Prediction of the Impact Sensitivity by Neural Networks

Hedi Nefati, Jean-Michel Cense, and Jean-Jacques Legendre*

Unité de Recherche Modélisation Appliquée à la Chimie et aux Procédés, ENSCP 11, rue Pierre et Marie Curie 75005 Paris, France

Received June 22, 1995[⊗]

A method for optimizing the prediction of impact sensitivity of explosive molecules by neural networks is presented. The database consists of 204 molecules of known sensitivity, containing C, H, N, and O and belonging to several chemical families. Pertinent molecular descriptors are selected by a preliminary evolutionary multiple linear regression treatment, and the effects of the network's topology and the extent of the training are examined and optimized. The predictions are satisfactory with a correlation coefficient $R = 0.94$ obtained through cross-validation. The neural networks approach proves more accurate than linear methods and more general than all previously used methods.

INTRODUCTION

The impact sensitivity is a property of primary importance for the handling of explosive compounds and its prediction stimulated numerous studies during the last decades. These studies stated the influence of various molecular parameters such as the oxygen balance,[1−4] the molecular electronegativity,[5−7] the lengths of the trigger bonds,[8−10] the charge dissymmetry around these bonds in the fundamental state[11−13] and in excited states,[14−17] and the presence of specific groups.[1] However, no fully satisfactory answer has been provided up to now: all these studies were focused on small families of several tens of organic explosive compounds and did not take into account simultaneously all these compounds. Moreover, for most of these studies, the influence of the significant parameters was considered to be linear.

In order to improve these former results we have undertaken work whose main purpose was the building of a rapid and easy to use method for the prediction of the impact sensitivity of pure compounds containing carbon, hydrogen, oxygen, and nitrogen. This tool had to be used in industry for a safe handling of explosive compounds and as an assistance for the design of low sensitivity explosives. Since rapid predictions were needed, the use of sophisticated quantum chemistry modeling methods such as configurations interaction of excited states[14−17] were avoided even if they provide usually good results.

Consequently, we decided to use a purely empirical method based on the statistical treatment of a database containing molecules of known sensitivity. As the impact sensitivity is related to numerous complex phenomena, we did not expect to obtain a good linear correlation between molecular descriptors and impact sensitivity and therefore turned to the use of neural networks which are known to perform nonlinear regressions.

DATABASE

The whole database is extracted from the literature; it consists of 204 molecules containing C, H, N, and O whose sensitivity was measured by the same method developed by the Explosive Research Laboratory (U.S.A.)[18] providing the largest number of impact sensitivity measurements currently available. The impact sensitivity is measured by the logarithmic 50% impact height, log $H_{50\%}$. Molecules with log $H_{50\%} > 2.5$ ($H_{50\%} > 320$ cm) were discarded since their sensitivity is usually measured with a very low precision. These molecules belong to several families: 59 nitroaromatics, 103 nitroaliphatics, 17 nitrotriazoles, 5 nitropyridines, 6 nitroimidazoles, 3 nitrofurazanes, 3 nitropyrazoles, and 8 other nitro compounds.

MOLECULAR DESCRIPTORS

Following the literature we considered that the impact sensitivity of a pure compound mainly depends on the molecular structure of this compound, and we neglected the effects of molecular packing and the effect of eventual impurities. This approximation is valid because the impact sensitivity seems principally connected with the ability of the molecule to decompose when submitted to a mechanical shock. The way the decomposition energy is transmitted to the other molecules, which depends on the molecular neighboring, seems to have only a weak influence on this property for the compounds in our database.

In a preliminary we tested several types of molecular descriptors: number of occurrence of first and second neighbor fragments[19] and autocorrelation with atomic number or Pauling electronegativity.[20−22] As we did not obtain a good correlation with the impact sensitivity we decided to use all the molecular parameters used in previous works and to select from among them the most informative.

Two types of parameters were tested: topological parameters which depend only on the graph of the molecule and geometric or electronic parameters calculated with the semiempirical package MOPAC.[23] Potential conformations were first screened using the Discover software with CVFF force field,[24] and the selected conformations were then refined using the AM1 Hamiltonian.[25]

Altogether, 39 parameters (Table 1) were calculated for each of the 204 molecules. Among these parameters, 27 are purely topological parameters, and 12 are obtained from the MOPAC calculations. The definition of some of these parameters is recalled here:

[⊗] Abstract published in *Advance ACS Abstracts,* March 1, 1996.

S0095-2338(95)00223-X CCC: $12.00

Table 1. The 39 Descriptors Used for the Building of the Models[a]

parameter	description	parameter	description
1	oxygen balance	21	number of C≡N bonds
2	molecular electronegativity	22	number of C atoms
3	number of CO_2 groups	23	number of H atoms
4	number of $NO_2-C_{sp^2}$ bonds	24	number of N atoms
5	number of $NO_2-C_{sp^3}$ bonds	25	number of O atoms
6	number of NO_2−N bonds	26	100/molecular weight
7	number of NO_2−O bonds	27	indicator of aromaticity (0 or 1)
8	number of rings	28	sum of X−NO_2 charge dissymmetry
9	number of NH_2 groups	29	average of the X−NO_2 charge dissymmetry
10	number of OH groups	30	sum of X−NO_2 charges dissymmetry/mol weight
11	number of $C(NO_2)_3$ groups	31	length of the longest X−NO_2 bond
12	−CH in α of a nitroaromatic	32	length of the shortest X−NO_2 bond
13	indicator of symmetry (0 or 1)	33	highest potential for a X−NO_2 bond
14	number of −C═O groups	34	smallest potential for a X−NO_2 bond
15	number of Y−O−X groups	35	average potential for a X−NO_2 bond
16	number of C═C bonds	36	average length of the X−NO_2 bonds
17	number of C≡C bonds	37	heat of formation
18	number of C═N bonds	38	dipole
19	number of N≡N bonds	39	ionization potential
20	number of N═N bonds		

[a] See text for some definitions.

Oxygen balance: OB = 100 (2*n*O-*n*H-2*n*C-2*n*COO)/M, M is the molecular weight. Indicator of symmetry: 1 if the topological graph presents any symmetry element; otherwise 0. Molecular electronegativity: $\chi = \sum n_i/\sum(n_i/\chi_i)$, χ_i is the Pauling electronegativity of atoms of type i. X-NO_2 charge dissymmetry: $q_X - q_N$, q_X and q_N are the charges of X and N atoms. Potential of an X−NO_2 bond: $(q_X + q_N)/d$, d is the length of the X−N bond.

It must be noticed here that topological descriptors are calculated without approximation, that they do not depend on the conformation of the molecule, and that they are not highly intercorrelated. Inversely, geometric and electronic descriptors are more sensitive to conformations and to small variations of geometry and are frequently intercorrelated.

DESCRIPTORS SELECTION

The accuracy of a prediction depends on the ability of the model to represent reality. It is well-known that a model with too many parameters will fit correctly with the data base but will predict poorly new molecules. In order to detect the most pertinent combinations of descriptors we performed an evolutionary search of the descriptor space similar to the methods recently proposed[26−28] but we used cross-validated (leave one out) standard error of prediction (sep) as criterion[29] instead of the classical *s* value (root mean squared error) and *F* value (Fisher's-ratio) adopted there. We found many combinations of descriptors giving models of roughly equivalent predictive ability and selected among all these models those having the highest overall *F* value together with *F*-to-remove values higher than 3 for each descriptor.

Starting from the pool of 27 topological descriptors, an 11 parameter model (sep = 0.198, parameters: 1 4 5 6 7 12 20 23 24 25 26) was selected, whereas starting from the pool of all the 39 parameters, a 13 parameter model (sep = 0.181, parameters: 1 2 4 6 12 13 14 17 18 20 29 30 32) was selected. These two models are somewhat different, with only five common descriptors and among these common descriptors only one parameter (parameter 20, number of N═N bonds) exhibiting similar regression coefficient and similar partial *F* value in both models.

Starting from the 13 parameter model, we later found that a three component PLS[30] model gave a substantially better prediction (sep = 0.174), while starting from the 11 parameter model, we did not find a better PLS model. These results comforted us in the idea that we needed to introduce nonlinearity in our model in order to improve further the quality of the prediction. This was achieved by using back-propagation neural networks.

NEURAL NETWORKS

Neural networks are computer based emulations of the brain's information processing capability. The use of neural networks in chemistry has grown constantly since their first utilization for the prediction of protein secondary structure.[31] Chemists soon recognized that the characteristics of these neural networks were well adapted to the processing of data in which the relation between the cause and its effects could not be exactly defined. In the field of structure−activity or structure−property relationships the back-propagation neural networks, particularly the three layer networks, have gained wide acceptance so that the theory and practice of these networks is now well documented.[32,33]

We used a back-propagation neural network simulator already successfully employed for the prediction of partition coefficients (log *P*) of organic compounds.[34] We submitted the 11 parameter and the 13 parameter models to neural networks having 11 or 13 neurons in the input layer and one to four neurons in the hidden layer. A bias neuron was added to the input and hidden layers. Descriptors and target value were scaled between 0.1 and 0.9, the transfer function for neurons was the usual logistic function, the weights were initialized to small random values between −0.05 and +0.05, and weights were revised after each presentation of a molecule. The learning rate, initially set to 1, was gradually decreased during training, and no momentum term was used. In this paper, an iteration corresponds to the successive presentation of the whole database, which is frequently called

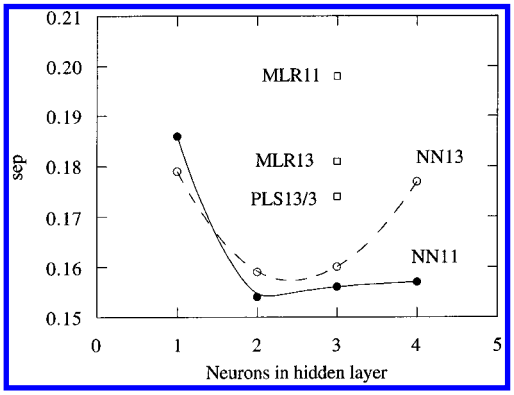

Figure 1. Evolution of the standard error of prediction (sep, leave one out method) with the number of neurons in hidden layer. NN11 is for the 11 topological descriptors, NN13 is for the 13 topological + geometrical + electronic descriptors. Sep for linear methods (MLR11, MLR13, and PLS13/3) are plotted for comparison.

an epoch by others. The neural network was trained for 3000 iterations, and all calculations were done on a Silicon Graphics R4000 using a software written in the C programming language.

The prediction accuracy was monitored by the cross-validation technique (leave one out) and assessed by the cross-validated R^2 (cvR^2), the standard error of prediction (sep), the Spearman rank correlation coefficient (srcc), Kendall's τ, and the number of outliers (molecules with | prediction−observation | > 0.3, i.e., a factor of 2 in the $H_{50\%}$ impact height). The predictions were repeated four times starting from different initial random sets of weights, and the mean prediction was used as the final value, a technique known to weaken the detrimental effects of local minima on the error surface.[35] No attempt was made to discard outliers as there is no recognized objective method to detect outliers with neural networks.

RESULTS AND DISCUSSION

All the models obtained are shown in Figure 1 and Table 2. Selected predictions are given in Table 3. A plot of predicted vs observed impact sensitivity (log $H_{50\%}$) for the best model (NN11.2 + NN13.2) is displayed in Figure 2. Clearly the best of the linear models is obtained with a three component PLS regression build on the 13 parameter model. Performing the PLS analysis on the original 39 descriptors did not give such a good model. This observation confirms that the selection of variables prior to PLS analysis largely affects the quality of the prediction.[26]

For neural network predictions the best models are those retaining only two neurons in hidden layers. These neural networks with two neurons in hidden layers give better prediction than those with more neurons in hidden layers despite the fact thatt these latter networks give slightly better fitting of the data set. In their study of dihydrofolate reductase inhibitors, Andrea and Kalayeh[36] pointed out that the ratio of the number of compounds to network connections (the so called ρ parameter) had to be higher than 2 in order to avoid over-fitting of the data. With our data the ρ parameter is much higher and, as the 11.2.1 and 13.2.1 neural networks have a connection to case ratio $\rho = 204/(12*2+3) = 7.6$ and $\rho = 204/(14*2+3) = 6.6$, we expect few risks of overtraining.[37,38]

The determination of the correct topology of the neural network is a lengthy process demanding one training phase

Table 2. Comparison of Predictive Ability for Multiple Linear Regression (MLR), Partial Least Square (PLS), Neural Networks (NN), and Combinations of These Methods[a]

method	sep	cv R^2	srcc	Kendall's τ	outliers
MLR11 (t)	0.198	0.799	0.910	0.745	16
MLR13 (t+g+e)	0.181	0.831	0.920	0.761	17
PLS13/3	0.174	0.844	0.926	0.767	17
NN13.4	0.177	0.840	0.926	0.770	13
NN13.3	0.163	0.864	0.937	0.785	13
NN13.2	0.159	0.870	0.937	0.786	12
NN13.1	0.179	0.835	0.919	0.759	18
NN11.4	0.157	0.874	0.938	0.782	11
NN11.3	0.156	0.876	0.938	0.785	11
NN11.2	0.154	0.878	0.941	0.788	11
NN11.1	0.186	0.822	0.911	0.738	23
NN11.2 + PLS13/3	0.151	0.884	0.943	0.794	7
NN11.3 + NN13.3	0.151	0.883	0.942	0.794	10
NN11.2 + NN13.2	0.149	0.886	0.946	0.801	10

[a] The 11 parameter models contain only topological descriptors, and the 13 parameter models contain topological, geometrical, and electronic parameters. Sep is the standard error of prediction, srcc is the Spearman rank correlation coefficient and outliers have |obs − pred| > 0.3. NN11.2 is for 11 neurons in the input layer, 2 neurons in the hidden layer, one bias unit being added in input and hidden layers. Neural networks are trained for 3000 iterations. Values given for neural networks are obtained by using the mean prediction of four runs (eight runs for the last two lines).

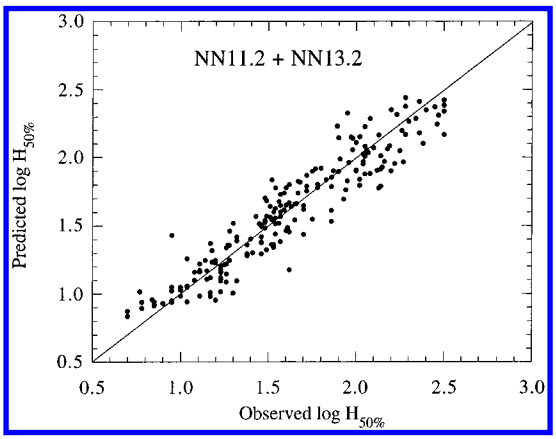

Figure 2. Predicted vs observed impact sensitivity using the mean value given by NN11.2 and NN13.2 neural networks.

for each molecule (leave one out), and it is not presently possible to perform the selection of the pertient descriptors directly with neural networks. In this context however, we observed that for all neural networks tested, the weights connecting the input neuron corresponding to parameter 20 (number of N═N bonds) to the hidden layer always had the highest magnitude. This is in agreement with the observation made when comparing multiple linear regression models that this parameter is the most informative.

Once the best neural network is found, predictions are very rapid, requiring only a forward propagation. The fact that purely topological descriptors give better prediction than topological + geometric + electronic descriptors strongly advocate the use of neural networks when a rapid estimation must be done because topological descriptors are calculated almost instantaneously and are not sensitive to the selected conformation.

Neural networks improve further the accuracy of the linear models by introducing nonlinear interaction terms between the descriptors. If this is a great advantage for interpolation of data, it is not clearly established that this advantage will

Table 3. Observed and Predicted Impact Sensitivity (log $H_{50\%}$) for MLR13, PLS13/3, and NN11.2 + NN13.2

code	name	obs	MLR	PLS	NN
al17	bis(2,2,2-trinitroethyl)nitramine	0.70	0.511	0.524	0.835
al19	*N*,*N*′-bis(2,2,2-trinitroethyl)MEDINA	0.70	0.614	0.636	0.873
al96	trinitroethylbis(trinitroethoxy)acetate	0.77	1.039	1.023	1.016
al18	2,2,2-trinitropropyl-3,3,3-trinitropropylnitramine	0.78	0.734	0.737	0.894
mi5	1-nitro-2,5-bis(trinitromethyl)pyrrolidine	0.78	0.871	0.874	0.940
al79	1,1,1,6,6,6-hexanitro-3-hexyne	0.84	0.811	0.790	0.957
al32	*N*-(2,2,2-trinitroethyl)nitraminoethylnitrate	0.85	0.923	0.915	0.922
al55	tris(2,2,2-trinitroethyl)orthoformate	0.85	0.943	0.949	0.942
al61	tetrakis(2,2,2-trinitroethyl)orthocarbonate	0.85	0.856	0.866	0.913
al49	1,1,1,3,5,5,5-heptanitropentane	0.90	0.876	0.870	0.931
al35	*N*-nitro-*N*-(3,3,3-trinitropropyl)2,2,2-trinitroethylcarbamate	0.95	0.904	0.902	0.952
al39	*N*,*N*′-dinitro-*N*,*N*′-bis(3,3,3-trinitropropyl)oxamide	0.95	1.045	1.033	1.048
tr21	4-nitro-1-picryl-1,2,3-triazole	0.95	1.359	1.354	1.430
al21	2,2,2-trinitroethyl-*N*-(2,2,2-trinitroethyl)nitraminoacetate	0.95	0.880	0.875	0.940
al16	methyl-2,2,2-trinitroethylnitramine	0.95	0.973	0.965	1.020
al26	1,1,1,3,6,6,9,11,11,11-decanitro-3,9-diazaundecane	1.00	1.007	1.030	1.048
al107	bis(trinitroethyl)-2,4-dinitrazapentanedioate	1.00	1.088	1.115	0.983
al120	1,9-dinitroxy-2,4,6,8-tetranitrazanonane	1.00	1.043	1.069	1.026
al41	1,1,1,5,7,10,14,14-nonanitro-3,12-dioxa-4,11-dioxo-5,7,10-triazatetradecane	1.04	0.992	0.994	1.059
al27	1,1,1,4,6,6,8,11,11,11-decanitro-4,8-diazaundecane	1.04	1.011	1.034	1.047
al99	trinitroethylcyanomethylnitrate	1.04	0.826	0.880	0.942
al20	*N*-methyl-*N*′-trinitroethylethylenedinitramine	1.04	1.286	1.294	1.256
al25	1,1,1,3,6,9,11,11,11-nonanitro-3,6,9-triazaundecane	1.08	0.991	1.026	1.099
al115	bis(trinitroethyl)-5,5-dinitro-2,8-dinitrazanonanedioate	1.08	1.201	1.195	1.160
al33	3-[*N*-(2,2,2-trinitroethyl)nitramino]propylnitrate	1.08	1.189	1.181	1.098
al1	methylenedinitramine	1.11	1.256	1.251	1.172
ar26	1-(2,2,2-trinitroethyl)-2,4,6-trinitrobenzene	1.11	1.142	1.124	1.218
al86	bis(trinitroethyl)oxamide	1.11	1.270	1.308	1.166
al111	bis(trinitroethyl)-2,4,6-trinitrazaheptanedioate	1.11	0.882	0.892	0.984
al113	*N*,*N*′-dinitromethylene-bis-(4,4,4-trinitro)butyramide	1.11	1.201	1.187	1.161
al95	bis(trinitroethyl)fumarate	1.14	1.413	1.396	1.246
al24	bis(2,2,2-trinitroethyl)-3-nitrazaglutarate	1.15	1.246	1.235	1.168
al10	*N*-3,3,5,5-pentanitropiperidine	1.15	1.155	1.170	1.106
al98	trinitroethylnitroguanidine	1.17	1.144	1.148	1.118
al112	bis(5,5,5-trinitro-3-nitrazapentanoyl)methylenedinitramine	1.17	0.829	0.846	1.011
al85	bis(trinitroethyl)oxalate	1.17	1.058	1.045	0.979
al87	trinitroethyl-2,2-dinitropropylcarbonate	1.17	1.448	1.438	1.369
al36	2,2,2-trinitroethyl-2,5-dinitrazahexanoate	1.18	1.405	1.406	1.317
ar1	2,3,4,5,6-pentanitroaniline	1.18	1.235	1.229	1.235
al109	trinitroethyl-5,5-dinitro-3-nitrazahexanoate	1.20	1.318	1.310	1.244
al116	1,4-bis-(5,5,5-trinitro-2-nitrazapentanoate)-2-butyne	1.20	1.228	1.230	1.227
al50	bis(2,2,2-trinitroethyl)carbonate	1.20	0.973	0.972	0.953
al100	*N*-nitro-*N*-methyl(trinitroethyl)carbamate	1.23	1.162	1.155	1.095
al114	bis(trinitroethyl)-2,5,8-trinitrazanonanedioate	1.23	1.189	1.196	1.204
al102	*N*-nitro-*N*-(trinitroethyl)glycinamide	1.23	1.274	1.264	1.179
al105	trinitropropyl(2,2,2-dinitropropyl)nitramine	1.23	1.206	1.210	1.162
al83	bis(trinitroethoxy)methane	1.23	1.086	1.089	1.016
ar112	pentanitrobenzene	1.23	1.065	1.062	1.105
al51	*N*,*N*′-bis(2,2,2-trinitroethyl)urea	1.23	1.216	1.203	1.120
al80	1,1,1,6,6,6-hexanitro-3-hexene	1.23	1.257	1.251	1.157
al88	*N*-trinitroethyl-4,4,4-trinitrobutyramide	1.25	1.322	1.298	1.213
al45	2,2,2-trinitroethylcarbamate	1.26	1.399	1.398	1.338
al38	2,2,2-trinitroethyl-2,4,6,6-tetranitro-2,4-diazaheptanoate	1.26	1.093	1.094	1.085
al64	*N*-(2,2,2-trinitroethyl)-3,3,5,5-tetranitropiperidine	1.26	1.344	1.342	1.220
al53	2,2,2-trinitroethyl-4,4,4-trinitrobutyrate	1.26	1.232	1.214	1.144
al03	trinitroethyl-*N*-ethyl-*N*-nitrocarbamate	1.27	1.465	1.457	1.359
al94	bis(1,1,1-trinitro-2-propyl)urea	1.28	1.581	1.612	1.459
al29	1,1,1,3,6,9,12,14,14,14-decanitro-3,6,9,12-tetrazatetradecane	1.28	1.072	1.131	1.245
al43	1,1,1,5,8,11,14,18,18,18-decanitro-3,16-dioxa-4,15-dioxo-5,8,11,14-tetrazaoctadecane	1.28	1.249	1.268	1.354
al119	nitroglycerine	1.30	1.268	1.228	1.003
py1	2,4,6-trinitropyridine-1-oxide	1.30	1.580	1.590	1.519
al34	3,5,5-trinitro-3-azahexylnitrate	1.32	1.520	1.512	1.416
ar27	1-(3,3,3-trinitropropyl)-2,4,6-trinitrobenzene	1.32	1.378	1.361	1.388
al110	*N*-nitro-*N*,*N*′-bis(trinitropropyl)-urea	1.32	1.135	1.127	1.096
al3	RDX	1.38	1.422	1.451	1.279
ar2	2,2,2-trinitroethyl-2,4,6-trinitrobenzoate	1.38	1.448	1.440	1.357
al89	1,5-bis-(trinitroethyl)biuret	1.38	1.426	1.402	1.292
tr19	4-nitro-1,2,3-triazole	1.40	1.708	1.719	1.403
al5	HMX	1.41	1.329	1.371	1.303
ar22	2,4,6-trinitrophloroglucinol	1.43	1.641	1.654	1.567
al46	methyl-2,2,2-trinitroethylcarbonate	1.45	1.569	1.568	1.514
al28	bis(2,2,2-trinitroethyl)4-nitraza-1,7-heptanedioate	1.46	1.572	1.556	1.500
al11	bis(2,2-dinitropropyl)nitramine	1.46	1.549	1.563	1.416
mi3	3,5-dinitroglycoluril	1.46	1.46	1.346	1.379

Table 3 (Continued)

code	name	obs	MLR	PLS	NN
al30	bis(2,2,2-trinitroethyl)3,6-dinitraza-1,8-octanedioate	1.46	1.372	1.372	1.293
mi7	*N*-(trinitropyl)-3,3,5,5-tetranitropiperidine	1.46	1.465	1.471	1.376
al75	2,2,4,4,6,6-hexanitroheptane	1.46	1.594	1.593	1.490
ar66	4,6-dinitrobenzofuroxan	1.48	1.654	1.657	1.703
al58	4,4,4-trinitrobutyric anhydride	1.48	1.602	1.575	1.481
al60	bis(2,2,2-trinitroethyl)succinate	1.48	1.609	1.584	1.538
mi1	1-ammonium-5-nitrotetrazole	1.48	1.935	1.937	1.523
ar28	1-(2,2,2-trinitroethyl)-2,4-dinitrobenzene	1.49	1.386	1.376	1.321
mi8	5-(3,5-diaminotrinitroanilino)tetrazole	1.49	1.874	1.878	1.681
ar60	1-(3,3,3-trinitropropyl)-2,4-dinitrobenzene	1.49	1.577	1.563	1.571
al72	bis(2,2,2-trinitroethyl)4,4,6,6,8,8-hexanitroundecanedioate	1.51	1.587	1.551	1.640
al8	tetryl	1.51	1.642	1.644	1.560
al47	1,1,1,3-teranitrobutane	1.52	1.421	1.412	1.361
ar104	5-aminostyphnic acid	1.52	1.827	1.835	1.836
py3	2,6-bis(picrylazo)-3,5-dinitropyridine	1.52	0.983	0.995	1.544
al56	1,1,1,7,7,7-hexanitroheptanone-2	1.53	1.447	1.416	1.342
al2	ethylenedinitramine	1.53	1.745	1.748	1.600
al91	1,1,1,7,7,7-hexanitroheptanone-4	1.53	1.498	1.477	1.361
tr29	5,6-dinitro-1-picrylbenzotriazole	1.54	1.522	1.535	1.437
tr30	1-picryl-4-picrylamino-1,2,3-triazole	1.54	1.590	1.598	1.555
tr7	4-(2-nitroethyl)-3,5-dinitro-1,2,4-triazole	1.54	1.814	1.788	1.775
al22	2,2,2-trinitroethyl-4-nitrazavalerate	1.54	1.689	1.689	1.626
al9	3,3-dinitro-1,5-pentanedinitramine	1.54	1.640	1.666	1.515
mi6	5-picrylaminotetrazole	1.56	1.701	1.712	1.517
ar29	2,4,6-trinitrobenzaldehyde	1.56	1.752	1.727	1.678
al15	2,2,5,7,7,10,12,12-octanitro-5,10-diazatridecane	1.57	1.562	1.586	1.564
ar64	bis(2,4,6-trinitrophenyl)diazine	1.57	1.377	1.378	1.382
im5	4,4′5,5′-tetranitrobiimidazole	1.57	1.697	1.682	1.600
al42	*N*,*N*′-dinitro-*N*,*N*′-bis(3,3-dinitrobutyl)oxamide	1.57	1.771	1.764	1.647
ar62	1,3,5-triazido-2,4,6-trinitrobenzene	1.57	1.685	1.752	1.731
ar61	hexanitrostilbene	1.59	1.889	1.874	1.740
ar51	2,2′,2″,4,4′,4″,6,6′,6″-nonanitro-*m*-terphenyl	1.59	1.872	1.863	1.738
al6	3-nitraza-1,5-pentanedinitramine	1.59	1.693	1.726	1.613
ar56	azobis-2,2′-4,4′,6,6′-hexanitrobiphenyl	1.60	1.453	1.469	1.460
ar24	3,3′-dihydroxy-2,2′,4,4′,6,6′-hexanitrobiphenyl	1.60	1.728	1.735	1.650
al82	4,4,4-trinitrobutyramide	1.60	1.752	1.739	1.777
ar4	2,3,4,6-tetranitroaniline	1.61	1.500	1.490	1.486
al104	trinitroethyl-2-methoxyethylnitrate	1.62	1.537	1.541	1.452
mi2	ethylenediammoniumdi-5-nitrotetrazole	1.62	1.121	1.277	1.175
ar33	2,4,6-trinitrobenzaldoxime	1.62	1.886	1.881	1.802
ar23	3-hydroxy-2,2′,4,4′,6,6′-hexanitrobiphenyl	1.62	1.713	1.711	1.668
ar3	2,4,6-trinitroresorcinol	1.63	1.675	1.687	1.639
al59	*N*,*N*′-bis(3,3,3-trinitropropyl)oxamide	1.65	1.744	1.713	1.659
ar5	2,2,2-trinitroethyl-3,5-dinitrosalicylate	1.65	1.580	1.578	1.542
im7	2,4-dinitro-1-picrylimidazole	1.66	1.708	1.695	1.662
ar102	5-(2-nitroguanidino)trinitro-1,3-benzenediamine	1.67	1.878	1.889	1.826
al40	2,2,6,9,9-pentanitro-4-oxa-5-oxo-6-azadecane	1.67	1.756	1.760	1.739
ar21	2,2′,4,4′,6,6′-hexanitrodiphenylamine	1.68	1.872	1.874	1.817
ar57	benzotrifuroxan	1.70	1.083	1.384	1.434
im3	ammonium-2,4,5-trinitroimidazole	1.70	1.611	1.576	1.621
al70	2,2-dinitropropane-1,3-diol-bis(4,4,4-trinitrobutyrate)	1.70	1.678	1.650	1.647
al7	3,6-dinitraza-1,8-octanedinitramine	1.72	1.872	1.913	1.787
ar32	2,4,6-trinitrophenylmethanol	1.72	1.881	1.878	1.869
ar35	3-methyl-2,2′,4,4′,6,6′-hexanitrobiphenyl	1.72	1.750	1.729	1.754
al74	2,2,4,6,6-pentanitroheptane	1.75	1.848	1.846	1.897
py2	3,3′,5,5′-tetranitro-2,2′-azopyridine	1.75	1.731	1.726	1.547
ar54	2,2′,2″,4,4″,6,6′,6″-octanitro-*p*-terphenyl	1.77	2.010	2.009	1.915
ar65	7-amino-4,6-dinitrobenzofuroxan	1.78	1.802	1.823	1.804
fu3	3-nitro-4-picrylaminofurazan	1.78	1.855	1.853	1.777
ar52	2,2″,4,4′,4″,6,6′,6″-octanitro-*m*-terphenyl	1.80	1.959	1.952	1.919
tr22	4-nitro-2-picryl-1,2,3-triazole	1.83	1.902	1.894	1.836
tr8	3-nitro-1-picryl-1,2,4-triazole	1.83	1.898	1.889	1.836
ar6	2,2,2-trinitroethyl-3,5-dinitrobenzoate	1.86	1.676	1.675	1.613
al13	2,2,4,7,9,9-hexanitro-4,7-diazadecane	1.86	1.609	1.632	1.530
al66	2,2-dinitropropyl-4,4,4-trinitrobutyramide	1.86	1.769	1.752	1.786
al69	*N*,*N*-bis(2,2-dinitropropyl)-4,4,4-trinitrobutyramide	1.86	1.797	1.776	1.855
ar105	5-formamidostyphnic acid	1.87	1.902	1.905	1.901
ar38	3,5-dimethyl-2,4,6-nitrophenol	1.89	2.163	2.155	2.228
al2	bis(2,2-dinitrobutyl)nitramine	1.90	1.911	1.929	1.893
al73	3,3,4,4-tetranitrohexane	1.90	2.005	2.003	2.144
al54	ethyl-2,2,2-trinitroethylcarbonate	1.91	1.876	1.875	1.987
ar9	2,2′,4,4′,6,6′-hexanitrobiphenyl	1.93	1.789	1.790	1.693
ar7	picric acid	1.94	1.751	1.757	1.761
ar42	methyl-2,4,6-trinitrobenzoate	1.95	2.218	2.217	2.326
al44	*N*,*N*′-dinitro-*N*,*N*′-bis(3-nitrazabutyl)oxamide	1.95	2.030	2.037	1.827

Table 3 (Continued)

code	name	obs	MLR	PLS	NN
py5	3,5-bis-(picrylamino)-2,6-dinitropyridine	1.96	2.025	2.016	1.964
ar108	dicyanopicramide	1.97	2.023	2.059	2.148
tr13	4-(2,4-dinitrobenzyl)-3,5-dinitro-1,2,4-triazole	1.98	2.183	2.163	2.140
fu7	3,5-bis-(picrylamino)-1,2,4-oxadiazole	1.98	2.135	2.130	2.054
mi4	1,4-dinitroglycoluril	2.00	1.881	1.935	1.899
ar49	1,4,5,8-tetranitronaphtalene	2.00	2.066	2.074	2.106
ar12	1,3,5-trinitrobenzene	2.00	1.824	1.823	1.907
ar58	1-dinitromethyl-3-nitrobenzene	2.02	1.780	1.791	1.795
im2	2,7-dinitroimidazole	2.02	1.683	1.671	1.840
im6	ammonium-4,4′5,5′-tetranitrobiimidazole	2.02	2.192	2.166	2.151
ar10	2,4,6-trinitrobenzoic acid	2.04	1.915	1.920	1.968
tr4	ammonium-3,5-dinitro-1,2,4-triazole	2.04	1.948	1.917	2.026
al81	2,2-dinitro-1,3-propanediol	2.04	1.859	1.895	1.952
pr2	4-niro-1-picrylpyrazole	2.05	1.975	1.972	2.007
ar43	hydroxy-3,5-diamino-2,4,6-trinitrobenzene	2.05	1.982	1.988	2.082
al62	methylene-bis(4,4,4-trinitrobutyramide)	2.05	1.894	1.864	1.877
al84	*N*-(2-propyl)trinitroacetamide	2.05	2.057	2.036	2.227
al4	*N*-methylethylenedinitramine	2.06	2.079	2.084	2.053
pr4	3,5-dinitro-1-methyl-4-picrylpyrazole	2.07	1.999	1.985	2.034
al63	ethylenebis(4,4,4-trinitrobutyrate)	2.08	1.852	1.835	1.881
fu4	3-amino-4-picrylaminofurazan	2.08	2.287	2.300	2.283
al52	5,5,5-trinitropentanone-2	2.10	1.939	1.922	2.069
ar25	3,3′-diamino-2,2′,4,4′,6,6′-hexanitrobiphenyl	2.12	1.933	1.935	1.904
ar48	ammonium picrate	2.13	2.048	2.046	2.163
ar36	3,3′-dimethyl-2,2′,4,4′,6,6′-hexanitrobiphenyl	2.13	1.842	1.830	1.775
al68	2,2,2-trinitroethyl-4,4-dinitrohexanoate	2.14	1.900	1.884	2.009
tr17	*N*,*N*′-dipicryl-5,5′-dinitro-3,3′-bi-1,2,4-triazole	2.14	2.131	2.095	1.790
ar8	2,4,6-trinitro-3-aminophenol	2.14	1.818	1.820	1.911
ar13	2,4,6-trinitrobenzonitrile	2.15	1.862	1.889	1.928
ar37	3-methyl-2,2′,4,4′,6-pentanitrobiphenyl	2.16	1.978	1.964	1.970
tr6	5,5′-dinitro-3,3′-bi-1,2,4-triazole	2.18	2.269	2.255	2.059
tr5	4-methyl-3,5-dinitro-1,2,4-triazole	2.19	1.978	1.952	2.084
ar31	trinitrotoluene	2.20	1.936	1.922	1.902
pr1	ammonium-3,5-dinitropyrazole	2.20	2.123	2.099	2.351
al106	1,7-dimethoxy-2,4,6-trinitrazaheptane	2.22	2.167	2.209	1.955
ar110	5-aminotrinitroisophthalamide	2.23	2.273	2.260	2.316
ar14	picramide	2.25	1.944	1.945	2.047
ar107	diamino-2,4,6-trinitrobenzonitrile	2.26	2.084	2.109	2.197
ar103	1,2-bis(3,5-diaminotrinitroanilino)ethane	2.27	1.978	2.004	1.966
py6	2,6-bis(picrylamino)pyridine	2.28	2.388	2.388	2.372
ar44	2,4,6-trinitrophenetole	2.28	2.345	2.350	2.438
ar17	2,4,6-trinitroanisole	2.28	2.070	2.076	2.168
tr26	2-picryl-1,2,3-triazole	2.30	2.330	2.336	2.264
ar109	3,5-diaminotrinitrobenzamide	2.34	2.210	2.210	2.286
tr12	3-amino-5-picrylamino-1,2,4-triazole	2.36	2.434	2.443	2.408
al76	bis(2,2-dinitropropyl)oxalate	2.36	2.094	2.089	2.179
tr16	3,5-bispicrylamino-1,2,4-triazole	2.38	2.202	2.198	2.103
ar18	1,3-dimethoxy-2,4,6-trinitrobenzene	2.40	2.203	2.212	2.348
al57	nitroisobutyl-4,4,4-trinitrobutyrate	2.45	2.182	2.178	2.371
tr2	4-nitro-1,2,4-triazole-5-one	2.46	2.006	2.028	2.243
ar16	2,4-dinitroresorcinol	2.47	2.107	2.133	2.307
im10	1-picrylimidazole	2.50	2.378	2.390	2.421
tr11	4-picrylamino-1,2,4-triazole	2.50	2.351	2.361	2.339
al97	*N*-nitro-*N*-methylformamide	2.50	2.144	2.128	2.382
ar20	1,3-diamino-2,4,6-trinitrobenzene	2.50	2.053	2.057	2.166

be so great for extrapolation of data. In this respect the fact that averaging the predictions of neural networks and PLS regression notably reduces the number of outliers is a hint to a way to temper the chaotic behavior of neural networks in extrapolation. These encouraging results will undoubtedly be ameliorated when a larger data base of more accurate measures of impact sensitivity will be available.

CONCLUSION

The impact sensitivity of 204 molecules of several different families has been successfully predicted using neural networks. The pertinent set of descriptors has been selected from a pool of 39 potential descriptors using an evolutionary linear regression technique. Reduced sets of 11 or 13 descriptors were provided to the neural network, giving predictions of much better quality than those obtained with linear regression.

Given the rather crude measure of impact sensitivity, we feel that these results fully confirm the interest of using neural network methods for structure−property studies.

REFERENCES AND NOTES

(1) Lothrop, W. C.; Handrick, C. R. The relationship between performance and constitution of pure organic explosive compounds. *Chem. Rev.* **1949**, *44*, 419−445.

(2) Kamlet, M. J. The Relationship of Impact Sensitivity with Structure of Organic High Explosives. I. Polynitroaliphatic Explosives. Proceedings of the 6th Symposium (International) on Detonation; San Diego, CA, 1976; pp 312−322.

(3) Kamlet, M. J.; Adolph, H. G. The Relationship of Impact Sensitivity with Structure of Organic High Explosives. II. Polynitroaromatic Explosives. *Propellants Explosives* **1979**, *4*, 30−34.
(4) Kamlet, M. J.; Adolph, H. G. Some comments regarding the sensitivities, thermal stabilities, and explosive performance characteristics of fluorodinitromethyl compounds. Proceedings of the 7th Symposium (International) on Detonation; Annapolis, MD, 1981; pp 84−92.
(5) Mullay, J. A. Relationship between Impact Sensitivity and Molecular Electronegativity. *Propellants, Explosives, Pyrotechnics* **1987**, *12*, 60−63.
(6) Mullay, J. Relationships between Impact Sensivity and Molecular Electronic Structure. *Propellants, Explosives, Pyrotechnics* **1987**, *12*, 121−124.
(7) Mullay, J. Atomic and Group Electronegativities. *J. Am. Chem. Soc.* **1984**, *106*, 5842−5847.
(8) Mullay, J. Calculation of Group Electronegativity *J. Am. Chem. Soc.* **1985**, *107*, 7271−7275.
(9) Mullay, J. A Simple Method for Calculating Atomic Charge in Molecules. *J. Am. Chem. Soc.* **1986**, *108*, 1770−1775.
(10) Kohno, Y.; Maekawa, K.; Tsuchioka, T.; Hashizume, T.; Imamura, A. A Relationship Between the Impact Sensitivity and Electronic Structures for the Unique N−N Bond in the HMX Polymorphs. *Combustion Flame* **1994**, *96*, 343−350.
(11) Politzer, P.; Seminario, J. M.; Bolduc, P. R. A proposed interpretation of the destabilizing effect of hydroxyl groups on nitroaromatic molecules. *Chem. Phys. Lett.* **1989**, *158*, 463−469.
(12) Murray, J. S.; Lane, P.; Politzer, P.; Bolduc, P. R. A relationship between impact sensitivity and the electrostatic potentials at the midpoints of C−NO_2 bonds in nitroaromatics. *Chem. Phys. Lett.* **1990**, *168*, 135−139.
(13) Politzer, P.; Murray, J. S.; Lane, P.; Sjoberg, P.; Adolph, H. G. Shock-sensitivity relationships for nitramines and nitroaliphatics. *Chem. Phys. Lett.* **1991**, *181*, 78−82.
(14) Delpuech, A.; Cherville, J. Relation entre la Structure Electronique et la Sensibilité au Choc des Explosifs Secondaires Nitrés. Critère Moléculaire de Sensibilité I. Cas des Nitroaromatiques et des Nitramines. *Propellants Explosives* **1978**, *3*, 169−175.
(15) Delpuech, A.; Cherville, J. Relation entre la Structure Electronique et la Sensibilité au Choc des Explosifs Secondaires Nitrés. Critère Moléculaire de Sensibilité II. Cas des esters nitriques. *Propellants Explosives* **1979**, *4*, 121−128.
(16) Delpuech, A.; Cherville, J. Relation entre la Structure Electronique et la Sensibilité au Choc des Explosifs Secondaires Nitrés. III. Influence de l'environnement cristallin. *Propellants Explosives* **1979**, *4*, 61−65.
(17) Delpuech, A.; Cherville, J.; Michaud, C. Molecular electronic structure and initiation of secondary explosives. Proceedings of the 7th Symposium (International) on Detonation; Annapolis, MD, 1981; pp 65−74.
(18) Storm, C. B.; Stine, J. R.; Kramer, J. F. Sensitivity relationships in energetic materials. In Chemistry and Physics of Energetic Materials; Bulusu S. N., Ed.; Kluwer Academic Publishers: Dordrecht, 1990; pp 605−609.
(19) Diawara, B.; Roullet, G.; Cense, J. M.; Legendre, J. J. Neuralchemist: un simulateur de réseaux de neurones à l'usage du chimiste. Proceedings of the 4th International Conference on Neural Networks and their Applications; **1991**, pp 777−781.
(20) Moreau, G.; Broto, P. The autocorrelation of topological structure: a new molecular descriptor. *Nouv. J. Chim.* **1980**, *4*, 359−360.
(21) Moreau, G.; Broto, P. Autocorrelation of molecular structures, application to SAR studies. *Nouv. J. Chim.* **1980**, *4*, 757−764.
(22) Broto, P.; Moreau, G.; Vandycke, C. Molecular structures: perception, autocorrelation descriptors and SAR studies. *Eur. J. Med. Chem.-Chim. Ther.* **1984**, *19*, 61−65.
(23) Stewart, J. J. P. MOPAC: A semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1.
(24) Biosym Technologies; San Diego, CA, U.S.A.
(25) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(26) Kubinyi, H. Variable selection in QSAR Studies. II. A Highly Efficient Combination of Systematic Search and Evolution. *Quant. Struct.-Act. Relat.* **1994**, *13*, 393−401.
(27) Luke, B. T. Evolutionary programming Applied to the Development of Quantitative Structure−Activity Relationships and Quantitative Structure−Property Relationships. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1279−1287.
(28) Rogers, D.; Hopfinger, A. J. Application of Genetic Function Approximation to Quantitative Structure−Property Relationships. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854−866.
(29) Cruciani, G.; Baroni, M.; Clementi, S.; Costantino, G.; Riganelli, D.; Skagerberg, B. Predictive Ability of Regression Models. Part I: Standard Deviation of Prediction Error. *J. Chemometrics* **1992**, *6*, 335−346.
(30) Geladi, P.; Kowalski, B. R. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* **1986**, *188*, 1−17.
(31) Qian, N.; Sejnowski, T. J. Predicting the Secondary Structure of Globular Proteins Using Neural Network Models. *J. Mol. Biol.* **1988**, *202*, 865−884.
(32) Gasteiger, J.; Zupan, J. Neural networks in chemistry. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 503−527.
(33) Burns, J.; Whitesides, G. M. Feed-Forward Neural Networks in Chemistry: Mathematical Systems for Classification and Pattern Recognition. *Chem. Rev.* **1993**, *93*, 2583−2601.
(34) Cense, J. M.; Diawara, B.; Legendre, J. J.; Roullet, G. Neural networks prediction of partition coefficients. *Chemometrics Intelligent Laboratory Systems* **1994**, *23*, 301−308.
(35) Hensen, L. K.; Salamon, P. Neural Networks Ensembles. *IEEE Trans. Pattern. Anal. Machine Intell.* **1990**, *12*, 993−1001.
(36) Andrea, T. A.; Kalayeh, H. Applications of Neural Networks in Quantitative Structure−Activity Relationships of Dihydrofolate Reductase Inhibitors. *J. Med. Chem.* **1991**, *34*, 2824−2836.
(37) Livingstone, D. J.; Manallack, D. T. Statistics Using Neural Networks-Chance Effects. *J. Med. Chem.* **1993**, *36*, 1295−1297.
(38) Manallack, D. T.; Ellis, D. D.; Livingstone, D. J. Analysis of Linear and Nonlinear QSAR Data Using Neural Networks. *J. Med. Chem.* **1994**, *37*, 3758−3767.

CI950223M