0(NO_2)$	$\sum Q_1(NO_2)$	$\sum Q_2(C)$	$\#NO_2$	h_{50} **(cm)**
2,4-DNP	0.578	1.98	7.022	2	90
2,4-DNR	0.547	1.908	7.095	2	296
4,6-DNR	0.584	1.942	7.092	2	320
ATNAN	0.88	2.995	7.242	3	325
ATNT	0.934	2.998	7.127	3	174
CLMTNB	0.822	2.995	6.981	3	44
CLTNB	0.789	2.991	6.762	3	79
DAHNBP	1.293	5.354	13.233	6	67
DATB	1.011	2.963	7.343	3	320
DATNP	0.972	2.884	7.016	3	112
DCLTNAN	0.78	2.9	6.942	3	75
DHTNA	0.842	2.829	7.133	3	32
DMOTNB	0.785	2.994	7.069	3	251
DMTNB	0.89	2.974	6.925	3	100
DNAN	0.542	1.976	7.169	2	220
DNB	0.569	2.021	6.972	2	100
ENTNB	0.754	3.022	6.863	3	79
EOTNB	0.792	2.948	6.988	3	190
HNB	0.696	5.406	6.444	6	11
HNBP	1.385	5.774	12.737	6	70
HNDPA	1.486	5.798	13.227	6	48
HNDPM	0.978	5.479	13.228	6	39
HNS	1.47	5.944	13.78	6	54
HTNA	0.839	2.906	7.157	3	138
LLM-105	0.47	1.948	6.513	2	117
m-ATETNT	0.915	3.819	6.969	4	37
METNBA	0.688	2.918	6.684	3	90
m-TETNT	0.827	3.816	6.704	4	19
MTNP	0.803	2.952	6.927	3	191
o-ATET	0.784	3.758	6.976	4	36
o-TETNT	0.718	3.752	6.684	4	15
PA	0.739	2.899	6.86	3	87
PATETNT	0.907	3.816	6.981	4	47
PNA	0.835	4.615	6.811	5	22
PNB	0.73	4.589	6.511	5	11
PNT	0.806	4.632	6.608	5	18
p-TETNT	0.824	3.795	6.724	4	25
TATB	1.131	2.931	7.518	3	490

TETNA	0.866	3.796	6.935	4	41
TETNB	0.757	3.78	6.63	4	28
THTNB	0.694	2.776	6.979	3	27
TMTNB	0.934	2.95	6.99	3	110
TNA	0.833	2.975	7.091	3	141
TNAN	0.775	2.968	6.956	3	192
TNAP	0.761	2.89	7.032	3	138
TNB	0.765	2.97	6.799	3	100
TNBA	0.676	2.895	6.673	3	109
TNBAL	0.704	2.222	6.728	3	36
TNBEtOH	0.835	3.007	6.866	3	68
TNBMeOH	0.798	2.944	6.774	3	52
TNBN	0.658	2.882	6.667	3	140
TNR	0.711	2.854	6.941	3	43
TNT	0.83	2.989	6.849	3	98

Table 3S – The 42 different algorithms tested employing the Lazy Predict Tool, and the respective root mean square errors (RMSE).

Algorithms	RMSE (cm)
XGBRegressor	25.65
RandomForestRegressor	27.25
ExtraTreesRegressor	30.83
AdaBoostRegressor	31.68
BaggingRegressor	32.08
GradientBoostingRegressor	39.25
DecisionTreeRegressor	46.05
ExtraTreeRegressor	46.54
PoissonRegressor	58.02
KNeighborsRegressor	60.52
PassiveAggressiveRegressor	62.47
HuberRegressor	63.32
SGDRegressor	69.63
LinearRegression	70.00
TransformedTargetRegressor	70.00
Lars	70.00
OrthogonalMatchingPursuitCV	70.02
LassoLarsIC	70.03
Ridge	70.08
Lasso	70.18
BayesianRidge	70.51
LarsCV	70.84
LassoLarsCV	70.84
ElasticNetCV	71.12
LassoCV	71.82
LassoLars	71.9
RidgeCV	73.83
GammaRegressor	74.64
ElasticNet	76.33
NuSVR	76.70
RANSACRegressor	76.71
SVR	76.73
OrthogonalMatchingPursuit	78.76
GeneralizedLinearRegressor	78.79
TweedieRegressor	78.79
HistGradientBoostingRegressor	82.70
DummyRegressor	82.70
LGBMRegressor	82.70
KernelRidge	93.12
LinearSVR	95.37
MLPRegressor	109.34

GaussianProcessRegressor	3295.80