Journal of the Serbian Chemical Society
JSCS-info@shd.org.rs • www.shd.org.rs/JSCS

ACCEPTED MANUSCRIPT

($RMSE_{ANN}$ = 0.5968, $RMSE_{SVM}$ = 0.5933). All statistical parameters of the model are satisfying and prove that the models are stable, robust and predictive. Then, the built models were used to predict the test set data.

Analysis of descriptors contribution in the ANN model and interpretation

To evaluate the influence of each descriptor on the calculated solubility, we based on relevance score. The Relevance Score is calculated by following all paths from the input neuron to the output neuron (including hidden layers). For each path, the product of all the connection weights (in absolute value) is added to the score. Afterwards, all relevance scores are normalized to be in the range between 0 and 100.[5]

The relative value of contributions of the five descriptors of the model has been determined (TABLE III). These values of contributions allow the following classification: RNCG > AlogP2 > MAXDN > MaxS8m > Mor26u. It should be noted that the difference in the descriptor contribution between any two descriptors used in the model is not significant, indicating that all of the descriptors are indispensable in generating the predictive model.

TABLE III. Relevance score

Index	Name	Relevance Score
0	RNCG	100
1	ALOGP2	76
2	MAXDN	39
3	MaxS8m	32
4	Mor26u	28

These values confirm the great effect of the RNCG and AlogP2 on the solubility. Another advantage of this method is determination of the impact of each descriptor on the analysis solubility.

The first important descriptor is the relative negative charge (RNCG) is the charge of the most negative atom divided by the total negative charge (Q_{neg}). The charge density on the ions plays an important role in the interactions of these ions and the water molecule, which reflects the influence of the negative charge on the aqueous solubility. The second important descriptor is the Squared-Ghose-Crippen-Viswanadhan octanol-water partition coefficient (AlogP2), is calculated from a regression equation based on the hydrophobic character of the molecule. It reflects both the interactions of the solute with the bulk of the surrounding solvent (macroscopic or non-specific solvent effects) and the specific bonding between the solute and individual solvent molecules (microscopic or specific solvent effects).[29]

MAXDN, is a topological descriptor, the maximal electrotopological negative variation (MAXDN) is calculated as the maximum negative value of DI_i (topologic distance) in the molecule.

Mats8m, is a 2D autocorrelation indices, is calculated by applying the Moran coefficient to the molecular graph.[7]

Mor26u, is a descriptor of the 3D-MoRSE descriptor class, the 3D-MoRSE descriptors are the 3D molecular representations of structure based on electron diffraction descripttor[30,31], which are calculated by summing atomic weights viewed by a different angular scattering function. The values of these descriptor functions are calculated at 32 evenly distributed values of scattering angle (s) in the range of 0 - 31 Å from the three-dimensional atomic coordinates of a molecule. The 3D-MoRSE descriptor is calculated using following expression:

$$\text{Morsw} = \sum_{i=1}^{nAT-1} \sum_{j=i+1}^{nAT} w_i w_j (\sin(sr_{ij}) / sr_{ij}) \quad (16)$$

where s is the scattering angle, nAT is the number of atoms, r_{ij} is the interatomic distance between the i^{th} and the j^{th} atoms, w is an atomic property, including atomic number, masses, van der Waals volumes, Sanderson electronegativities, and polarizabilities.

TABLE IV. Golbraikh and Tropsha criteria

	Validation set n = 22					
Method	R^2_{test} / %	Q^2_{ext} / %	$(R^2-R^2_0) / R^2<0.1$	$(R^2-R'^2_0) / R^2<0.1$	$0.85<k<1.15$	$0.85<k'<1.15$
ANN	74.14	73.94	-0.3379	-0.3461	0.9628	0.9794
SVM	70.70	70.61	-0.4043	-0.356	1.0333	0.9050

The statistical parameters obtained for the test set,[32] demonstrate the power of the predictivity of the models.

Comparison of the results with other modeling methods

We have compared our results with other modeling methods in previous publications. Table V summarize those comparisons, it shows that our SVM model gives better prediction than most of the other methods, because we used less descriptors compared the other models, and we obtained an almost similar result. In addition, our models were evaluated by using different statistical parameters compared with other model in the literature.[33]

TABLE V. Comparison of our results and other modeling methods

References	Method	Test set	Training set	Number of descriptors	R^2_{tr}	R^2_{test}	RMSE
Our results	ANN	22	58	5	80.97	74.12	0.5968
	SVM	22	58	5	84.03	70.68	0.5933
Deeb and Goodarzi [3]	PLS		219	22	79.98	79.44	
	PC- ANN		219	22	84.35	81.93	
Bouakkadia *et al.*[29]	MLR	19	58	6	88.95	85.11	0.5200

The difference between this work and the previously published work on this data set, that the number of compounds in the validation set is not the same, as

well as the two training and validation sets do not contain the same compounds, because the method of separation is not the same. And the descriptors selected by genetic algorithm are different except AlogP2.

CONCLUSION

A quantitative – structural property relationship analysis has been performed on the logarithm of solubility in water for 80 pesticides compounds by using ANN and SVM. The built models clearly demonstrate good correlations between the structure and aqueous solubility of the studied compounds. Five descriptors were selected with genetic algorithm. The selected descriptors, which are MATS8m, RNCG, AlogP2, MAXDN and Mor26u, were found to be important factors controlling the aqueous solubility. Comparison between ANN and SVM methods demonstrates that the performance of SVM model is better than that of ANN, but the ANN model is more general than SVM because of the great value of R^2_{test}.The proposed models will help identifying new pesticides and provide insight to guide their development and may be useful for predicting their solubility.

SUPPLEMENTARY MATERIAL

Supplementary Material is available from Journal Web site http://www.shd.org.rs/JSCS/, or from the corresponding author on request.

ИЗВОД

УПОТРЕБА GA-ANN И GA-SVM ЗА QSPR СТУДИЈУ РАСТВОРЉИВОСТИ ПЕСТИЦИДА У ВОДИ

AMEL BOUAKKADIA[1,2], NOUREDDINE KERTIOU[1,2], RANA AMIRI[1], YOUSSOUF DRIOUCHE[1] и DJELLOUL MESSADI[1]

[1]Environmental and Food Safety Laboratory, Department of Chemistry, Badji Mokhtar University - Annaba, BP. 12, 23000 Annaba, Algeria; [2]Abbes Laghrour University, Faculty of Sciences and Technology - Khenchela, BP 1252 Route de Batna 40004 Khenchela, Algeria

Тежња пестицида, у овој студији нарочито херницида, за расподелу по различитим одељцима животне средине, зависи углавном од физичкохемијских особина самих пестицида. Растворљивост у води (S) указује на тенденцију пестицида да се уклоне испирањем или иригацијом да би завршили у површинским водама. Експериментални поступак за одређивање растворљивости пестицида у води је веома скуп и тежак. QSPR методе се често користе за процену растворљивости хербицида у води. Методе вештачке неуронске мреже (ANN) и векторске машине за подршку (SVM), сваки пут повезане са селекцијом помоћу генетичког алгоритма (GA) за избор најзначајније варијабле, биле су коришћене за развој QSPR модела за предвиђање растворљивости у води серије од 80 хербицида. Вредности $\log S$ проучаваних једињења добро су корелисане са дескрипторима. Разматрајући погодне дескрипторе, квадратни Пирсонов коефицијен (R^2) 0,8 добијен је за ANN модел за структуру 5-3-1 и 0,8 је добијен за SVM модел користећи RBF функцију за оптималне вредности параметара: C = 11,12; σ= 0,1111 and ε = 0,222.

(Примљено 18. јуна; ревидирано 19. августа; прихваћено 6. октобра 2020)

REFERENCES

1. P. Gramatica, A. Di Guardo, *Chemosphere* **47** (2002) 947 (https://doi.org/10.1016/S0045-6535(02)00007-3)
2. O. C. Hansen, *Quantitative Structure-Activity Relationships (QSAR) and Pesticides*, Teknologisk Institute, Pesticides Research No. **94** (2004) (https://www2.mst.dk/udgiv/publications/2004/87-7614-434-8/html/helepubl_eng.htm)
3. M. Zine, A. Bouakkadia, L. Lourici, D. Messadi, *J. Serb. Chem. Soc*. **84** (2019) 1405 (https://doi.org/10.2298/JSC190306059Z)
4. *Hyperchem*TM, Release 6.03 for Windows, Molecular Modeling system, 2000 (http://www.hyper.com)
5. *Molegro Data Modeller (*MDM), v.2.0. Copyright Molegro (2009) (https://www.scientific-computing.com/press-releases/molegro-data-modeller-v20)
6. I. N. Levine, *Quantum Chemistry*. 5thed. New Jersey: Prentice Hall, 2000, p. 626- 739 (ISBN-10: 0136855121)
7. R. Todeschini, V. Consonni, M. Pavan, Dragon, *Software for the Calculation of Molecular Descriptors*, Release 5.3 for windows, Milano, 2006 (http://www.talete.mi.it)
8. R. Todeschini, D. Ballabio, V. Consonni, A. Mauri, M. Pavan, *MOBYDIGS, Software for Multilinear Regression Analysis and Variable Subset Selection by Genetic Algorithm*, Release 1.1 for Windows, Milano, Italy, 2009 (http://www.talete.mi.it/)
9. R. Amiri, D. Messadi, A. Bouakkadia, *J. Serb. Chem. Soc*. **85** (2020) 467 (https://doi.org/10.2298/JSC190610090A)
10. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, London: Addison Wesley, 1989, p. 89- 94 (https://dl.acm.org/citation.cfm?id=534133)
11. A. Beheshti, E. Pourbasheer, M. Nekoei,S. Vahdani, *J. Saudi Chem. Soc.* **28** (2016) 282 (http://dx.doi.org/10.1016/j.jscs.2012.07.019)
12. J. Xu, H. Zhang, L. Wang, G. Liang, L. Wang, X. Shen, W. Xu, *Spectrochim. Acta Part A* **76** (2010) 239 (https://doi.org/10.1016/j.saa.2010.03.027)
13. S. Jothilakshmi, V. N. Gudivada, *Handbook of Statistics* **35** (2016) 301-340 (http://dx.doi.org/10.1016/bs.host.2016.07.005)
14. M. Safamirzaei, H. Modarress, M. Mohsen-Nia, *Fluid. Phase. Equilibria*. **266** (2008) 187 (https://doi.org/10.1016/j.fluid.2008.01.022)
15. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, New Delhi, 2006 (ISBN-13: 978-0023527616)
16. R. Rojas, *Neural Network*, Springer-Verlag, Berlin, 1996, p 229-261 (http://page.mi.fu-berlin.de/rojas/neural/)
17. O. Deeb, M. Drabh, *Chem. Biol. Drug. Des*. **76** (2010) 255 (https://dx.doi.org/10.4236/aces.2012.21010)
18. V. N. Vapnik, *Statistical Learning Theory*. John Wiley & Sons: New York, 1998, p. 375-473 (ISBN: 978-0-471-03003-4)
19. V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer2000, p. 267-287 (ISBN: 978-1-4757-3264-1)
20. C. J. Lu, T. S. Lee, C. C. Chiu, *Decision. Support. Syst*. **47** (2009) 115 (https://doi.org/10.1016/j.dss.2009.02.001)
21. N. Cristianini, J. Shawe-Taylor, *An introduction to support vector machines and other kernel- based learning methods*, Beijing: Publishing House of Electronics Industry, 2005, p. 93-122 (ISBN: 0521 780195).
22. R. Amiri**,** D. Messadi, A. Bouakkadia**,** L. Lourici, *Egypt. J. Chem.* **62** (2019) 1563 (https://doi.org/ 10.21608/ejchem.2019.4976.1446)

23. V. Vapnik, *The nature of Statistical Learning Theory*, Springer, New York, 1995, p. 70-71 (ISBN: 0-387-94559-8).
24. N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University Press: Cambridge, UK, 2000, p. 32-42 (ISBN: 0-521-78019-5).
25. B. Scholkopf, A. Smola, *Learning with Kernels*, MIT Press: Cambridge, MA, 2002, p.13-17 (ISBN: 0262194759)
26. A. Golbraikh, A. Tropsha, *J. Mol. Graph. Model*. **20** (2002) 269 (https://doi.org/10.1016/S1093-3263(01)00123-1)
27. N. Kertiou, A. Bouakkadia**,** D. Messadi, *Res. J. Pharm. Biol. Chem. Sci.* **8(6)** (2017) 251 (https://www.rjpbcs.com/pdf/2017_8(6)/[29].pdf)
28. A. Bouakkadia**,**Y. Driouche, N. Kertiou, D. Messadi, *Int. J. Saf. Secur. Eng.* **10** (2020) 389 (https://doi.org/10.18280/IJSSE.100311)
29. A. Bouakkadia, H. Haddag, N. Bouarra, D. Messadi, *Synthèse: Revue des Sciences et de la Technologie* **32** (2016) 12 (in French) (https://www.ajol.info/index.php/srst/issue/view/13942)
30. J. Gasteiger, J. Sadowski, J Schuur, P. Selzer, L. Steinhauer, V. Steinhauer, *J. Chem. Inf. Comput. Sci*. **36** (1996) 1030 (https://doi.org/10.1021/CI960343+)
31. J. Schuur, P. Selzer, J. Gasteiger, *J. Chem. Inf. Comput. Sci.* **36** (1996) 334 (https://doi.org/10.1021/CI950164C)
32. W. J. Wang, Z. B. Xu, W. Z. Lu, X. Y. Zhang, *Neurocomputing* **55** (2003) 643 (https://doi.org/10.1016/S0925-2312(02)00632-X)
33. O. Deeb, M. Goodarzi, *Mol. Phys.* **108** (2010) 181 (https://doi.org/10.1080/00268971003604575)

SUPPLEMENTARY MATERIAL TO

Use of GA-ANN and GA-SVM for a QSPR study on aqueous solubility of pesticides

AMEL BOUAKKADIA[1,2], NOUREDDINE KERTIOU[1,2], RANA AMIRI[1], YOUSSOUF DRIOUCHE[1] and DJELLOUL MESSADI[1]

[1]*Environmental and Food Safety Laboratory, Department of Chemistry, Badji Mokhtar University - Annaba, BP. 12, 23000 Annaba, Algeria; 2Abbes Laghrour University, Faculty of Sciences and Technology - Khenchela, BP 1252 Route de Batna 40004 Khenchela, Algeria*

TABLE S-I. Pesticides compounds used in this study and their solubility

No	Name	log (*S* / mg l^{-1})
1	2-ethylamino-4-(isopropylamino)-6-(methylthio)-s triazine	2.27
2	methyl α-[(4,6-dimethoxypyrimidin-2-ylcarbamoyl) sulfamoyl]-*o*-toluate	2.08
3	S-Ethyldiisobutylcarbamothioate	1.64
4	ethyl 2-(4-chloro-6-methoxypyrimidin-2-ylcarbamoylsulfamoyl) benzoate	3.08
5	3-[4-(4-chlorophenoxy)phenyl]-1,1-dimethyl-urea	0.4
6	1-(2-chlorophenylsulfonyl)-3-(4-methoxy-6-methyl-1,3,5-triazin-2-yl)urea	4.45
7	3,6-dichloropyridine-2-carboxylic acid	5.16
8	(2-hydroxyethyl)ammonium 3,6-dichloropyridine-2-carboxylate	5.75
9	2-{[4-chloro-6-(ethylamino)-1,3,5-triazin-2-yl]amino}-2-methylpropanenitrile	2.23
10	S-ethyl N-cyclohexyl-N-ethylcarbamothioate	1.98
11	2,4-Dichlorophenoxyacetic acid	2.95
12	diethylammonium (2,4-dichlorophenoxy) acetate	5.9
13	methyl 2-(2,4-dichlorophenoxy)acetate	2
14	2,3,6-trichlorobenzoic acid	3.89
15	ethyl 3-phenylcarbamoyloxycarbanilate	0.95
16	(2R)-2-(2,4-dichlorophenoxy)propionic acid	2.77
17	3-[4-(4-methoxyphenoxy)phenyl]-1,1-dimethylurea	1.3
18	6-ethylthio-N^2,N^4-diisopropyl-1,3,5-triazine-2,4-diamine	1.2
19	3-(3,4-dichlorophenyl)-1,1-dimethylurea	1.62
20	S-ethyl dipropyl(thiocarbamate)	2.54
21	methyl 2-[(4-ethoxy-6-methylamino-1,3,5-triazin-2-yl)carbamoylsulfa-moyl]benzoate	1.7
22	ethyl (RS)-2-[4-(6-chloro-1,3-benzoxazol-2-yloxy)phenoxy] propionate	-0.1
23	butyl (RS)-2-{4-[5-(trifluoromethyl)-2-pyridyloxy]phenoxy}propionate	0
24	butyl (R)-2-{4-[5-(trifluoromethyl)-2-pyridyloxy]phenoxy}propionate	0.3
25	(*RS*)-1-methylheptyl 4-amino-3,5-dichloro-6-fluoro-2-pyridyloxyacetate	-1.05
26	N-(phosphonomethyl) glycine	4.08

27 (RS)-2-{4-[3-chloro-5-(trifluoromethyl)-2-pyridyloxy]phenoxy}propionic acid	1.64
28 3-(4-isopropylphenyl)-1,1-dimethylurea	1.81
29 3-cyclohexyl-1,5,6,7-tetrahydrocyclopentapyrimidine-2,4(3H)-dione	0.78
30 3-(3,4-dichlorophenyl)-1-methoxy-1-methylurea	1.88
31 4-(4-chloro-o-tolyloxy)butyric acid	1.64
32 (RS)-2-(4-chloro-o-tolyloxy)propionic acid	2.87
33 4-amino-4,5-dihydro-3-methyl-6-phenyl-1,2,4-triazin-5-on	3.23
34 2-(3,4-dichlorophenyl)-4-methyl-1,2,4-oxadiazolidine-3,5-dion	0.18
35 3-(3-chloro-4-methoxyphenyl)-1,1-dimethylurea	2.83
36 methyl 2-(4-methoxy-6-methyl-1,3,5-triazin-2-ylcarbamoylsulfamoyl)benzoate	3.98
37 (*RS*)-*N*,*N*-diethyl-2-(1-naphthyloxy)propionamide	1.87
38 S-propyl butyl(ethyl) thiocarbamate	2
39 4-amino-3,5,6-trichloropyridine-2-carboxylic acid	2.63
40 N^2,N^4-diisopropyl-6-methoxy-1,3,5-triazine-2,4-diamine	2.86
41 N^2,N^4-diisopropyl-6-methylthio-1,3,5-triazine-2,4-diamine	1.52
42 6-chloro-N^2,N^4-diisopropyl-1,3,5-triazine-2,4-diamine	0.93
43 S-benzyl dipropyl (thiocarbamate)	1.12
44 ethyl (2RS)-2-[4-(6-chloroquinoxalin-2-yloxy)phenoxy] propionate	-0.51
45 1-(4,6-dimethoxypyrimidin-2-yl)-3-(3-ethylsulfonyl-2-pyridylsulfonyl)urea	3.86
46 6-chloro-N^2,N^4-diethyl-1,3,5-triazine-2,4-diamine	0.79
47 1-(5-tert-butyl-1,3,4-thiadiazol-2-yl)-1,3-dimethylurea	3.4
48 3-tert-butyl-5-chloro-6-methyluracil	2.85
49 N^2-*tert*-butyl-N^4-ethyl-6-methylthio-1,3,5-triazine-2,4-diamine	1.34
50 methyl 3-(4-methoxy-6-methyl-1,3,5-triazin-2-ylcarbamoylsulfamoyl)thiophene-2-carboxylate	3.8
51 *S*-4-chlorobenzyl diethyl(thiocarbamate)	1.45
52 S-2,3,3-trichloroallyl diisopropyl(thiocarbamate)	0.6
53 methyl 2-[4-methoxy-6-methyl-1,3,5-triazin-2-yl(methyl)carbamoylsulfamoyl] benzoate	3.18
54 3,5,6-trichloro-2-pyridyloxyacetic acid	3.91
55 2-butoxyethyl 3,5,6-trichloro-2-pyridyloxyacetate	1.36
56 methyl 2-[4-dimethylamino-6-(2,2,2-trifluoroethoxy)-1,3,5-triazin-2-ylcarbamoylsulfamoyl] totoutoluate	2.04
57 1,1-dimethyl-3-(α,α,α-trifluoro-m-tolyl)urea	2.04
58 iso-octyl 4-chloro-o-tolyloxyacetate	0.7
59* methyl 2-[4,6-bis(difluoromethoxy)pyrimidin-2-ylcarbamoylsulfamoyl] benzoate	2.39
60* 1-(2-methylcyclohexyl)-3-phenylurea	1.26
61* 6-chloro-N-ethyl-N'-isopropyl-1,3,5-triazin-2,4-diamin	1.52
62* 5-bromo-3-(butan-2-yl)-6-methylpyrimidine-2,4(1H,3H)-dione	2.85
63* (R)-1-(Ethylcarbamoyl)ethylcarbanilate	3.54
64* isopropyl (3-chlorophenyl)carbamate	1.95
65* 4-(2,4-dichlorophenoxy)butanoicacid	1.66
66* (2,4,5-trichlorophenoxy)acetic acid	2.18
67* N^2-isopropyl-N^4-methyl-6-methylthio-1,3,5-triazine-2,4-diamine	2.76
68* (2RS)-2-(2,4-dichlorophenoxy)propionic acid	2.54

69* Haloxyfopethoxyethyl	0.11
70* 4-chloro-*o*-tolyloxyacetic acid	2.87
71* (R)-2-(4-chloro-o-tolyloxy)propionic acid	2.93
72* 1-(1,3-benzothiazol-2-yl)-1,3-dimethylurea	1.77
73* 4-amino-6-tert-butyl-4,5-dihydro-3-methylthio-1,2,4-triazin-5-one	3.09
74* methyl 3-(3-methylcarbaniloyloxy)carbanilate	0.67
75* 1-(4-methoxy-6-methyl-1,3,5-triazin-2-yl)-3-[2-(3,3,3-trifluoropropyl)phenylsulfonyl] urea	3.6
76* N^2-tert-butyl-6-chloro-N^4-ethyl-1,3,5-triazine-2,4-diamine	0.93
77* 1-[2-(2-chloroethoxy)phenylsulfonyl]-3-(4-methoxy-6-methyl-1,3,5-triazin-2-yl) urea	2.91
78* S-propyl dipropyl(thiocarbamate)	2.03
79* (RS)-3-(3,5-dichlorophenyl)-5-methyl-5-vinyloxazolidine-2,4-dione	0.53
80* 2-isopropylideneaminooxyethyl (R)-2-[4-(6-chloroquinoxalin-2-yloxy)phenoxy] propionate	-0.2

*validation set

TABLE S-II. Experimental, predicted values of log *S*, and descriptors values for the training and validation sets

N°	log (S_{exp} / mg l^{-1})	log (S_{pred} / mg l^{-1})		MATS8m	RNCG	ALOGP2	MAXDN	Mor26u
		ANN	SVM					
1	2.27	1.97007	2.04153	0.166	0.141	7.524	1.042	0.495
2	2.07918	2.80591	2.84972	-0.1	0.136	3.054	5.054	-0.081
3	1.64345	1.93482	2.03343	0	0.161	12.345	1.435	0.635
4	3.07918	2.98359	2.8582	-0.004	0.143	5.956	5.299	0.103
5	0.39794	0.800011	0.700217	-0.277	0.153	11.403	1.837	-0.136
6	4.4456	3.18637	3.12269	-0.042	0.162	3.456	5.049	-0.053
7	5.15534	4.23823	4.63397	0	0.333	5.064	2.85	-0.088
8	5.74819	4.72705	6.22588	0	0.422	1.409	1.403	0.05
9	2.23	1.72042	1.45922	0.123	0.134	5.456	2.006	-0.003
10	1.97772	1.85824	1.78546	0.13	0.167	11.136	1.387	0.349
11	2.94939	3.18143	2.97726	-0.15	0.262	7.907	2.717	-0.17
12	5.90091	4.34909	5.12943	0	0.349	0.232	0.406	0.139
13	2	3.03824	2.78796	0.049	0.251	9.381	2.129	-0.011
14	3.88649	3.71569	3.90686	0	0.33	11.748	2.854	-0.213
15	0.95424	1.27221	1.30116	0.685	0.115	10.775	2.278	-0.057
16	2.77085	2.2878	1.99897	-0.064	0.208	10.172	2.712	-0.126
17	1.30103	1.0429	0.869131	-0.452	0.14	7.268	1.834	-0.041
18	1.2	1.08567	1.57748	0.144	0.109	12.036	1.033	0.615
19	1.62325	1.83547	1.67145	-1.18	0.222	6.155	1.872	-0.088
20	2.53656	2.3036	2.1091	0	0.187	8.265	1.429	0.263
21	1.69897	2.91823	2.91128	-0.157	0.133	3.143	5.299	0.02
22	-0.09691	-0.127375	-0.117151	-0.107	0.132	22.584	2.076	-0.05
23	0	0.116869	0.503699	-0.135	0.099	28.423	5.696	0.176
24	0.30103	0.116869	0.503699	-0.135	0.099	28.423	5.696	0.176
25	-1.04576	-0.252303	-0.273216	-0.017	0.114	28.798	2.679	0.285
26	4.07918	3.90895	3.96613	0	0.242	4.273	4.902	-0.112
27	1.63749	0.900297	1.70448	-0.002	0.118	19.501	5.797	-0.174
28	1.81291	1.94637	1.76373	-0.595	0.168	5.506	1.771	0.243
29	0.77815	1.93203	1.64045	0.269	0.133	4.686	1.859	0.05
30	1.87506	2.29495	1.96685	-0.564	0.2	5.414	2.059	0.029
31	1.64345	2.04425	1.72174	-0.246	0.184	8.938	2.466	0.036
32	2.8657	2.68557	2.29279	0.842	0.183	9.067	2.653	-0.131
33	3.23	3.26032	2.69202	0.429	0.188	0.284	2.006	-0.032
34	0.17609	1.40374	1.18311	-0.665	0.188	8.466	2.406	-0.249
35	2.83123	2.84747	2.42529	-0.079	0.198	3.241	1.865	0.067
36	3.97772	3.53736	3.24333	-0.069	0.142	1.103	5.273	0.269
37	1.86923	1.0971	1.00116	0.023	0.137	10.733	1.805	0.059
38	2	1.93799	1.82598	0.141	0.173	11.096	1.419	0.319
39	2.63347	3.51021	3.27951	0	0.256	4.701	2.977	-0.203
40	2.86	2.23532	2.0885	0.516	0.13	6.565	1.36	0.449
41	1.52	1.67701	1.93075	0.154	0.131	9.738	1.044	0.593

N°	log (S_{exp} / mg l^{-1})	log (S_{pred} / mg l^{-1})		MATS8m	RNCG	ALOGP2	MAXDN	Mor26u
		ANN	SVM					
42	0.934	1.54919	1.47091	0.151	0.134	8.487	1.476	0.25
43	1.12057	0.544881	0.618408	0.104	0.152	16.889	1.461	-0.001
44	-0.50864	-0.298238	-0.16316	0.023	0.12	21.98	2.07	-0.285
45	3.86332	2.64188	2.9698	-0.125	0.103	1.32	5.522	-0.083
46	0.792	2.3567	2.01921	0.178	0.164	4.658	1.47	0.153
47	3.39794	3.40434	3.06829	0.187	0.159	2.054	1.855	0.748
48	2.85126	2.54508	2.07957	0	0.151	1.016	2.111	0.074
49	1.34	1.40891	1.43875	0.154	0.126	8.691	1.305	0.301
50	3.79727	3.38015	3.17407	-0.06	0.136	1.007	5.219	0.183
51	1.44716	1.28912	1.2225	0.13	0.182	13.888	1.528	-0.038
52	0.60206	0.645302	0.943595	-1.126	0.186	15.682	1.683	0.242
53	3.17609	3.08267	3.19066	-0.075	0.143	1.578	5.285	-0.122
54	3.90849	3.10932	2.99134	-0.415	0.274	9.449	2.797	-0.114
55	1.36173	0.609083	0.590257	-0.068	0.159	20.456	2.21	0.174
56	2.04139	2.43526	2.38896	-1.27	0.167	4.387	5.654	-0.093
57	0.69897	-0.319298	-0.0707111	0.337	0.129	34.033	1.992	0.418
58	2.04139	2.14982	2.2675	-0.34	0.122	10.701	5.922	0.123
59*	1.52	2.10053	1.9355	0.163	0.143	6.43	1.473	0.362
60*	2.8451	2.35458	1.93893	0	0.146	1.99	2.009	0.096
61*	3.54407	3.14343	2.65957	0.368	0.135	0.899	2.281	0.475
62*	1.94939	1.15193	1.07316	-0.042	0.148	10.255	2.59	-0.233
63*	1.66276	2.2624	1.95081	-0.122	0.205	10.036	2.499	-0.034
64*	2.17609	2.79982	2.72653	-0.383	0.275	12.085	2.757	-0.172
65*	2.76	2.32061	2.35449	0	0.152	5.732	1.076	0.557
66*	2.54407	2.2878	1.99897	-0.064	0.208	10.172	2.712	-0.126
67*	0.11394	0.558332	1.07797	0.028	0.101	23.862	5.797	0.082
68*	2.8657	3.45844	2.9976	0.681	0.228	6.936	2.658	-0.058
69*	2.92942	2.68557	2.29279	0.842	0.183	9.067	2.653	-0.131
70*	1.77085	2.57556	2.12201	0.163	0.171	3.16	1.825	0.029
71*	3.09	3.46152	3.13894	0	0.177	2.15	1.947	0.688
72*	0.6721	1.15908	1.27212	0.767	0.116	11.696	2.277	-0.122
73*	3.60206	2.56352	2.75728	-0.086	0.137	8.116	5.697	-0.039
74*	0.929	2.03968	1.98117	0.151	0.136	7.511	1.479	0.494
75*	2.91116	3.07567	3.00836	-0.063	0.148	3.395	5.116	-0.003
76*	2.03342	1.90878	1.81662	0.141	0.173	11.55	1.412	0.338
77*	0.53148	0.304822	0.214497	-0.789	0.14	13.412	2.605	-0.164
78*	-0.20066	0.0957958	0.142756	-0.006	0.102	17.699	2.16	0.077
79*	2.38561	2.2525	2.3682	0.047	0.129	11.456	5.561	0.053
80*	1.25527	1.04243	0.975408	0.186	0.141	9.759	1.755	-0.189

*Validation set