

COMPUTATIONAL SENSOR FOR ANALYSIS OF ENVIRONMENTAL POLLUTANTS

R. K. Tiwari, Amit Kumar Sharma, M. S. Gaur
Hindustan College of Science and Technology, Farah, Mathura.
twr_uv@yahoo.co.in

Abstract: We will present a computational sensor for the detection of environmental pollutants using a neural network method that takes into account the sensor or statistical parts interaction to decrease the signal dynamic range and to increase the model coverage (MLR & PLR model).

1. INTRODUCTION

Over the past two decades, the optical measurement has gained increasing popularity in all aspects of manufacturing and quality inspection processes. Despite its wide usage, the optical measurement is very sensitive to the relative position and orientation between the sensor components. Sensors are often designed without a detailed analysis of geometrical and optical properties of the parts. This may lead to high signal dynamic range and consequently poor model coverage or low measurement accuracy. But we can use to compute and analyze absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties [1]-[3].

2. METHODS

Three computational models (MLR, PLS and ANN) were selected to analysis of molecular properties and ADMET tests.

2.1 Artificial Neural Network (ANN) method using back propagation method

To create back propagation Neural Network model protocol produces a neural network model for using the selected molecular descriptors. It appears in the other group in the molecular properties dialog to calculate the molecular properties protocol and used to compute the property for future ligands.

2.2 Multiple Linear Regression (MLR) method

The Multiple Linear Regression model protocol builds a model for a dependent property using the selected molecular descriptors. It is used to compute the property for any ligands.

Multiple Linear Regression requires at least as many molecules as independent variables. Reliable results typically require 5 times as many molecules as independent variables. We use Partial Least Squares when there are a large number of independent variables.

2.3 Partial Least Square (PLS) method

The Partial Least Squares model protocol will build a Partial Least Squares model for a dependent property using the selected molecular descriptors. It is once time created, that appears in the other group in the molecular properties dialog to calculate molecular properties

protocol and is used to compute the property for any other ligands

3. RESULTS AND DISCUSSION

3.1 Analysis of chemical structure of Environmental Pollutants

The Chemical structure of all the three pesticides downloaded from the Pubchem Compound Search of NCBI database [4]. All the pesticides structures were downloaded from PubChem database as *.sdf file. The three dimensional (3-D) structure of pesticides has its own Pubchem Id number, i.e. Malathion PubChem Id is CID: 4004, Ethyl parathion: CID: 4130 and Parathion: CID: 991. The 3-D structure of pesticides is shown in the RasMol V2.7.2.1.1 [5] visualization tool.

Table 1: Toxicity prediction (ADMET) properties of methyl parathion, malathion and parathion pesticides through TOPKAT protocol of D.S.-2.1.

Tests	Malathion	Parathion	Methyl parathion
ADMET_BBB	-0.774	-0.396	-0.18
ADMET_BBB_Level	3	2	2
ADMET_Absorption_Level	0	0	0
ADMET_Solubility	-2.093	-3.585	-4.091
ADMET_Solubility_Level	3	3	2
ADMET_Hepatotoxicity	0	1	1
ADMET_Hepatotoxicity_Probability	0.139	0.629	0.523
ADMET_CYP2D6	0	0	0
ADMET_CYP2D6_Probability	0.247	0.029	0.108
ADMET_PPB_Level	0	0	0
ADMET_AlogP98	1.595	2.524	3.221
ADMET_Unknown_AlogP98	1	1	1
ADMET_PSA_2D	70.321	64.595	64.595
ADMET_PPB_Level	0	0	0

It is a molecular graphics visualization tool for proteins and any macromolecules and is freely available on Internet.

3.2 Toxicity predictions (ADMET)

The final 3-D structure was further used to analyze through Toxicity Prediction (TOPKAT) protocol of D. S. 2.1 (Accelrys). TOPKAT computes and validates assessments of the toxic and environmental effects of chemicals solely from their molecular structure.

The ADMET_BBB result that predicts blood-brain penetration (Blood Brain Barrier, BBB) after oral administration of any lead molecules in a human body. In case of Malathion, Parathion and Methyl parathion all shows the low penetration in blood-brain barrier level in human body. The ADMET_Absorption_Level in the lead compounds have good absorption level in human intestinal absorption (HIA) after oral administration.

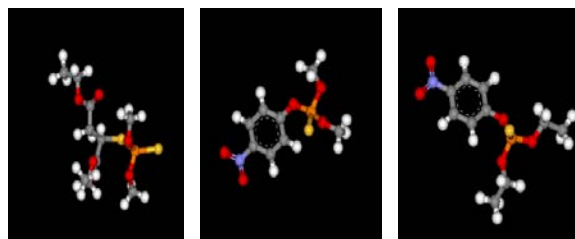
The ADMET_Solubility are the base 10 logarithm of the molar solubility as predicted the solubility of each compound in water at 25 °C by the regression analysis. In case of malathion the solubility level -2.093 that shows the good solubility level but in parathion and methyl parathion the -3.585 and -4.091 respectively the defines the low solubility level in aqueous medium, as shown in Table 1.

The ADMET_Hepatotoxicity model of TOPKAT protocol predicts potential organ toxicity for a wide range of structurally diverse compounds. The result shows the malathion is not toxic but other two pesticides are toxic. ADMET_Hepatotoxicity_Probability is the hepatotoxicity score is the sum of the predicted values (0 and 1) from all individual trees that comprise the ensemble recursive partitioning model, divided by the total number of trees in that model. The parathion and methyl parathion are closer to toxic and malathion is nontoxic because their value is near to zero value.

3.3 Analysis of Molecular properties and MLR, PLS, ANN model analysis

The molecular properties prediction of all three pesticides is predicted by using Calculate Molecular Properties protocol of D. S. 2.1. This protocol can calculate many properties that can be used to create a QSAR model. These range from the traditional molecular descriptors, Semi empirical QM descriptors, Density Functional QM descriptors, and properties that can be calculated from user-built QSAR models.

In addition, we can apply specific rules to remove ligands that are not likely drug-like, unsuitable leads, etc based on the presence or absence and frequency of certain chemical groups. This is especially true for minor change in molecular structure of some pollutants like methyl parathion, malathion and parathion, where the structure passes different phenomena to the other compound. The 3-D structure of methyl parathion, malathion and parathion are shown in Figure1.



Malathion Parathion Methyl parathion

Figure1: The three dimensional structures of all the three pesticides like Malathion Parathion and Methyl parathion by using the RasMol visualization tool.

The 3-D structure of all the three pesticides contains Hydrogen (white), Oxygen (Red), Carbon (Grey),

Sulphur (Yellow), Nitrogen (Purpule) and Phosphorous (Orange). The malathion structural formula is $C_{10}H_{19}O_6PS_2$, parathion $C_{10}H_{14}NO_5PS$ and methyl parathion $C_8H_{10}NO_5PS$. This is especially true for minor change in molecular structure of some pollutants like methyl parathion, malathion and parathion, where the structure posses different phenomena to the other compound. In such a model, the data from the structure is very closely with the compounds and the accuracy can be increased using computational model. The data is controlled and the acquisition the output signal is achieved using a neural network method and the correlation coefficient is achieved in the molecular structure of pollutants by using a computational toxicity model [Figure 2, 3 & 4; & Table1, 2].

Name	A LogP	Molecular Weight	Num_Rotatable Bonds	Num_H Acceptors	Num_H Donors	Molecular Volume	Molecular Fractional Polar Surface Area
Melathion	1.423	331.366	11	8	1	223.97	0.416
Parathion	0.394	267.239	5	7	2	162.92	0.467
Methylparathion	1.092	295.292	7	7	2	186.93	0.415

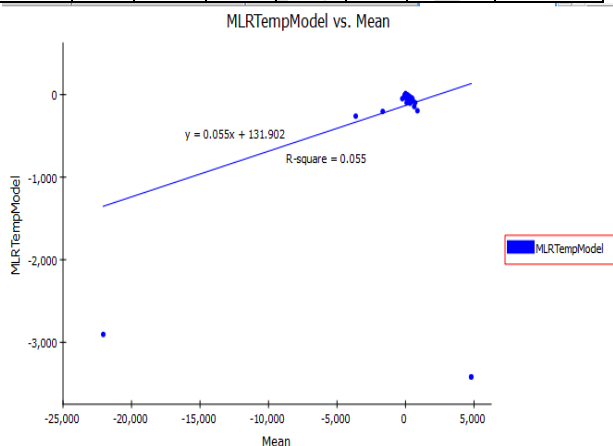


Figure 2: MLR model

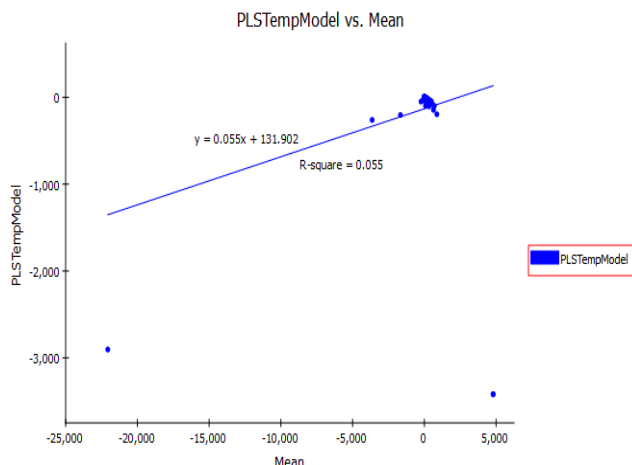


Figure 3: PLS model

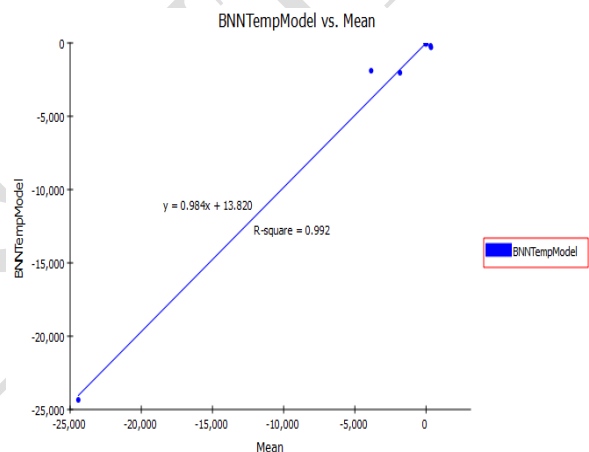


Figure 3: ANN model

Table 2: Molecular properties of methyl parathion, malathion and parathion.

4. CONCLUSION

The overview presented here to discuss the advance in and the applicability of predictive computational sensing techniques for the detection of environmental pollutants. The evaluation of pollutants target affinity and specificity, identification and to predict their fate in the body through ADME and toxicity studies. In the study, three computational models were studied, in which, ANN model is found to be suitable. The correlation coefficients are calculated 0.055 for MLR, 0.055 for PLS and 0.992 for ANN models. It means that the MLR and PLS models have very less significant than ANN model. The produced data can be utilized for the detection of environmental pollutants.

5. REFERENCES

- [1] S.P. DeWeerth, "Analog VLSI Circuits for Stimulus Localization and Centroid Computation,"*Intl. Jour. of Comp. Vision*, **8(3)**, 202 (1999).
- [2] P. Ceccato, S. Flasse, S. Tarantola, S. Jacquemoud and J.M. Gr_egoire, "Detecting vegetation Water content using reactance in the optical domain," *Remote Sensing of Environment*, **77**, 22 (2001).
- [3] D. Combes, L. Bousquet, S. Jacquemoud, H. Sinoquet, C. Varlet-Grancher and I. Moya, "A new spectrophotogoniometer to measure leaf spectral and directional optical properties,"*Remote Sensing of Environment* **1**, 107 (2007).
- [4] <http://pubchem.ncbi.nlm.nih.gov/search/search.cgi>
- [5] <http://www.umass.edu/microbio/rasmol>

NOT PRESENTED