

Prediction of the impact sensitivity by neural networks

H. Nefati, J.M. Cense, J.J. Legendre, Christian Michot

► **To cite this version:**

H. Nefati, J.M. Cense, J.J. Legendre, Christian Michot. Prediction of the impact sensitivity by neural networks. 20. International Pyrotechnics Seminar (IPS 1994), Jul 1994, Colorado Springs, United States. pp.759-767. ineris-00971902

HAL Id: ineris-00971902

<https://hal-ineris.archives-ouvertes.fr/ineris-00971902>

Submitted on 3 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PREDICTION OF THE IMPACT SENSITIVITY BY NEURAL NETWORKS

Hedi NEFATI*, Jean-Michel CENSE*, Jean-Jacques LEGENDRE*
and Christian MICHOT**

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

** INERIS Parc Technologique ALATA
BP n°2 60550 - Verneuil en Halatte - France

ABSTRACT

A method for optimizing the prediction of impact sensitivity of explosive molecules by neural networks is presented.

The database we used consisted of 272 molecules containing C,H,N,O of known sensitivity and belonging to several chemical families.

Pertinent molecular descriptors were selected by a preliminary multilinear treatment.

The effects of the network's topology, the extent of the training, the choice of descriptors were examined and optimized.

The predictions are satisfactory with a correlation coefficient of 0.94 obtained through cross-validation. Moreover 95% of compounds are correctly classified in a 3-sensitivity scale and the remaining 5% are classified as ambiguous which is very encouraging for a real world implementation.

The neural networks approach proves more accurate and more general than previous methods.

I-INTRODUCTION :

The purpose of this work was the building of an efficient tool for the impact sensitivity prediction in order to minimize the risks during the handling of explosive compounds.

In previous works the field concerned separate families of compounds and revealed the influence of various molecular parameters such as the oxygen balance (Ref. 1) , the electronegativity (Ref. 2), the bond lengths (Ref. 3), the charge dissymmetry (Ref. 4), the presence of specific groups (Ref. 5),... Moreover, for most of these studies, the influence of these parameters was considered to be linear and taken into account separately which certainly induces a loss of precision for the predictions.

In order to overcome these difficulties, we decided to build an experimental data base as large as possible and we used data processing methods able to take into account simultaneously several non-linear correlations: the neural networks.

II-THE DATA BASE :

The initial data base was built from the literature. It consisted of 204 molecules $C_xH_yN_zO_t$ the sensitivity of which was measured by the same method developed by the Explosive Research Laboratory (USA) (Ref. 6). These molecules belong to nine families: nitroaromatics, nitroaliphatics, nitramines, nitric esters, nitrotriazoles, nitropyridines, nitroimidazoles, nitrofurazanes, others. The others 68 compounds were discarded because the sensitivity was measured with a different protocol or because they were clearly non explosive (musk).

The geometry of each of these molecules has been optimized using the semi-empirical program MOPAC (Ref. 7).

Some molecular parameters have been taken out from these quantum mechanics calculations and added to the list of the parameters which are the most frequently used in the literature.

Molecular Parameters

1: oxygen balance	19: number of N≡N bonds
2: molecular electronegativity	20: number of N=N bonds
3: number of CO ₂ groups	21: number of C≡N bonds
4: number of NO ₂ -Csp ² bonds	22: number of C atoms
5: number of NO ₂ -Csp ³ bonds	23: number of H atoms
6: number of NO ₂ -N bonds	24: number of N atoms
7: number of NO ₂ -O bonds	25: number of O atoms
8: number of rings	26: 100/molecular weight
9: number of -NH ₂ groups	27: indicator of aromaticity (0 or 1)
10: number of -OH groups	28: sum of the nitro's charges
11: number of -C(NO ₂) ₃ groups	29: average of the nitro's charges disymmetry
12: -CH in α of a nitroaromatic	30: nitro's charges disymmetry per molecular weight
13: indicator of symmetry	31: length of the longest X-NO ₂ bond
14: number of -C=O groups	32: length of the shortest X-NO ₂ bond
15: number of Y-O-X groups	33: highest potential for a X-NO ₂ bond
16: number of -C=C bonds	34: smallest potential for a X-NO ₂ bond
17: number of -C≡C bonds	35: average potential for a X-NO ₂ bond
18: number of -C=N bonds	36: average length of the X-NO ₂ bonds
37: formation energy	38: dipole
39: ionization potential	

Parameters 1 to 27 are obtained from the topology of the molecule whereas parameters 28 to 39 are calculated with MOPAC package.

III- DATA PREPROCESSING:

The determination of the pertinent parameters was carried out by a classical multivariate linear regression. After the elimination of the non-significant parameters, the predictions were tested by cross validation. In that procedure, the linear regression is performed on all the molecules except one, then the prediction of the sensitivity is made for the discarded molecule. This operation is repeated for all the data base.

In order to obtain a predicting method accessible to users who do not possess any quantum chemical software, two different processings were performed. The first one concerns only the parameters 1 to 27: i.e. no quantum mechanics calculations are involved; it is a purely topological coding. The second one takes into account all of the 39 parameters; it is a quanta-topological coding.

A- Topological coding

After the multilinear regression analysis, 16 parameters were discarded as having partial $F < 2$.

The following table presents the remaining parameters :

Table 1: Multilinear regression results for the topological coding

Correlation factor = 0,916 n = 204			
Standard-deviation s=0.18			
Equation		Stand. dev.	Partial-F
y =	1,065		
	- 0,140 x1	0,015	87,321
	- 0,146 x4	0,026	31,611
	- 0,205 x5	0,028	54,974
	-0,351 x6	0,033	113,078
	- 0,241 x7	0,059	16,649
	- 0,146 x12	0,054	7,288
	- 0,441 x20	0,052	72,001
	+ 0,039 x23	0,009	20,188
	+ 0,049 x24	0,001	25,235
	+ 0,064 x25	0,011	36,230
	+ 0,082 x26	0,019	18,662

Figure 1 presents the quality of the prediction tested by cross validation. A perfect value corresponds to a point located on the diagonal.

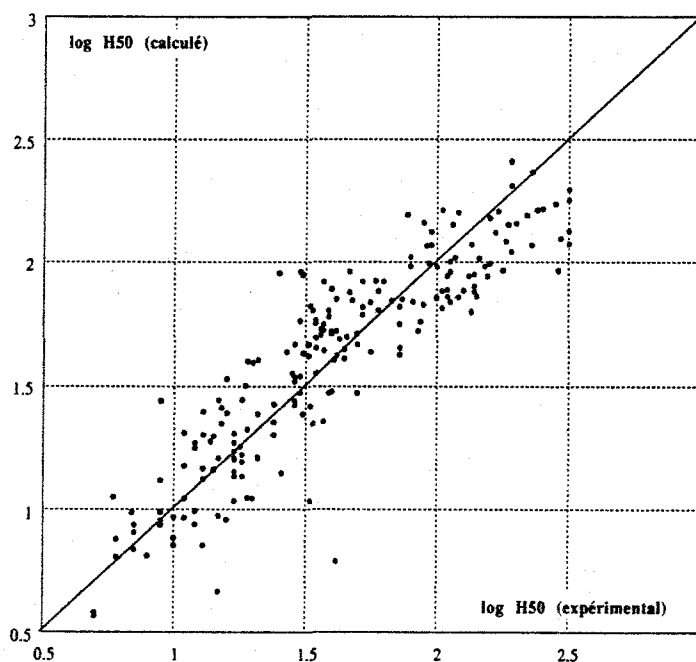


Figure 1: MLR Cross validation for the topological coding. (R=0.89, s=0.18, n=204)

An average correlation factor of 0.89 is obtained by cross validation.

B-Quanto-topological coding

The same procedure has been carried out with all the molecular parameters including those calculated from MOPAC.

In this case only 13 parameters remain.

Table 2: Multilinear regression results for the quonto-topological coding

Correlation factor = 0,918 n=204		
Standard deviation = 0,18		
Equation	Stand. dev.	Partial-F
y =	2,411	
	- 0,133 x1	0,015
	- 0,129 x4	0,040
	- 0,181 x5	0,041
	- 0,267 x6	0,054
	- 0,200 x7	0,071
	- 0,230 x12	0,052
	- 0,390 x20	0,052
	+ 0,035 x23	0,010
	+ 0,075 x25	0,017
	+ 0,075 x26	0,019
	- 0,219 x33	0,082
	+ 0,001 x37	0,001
	- 0,086 x39	0,020

The cross validation procedure gives a correlation coefficient of 0.89 which is equivalent to what was obtained by the purely topological coding.

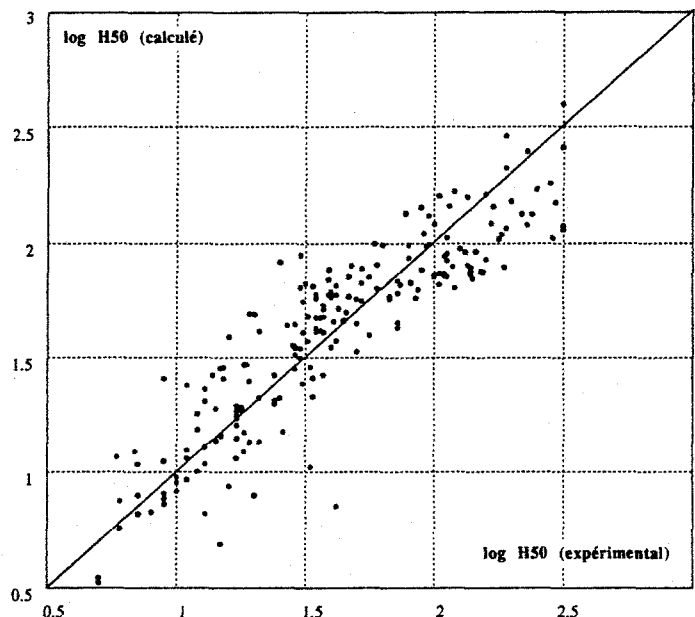


Figure 2: MLR Cross validation for the quonto-topological coding (R=0.89, s=0.18, n=204)

IV-NEURAL PROCESSING

A branch of artificial intelligence is now growing fast: the connexionism. It concerns the study of the behavior of assemblies of formal neurons which are inspired by a very crude model of the brain. Recent progresses in this field gave rise to promising results in pattern recognition, diagnosis, data processing....

A-The formal neuron:

A formal neuron is an automaton which is characterized by a state of activity rated between 0 and 1. This state is defined by the excitations received from other neurons through the synaptic connections which are the bonds between the neurons of a network.

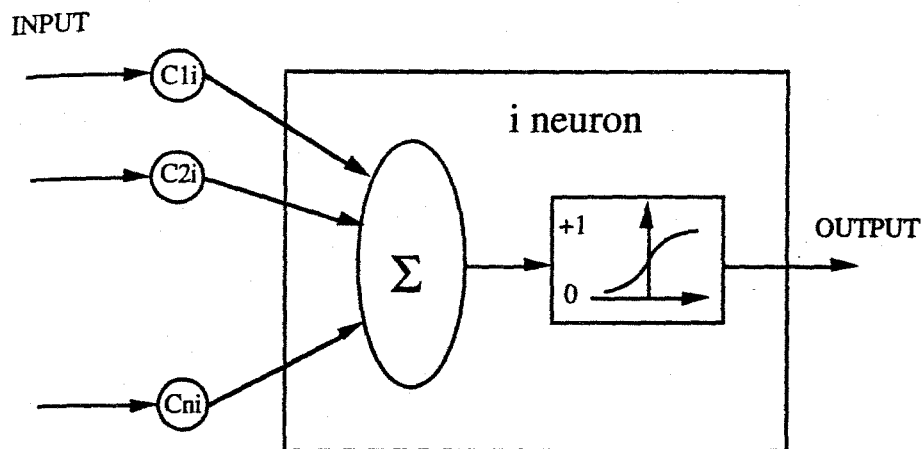


Figure 3: Scheme of a formal neuron

The excitations reaching a neuron are weighted by coefficients depending on the synaptic connections (the synaptic weights) then they are added.

The activity of the neuron is then fixed to a value between 0 and 1 by using a sigmoidal function of the sum of the weighted excitations and is transmitted to the other neurons through synaptic connections.

B-Layered Networks:

Various types of neural networks may be considered depending on the interaction schemes between the neurons: layered networks, fully or partially connected networks, recurrent networks....(Ref. 8)

Each of these types have advantages and disadvantages and its own domain of application.

Our purpose (prediction of a property) led us to choose the layered networks since their structure is very well fitted to this problem.

These networks are made of neurons arranged in layers. The neurons of a layer are not connected together. They receive excitation only from the neurons of the former layer and they transmit their activity only to the neurons of the next layer (Figure 4).

Usually a bias neuron whose activity state is always 1 is added in the input and hidden layers.

Two layers play a special role: the layer which do not receive any excitation which is called the input layer and the layer which do not transmit any excitation and which is called the output layer.

It is important to notice that the state of activity of the neurons of the output layer only depends on the state of activity of the neurons located in the input layer and of the synaptic weights of the network.

Therefore such a network may be considered as a "black box" which answers with the excitations of its output neurons to the excitations of its input neurons. The choice of the synaptic weights governs these answers.

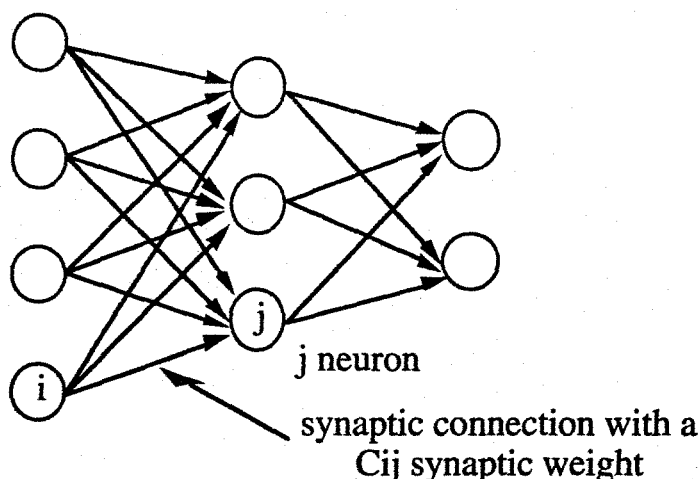


Figure 4: Scheme of a layered neural network

C-Training and Generalization:

The prediction of a property of an object is the association of a data set describing the property to another data set describing the object. The data set describing the object are called the descriptors.

For example the value of the impact sensitivity may be associated with the oxygen balance, the presence of some groups, etc....

The prediction of course has to be supported by experimental data: the prototypes for which the association description-prediction is already known. Therefore, in our case, we had to possess a set of molecules for which the sensitivity was known.

Before any prediction, the synaptic weights have to be modified in such a way that, for each prototype, the difference between the prediction and the known value is as small as possible. This step is the "training" of the network. It is performed by an iterating procedure called "Gradient Backpropagation" (Ref. 8). After this stage, the synaptic weights contain the data extracted from the prototypes.

It is then possible to make predictions. In that case, the descriptors of a molecule which does not belong to the training set are the inputs of the network. The information is propagated through the network towards the output layer where the prediction is collected.

D- Precautions for using networks:

The generalization aptitude, in other words, the level of prediction accuracy that can be obtained, is of course the essential quality of a network. This aptitude depends on many factors: the quality of the experimental data (number, repartition), the adequacy of the network structure to the complexity of the problem, the pertinence of the descriptors...

A rigorous optimization of these factors has not yet been proposed.

Nevertheless, a certain number of constraints on the structure of the network has to be taken into account.

The number of neurons in the input layer is determined by the number of parameters used to describe the prototype (here a molecule) and the number of neurons in the output layer depends on the descriptors associated with the prediction.

Furthermore, the number of neurons in the hidden layer must be chosen in such a way that the number of synaptic weights to be determined during the learning phase does not exceed the number of the prototypes (no more unknowns than equations).

Generally speaking, for a given problem, the optimal network architecture will be that which minimizes the number of connections, while allowing the use of descriptors that are rich enough to enable the discrimination of the prototypes.

Indeed, we can verify that, for a given problem, the generalization aptitude of this type of network first increases with the number of its connections until the network is flexible enough to fully take into account the complexity of the problem and enables proper learning conditions.

When this aptitude is reached, any increase in the number of connections will induce an excess in the network flexibility and thus reduce the generalization performance.

These remarks are comparable to those made for a polynomial interpolation.

Generally speaking, a statistical preprocessing of the prototypes submitted to a network is very useful. By lowering the number of descriptors, it allows the minimizing of the number of neurons in the input layer, and therefore the number of connections in the network.

E- Results:

After the optimization of the structure of the network with 204 molecules, the following results were obtained:

a-Topological coding:

The descriptors were those previously chosen by a statistical treatment. The optimal network was a 11-4-1 (11 neurons in the input layer, 4 neurons in the hidden layer and 1 neuron in the output layer).

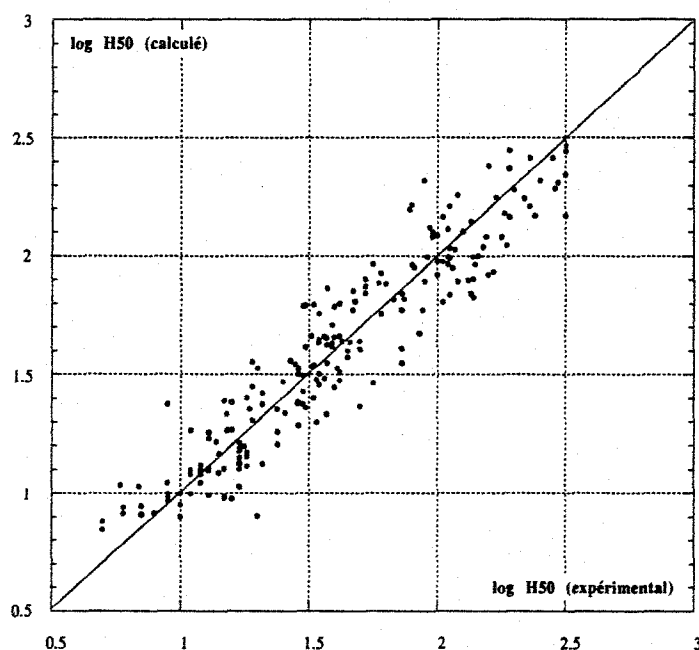


Figure 5: NN Cross validation for the topological coding ($R=0.94$, $s=0.13$, $n=204$).

A cross validation gave a correlation coefficient of 0.94 which is the best value we have obtained.

The figure 5 presents these results. One can notice that the correlation coefficient is clearly better (0.94) than in the multivariate analysis (0.89) and that

there are no more badly predicted compounds which is very promising for a real use of this method.

b- Quanto-topological coding:

A similar treatment had been performed on these data.

The best network has the 13-3-1 structure. A correlation factor of 0.92 has been obtained which is slightly inferior to the purely topological coding.

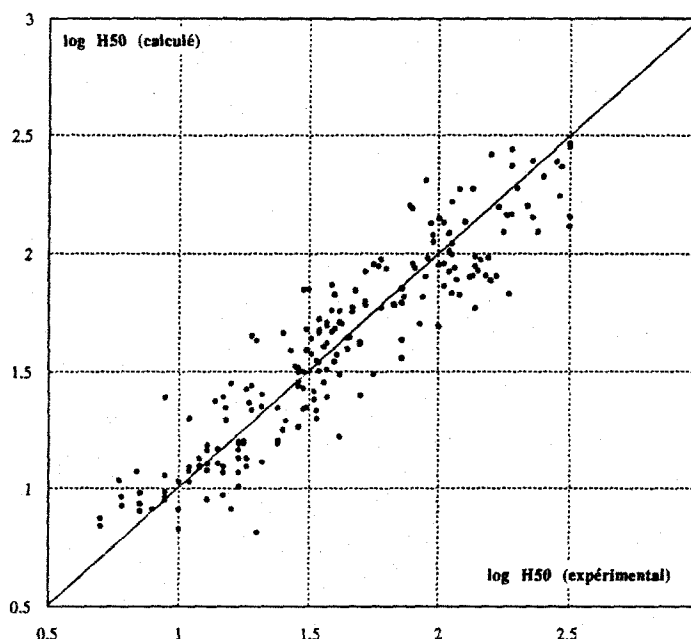


Figure 6: NN Cross validation for the quanto-topological coding (R=0.92, s=0.17, n=204).

Therefore the results obtained by a neural treatment are clearly better than those obtained by a classical multilinear treatment. They reveal that this new data processing method is very promising by allowing non linear correlations.

This method has also been tested on a separate prediction set composed of 68 molecules that were either non explosive (musk) or whose sensitivity was measured with other protocols. The results are quite satisfactory and the quanto-topological treatment seems to give better results in these cases which promote the use of such a coding even if it needs the use of a semi-empirical quantum chemistry program.

F- Improvement of the results:

The main purpose of this work was the building of a prediction tool dedicated to a real use. Therefore it was of primary importance to estimate the robustness of the prediction.

Therefore, rather than calculating a value of the sensitivity, we chose to classify a molecule in 3 classes:

class 1: High sensitivity ($5\text{cm} < H < 87\text{cm}$)

class 2: Medium sensitivity ($88\text{cm} < H < 160\text{cm}$)

class 3: Low sensitivity ($160\text{cm} < H < 320\text{cm}$)

The output layer of the network contains 3 neurons; one for each class. During the training step, the activities of these neurons are set to values that are the probabilities of belonging to a class. They depend on the value of the sensitivity compared to the extreme values of a class. For example, a sensitivity $H=160\text{cm}$ can

be classified in class 2 and class 3 with the same probability. On the other hand, a sensitivity $H=123\text{cm}$ which is the center of class 2 has a maximum probability of being classified in class 2.

During the prediction, the analysis of the activities of the 3 neurons of the output layer give informations on the quality of the prediction.

For example, in the case of the prediction of the sensitivity of a class 3 compound, the perfect prediction should be $C1=C2=0$ and $C3=1$ ($C1$: activity of neuron 1, $C2$: activity of neuron 2, $C3$: activity of neuron 3 with $C1+C2+C3=1$). In reality these values are never reached.

Several configurations may be obtained:

1- If $C1=C2=e$ and $C3=1-2e$, with $e<0.15$ then the network considers that the sensitivity belongs to class 3 and that the prediction is of good quality.

2- If $C1=0$ and $C2=C3$ the prediction is considered to be relatively ambiguous.

3- If $C1=C2=C3$ the prediction is considered to be completely ambiguous.

4- $C1>1-2e$ and $C2=C3=e$ or $C2>1-2e$ and $C1=C3=e$

The first case is the most favorable: the network provides a good classification without ambiguity.

Cases 2 and 3 are also acceptable since the network tells the user to be careful with the prediction.

Case 4 is unacceptable: the network provides a wrong prediction without ambiguity.

This method has been tested for both coding schemes. The structures of the networks were 11-2-3 and 13-3-3.

The results were satisfactory: 95% of the molecules were correctly classified and the remaining 5% were ambiguous answers. Case 4 was never observed which is encouraging for a real use of this method.

V-CONCLUSION:

The methodology we describe here seems to be promising: it concerns a large variety of molecules and the most significant parameters are simultaneously taken into account. Moreover, non linear correlations are directly considered which is certainly favorable for a complicated phenomenon such as the impact sensitivity.

However, the quality of the prediction should certainly be much higher if the experimental data base were improved by using a more accurate way for measuring the impact sensitivity, especially for the low sensitivity compounds.

Nevertheless, this new methodology could easily be used with this improved data base which make us think that neural networks will certainly have a brilliant future in the field of explosives and pyrotechnics.

REFERENCES

1. Kamlet, M.J. and Adolph, H.G. (1979). *Propellants and Explos.* **4**, 30.
2. Mullay, J. (1987). *Propellants, Explosives, Pyrotechnics* **12**, 60.
3. Polizer, P., Murray, J.S. and Lane, P. (1991). *Chem. Phys. Letters* **181**, 78.
4. Cheret, R., Delpuech, A., and Michaud, C. (1989). *Detonation in condensed explosives*. (T1) Masson, Paris, 140.
5. Pivina, T.S., Petrov, E.A., Agranov, G.K. and Shlyapochnikov, V.A. (1991). *Com. Exp. and Shock Waves* **27**, 242.
6. Encyclopedia of explosives and related items (Picattiny Arsenal) - Vol 7, 135
7. Stewart J.S. (1990). *J. Computer -Aided Molecular Design* **4**, 1.
8. Rumelhart D.E., McClelland J.L. (1988). *Parallel distributed processing*. The MIT Press Cambridge Massachussets.