Studii şi Cercetări Ştiinţifice Universitatea Bacău Seria Biologie
ISSN 1224-919X
Vol. 12, Issue 1, 2007

TRIAZINES HERBICIDAL ASSESSED ACTIVITY

Lorentz JÄNTSCHI, Sorana Daniela BOLBOACĂ

Abstract
Correlation coefficients and associated squared values are used as assessment parameters in validation of structure-activity relationships. Seven correlation coefficients were calculated for models that characterized the relationship between chemical structure the herbicidal activity of a triazine analogues series. Three previously reported models were compared by using Pearson (r_{Prs}), Spearman (ρ_{Spm}), Kendall's (τ_a, τ_b, τ_c), Gamma correlation coefficients (Γ), and a new proposed Semi-Quantitative (r_{sQ}). The results of investigation, expressed as correlation coefficients and associated 95% confidence intervals, squared correlation coefficient, Student's t parameters, respectively the Z parameter were calculated for each model, and the values are discussed. This approach proved to be a good tool for chemical structure versus biological activity relationships assessment.
Keywords: Triazine Analogues, Herbicidal Activity, Activity Assessment, Structure-Activity Relationships

INTRODUCTION

Structure-activity relationships (SARs) and quantitative structure-activity relationships (QSARs), collectively referred to as (Q)SARs, are theoretical models usually used to predict the physicochemical properties and/or biological activities of chemicals from the knowledge of structure (see Rogers 1994, Diudea 2001).

The validation of (Q)SARs, defined as the process by which the reliability (yielding the same or compatible results for a compounds or class of compounds), and relevance (the capability to correctly predict the biological effect of interest), are still debated at academic regulatory and regulated communities (see Worth 2004, Gramatica 2004).

To be consider valid, a (Q)SAR model must (see Worth 2004) had:
(1) a defined endpoint that is serves to predict;
(2) an unambiguous and easily applicable algorithm;
(3) a clear mechanistic basis;
(4) a well defined domain of applicability;
(5) an associated measures of its goodness-of-fit and internal goodness-of-prediction;
(6) been assessing in terms of its predictive power by using data that were not used in the development of the model.

Several approaches have been proposed for QSAR development, for which the simple and multiple linear regressions are frequently used by many researchers (Hansch 1995, Afantitis 2006, Deconinck 2007). The Pearson correlation coefficients and associated squared values are the most measures used in analyzing the statistical fit of linear regression models (Ma 2006, Isayev 2006, Si 2006, Lu 2007). Pearson correlation coefficient is computed based on the assumption that the measured and estimated/predicted activities of interest are quantitative variables (interval or ratio variables, Rosner 1995). In practice, the measured activity depends on many external factors (from which the most important are the measurement and experimental abilities of the researcher, and the performances of device used) and it could be classified more correctly as a semi-quantitative variable than as a quantitative variable. In these conditions, is the Pearson correlation coefficient the best statistical estimator for quantifying the relationship between dependent (e.g. herbicidal activity of triazine analogues) and one or more independent variables (e.g. measured herbicidal activity)?

The herbicidal activity of a sample of triazine analogues was previous modeled and some models that used topological descriptors (Diudea 2002) and molecular descriptors (Bolboaca 2006 - using the molecular descriptors family on the structure-activity relationship approach - Jäntschi 2005) were reported. Starting from the hypothesis that the herbicidal activity of triazine analogue is not pure quantitative variable, the aim of the research was to analyzed three previous reported models by using Pearson, Semi-Quantitative, Spearman, Kendall's and Gamma correlation coefficients.

MATERIAL & METHOD

Triazine analogues act as inhibitors of photosynthesis, by attaching to specific proteins in thylakoids and blocking electron transport between the photosystems. Herbicidal activity of a series of triazine analogue was previous studied and three models were included into analysis: one model reported by Diudea (2002) and two models reported by Bolboaca (2006). The models and associated statistical characteristics are presented in Eq.(1) - Eq.(3):

$$\hat{Y}_{pI50} = 9.614 - 0.153 \cdot X_5 - 58.888 \cdot 1/V_5 - 2.430 \cdot 1/N_3$$
$$n = 30;\ r^2 = 0.9694;\ F = 274.3;\ r_{loo} = 0.9778 \quad (1)$$

where $\hat{Y}_{pI50}$ = herbicidal activity estimated by Eq.(1), pI_{50} = 50% inhibition of Hill reaction on Chlorella, X_5

= topological descriptor for substituent number 5, V_5 = fragmental volumes of the substituent in the position 5 (cm^3/mol), N_3 = total number of hydrogen's in the substituent 3, r^2 = squared correlation coefficient, F = Fisher parameter, and r_{loo} = squared correlation coefficient obtained after leave-one-out analysis (see Diudea 2002).

$$\hat{Y}_{2v} = 5.52 - 8112.2 \cdot iSMMWHg + 194.35 \cdot iSMmEQt \quad (2)$$
$$n = 30;\ r = 0.987,\ r^2 = 0.975;F_{est} = 533\ (p < 0.001);\ r_{cv\text{-}loo}^2 = 0.971$$

where $\hat{Y}_{2v}$ = herbicidal activity of trianzine analogues estimated by the MDF-SAR model with two descriptors; *iSDRFHg*, and *iSMmEQt* = molecular descriptors; n = sample size; r = correlation coefficient; r^2 = squared correlation coefficient; F_{est} = Fisher parameter (and its significance p); $r_{cv\text{-}loo}^2$ = cross-validation leave-one-out score (see Diudea, 2001).

$$\hat{Y}_{4v} = 5.75 + 199 \cdot iSMmEQt - 9010 \cdot iSMMWHg - 0.071 \cdot LADmkQt + 2.86 \cdot INPRJQg \quad (3)$$
$$n = 30;\ r = 0.994,\ r^2 = 0.988;\ F_{est} = 537\ (p < 0.001);\ r_{cv\text{-}loo}^2 = 0.985$$

where $\hat{Y}_{4v}$ = herbicidal activity of trianzine analogues estimated by the MDF-SAR model with four descriptors; *iSMmEQt*, *iSMMWH*, *LADmkQt*, and *INPRJQg* = molecular descriptors (see Bolboaca 2006).

The measured herbicidal activity of triazine analogues (pI_{50}) and the value estimated by the models from Eq.(1) ($\hat{Y}_{pI50}$), Eq.(2) ($\hat{Y}_{2v}$) and Eq.(3) ($\hat{Y}_{4v}$) are presented in Table 1 (see Diudea 2002, Bolboaca 2006).

Table 1. Herbicidal activity of triazine analogues: measured and estimated by Eq.(1), Eq.(2), and Eq.(3)

Mol	pI_{50}	$\hat{Y}_{pI50}$ - Eq.(1)	$\hat{Y}_{2v}$ - Eq.(2)	$\hat{Y}_{4v}$ - Eq.(3)
t01	3.82	3.88	3.88	3.83
t02	5.20	5.09	5.14	5.20
t03	5.34	5.50	5.42	5.33
t04	5.83	5.70	5.84	5.83
t05	6.01	6.10	5.83	5.96
t06	6.39	6.51	6.39	6.43
t07	6.75	6.71	6.70	6.76
t08	6.76	6.71	6.81	6.72
t09	6.74	6.83	6.65	6.74
t10	6.76	6.83	6.74	6.78
t11	6.78	6.83	6.80	6.85
t12	7.12	6.91	7.14	7.07
t13	6.82	6.59	6.84	6.80
t14	6.74	6.79	6.92	6.87
t15	6.89	6.79	6.96	6.86
t16	6.95	6.91	6.83	6.83
t17	7.01	6.91	6.87	7.02
t18	6.87	6.91	6.94	6.89
t19	6.97	6.91	7.04	6.95
t20	6.94	6.99	6.98	6.95
t21	7.21	7.05	7.08	7.08
t22	7.01	7.09	6.96	7.03
t23	6.81	7.13	6.87	6.95
t24	6.45	6.52	6.68	6.66
t25	6.75	6.65	6.74	6.67
t26	6.75	6.65	6.77	6.76
t27	6.71	6.77	6.67	6.62
t28	6.88	6.77	6.76	6.77
t29	6.70	6.77	6.72	6.76
t30	6.69	6.85	6.66	6.65

The relationships between measured herbicidal activity and estimated by Eq.(1) - Eq.(3) have been investigated by using seven correlation coefficients (Eq.(4) – Eq.(10)), associated 95% confidence intervals (95%CI), and squared correlation coefficients:

÷ Pearson product-moment correlation coefficient (r_{Prs}):

$$r_{Prs} = \frac{\sum\left(Y_{m_i} - \overline{Y}_m\right)\left(Y_{est_i} - \overline{Y}_{est}\right)}{\sqrt{\left[\sum\left(Y_{m_i} - \overline{Y}_m\right)^2\right]\left[\sum\left(Y_{est_i} - \overline{Y}_{est}\right)^2\right]}} \quad (4)$$

where Y_{mi} = value of the measured herbicidal activity for compound i (where i = *1, 2, …, 30*); $\overline{Y}_m$ = average of the measured herbicidal activity; Y_{esti} = value of the estimated herbicidal activity for compound i, and $\overline{Y}_{est}$ = average of the estimated herbicidal activity.

÷ Spearman rank correlation coefficient (ρ_{Spm}):

$$r_{Spm} = \frac{\sum\left(R_{Y_{m_i}} - \overline{R}_{Y_m}\right)\left(R_{Y_{est_i}} - \overline{R}_{Y_{est}}\right)}{\sqrt{\left[\sum\left(R_{Y_{m_i}} - \overline{R}_{Y_m}\right)^2\right]\left[\sum\left(R_{Y_{est_i}} - \overline{R}_{Y_{est}}\right)^2\right]}} \quad (5)$$

where R_{Ymi} = rank of the measured herbicidal activity for compound i, $\overline{R}_{Y_{m_i}}$ = average rank of the measured herbicidal activity, R_{Yesti} = rank of the estimated herbicidal activity for compound i (where i *=1, 2, …, 30*), and $\overline{R}_{Y_{est_i}}$ = average rank of the estimated herbicidal activity.

÷ Semi-Quantitative correlation coefficient (r_{sQ}):

$$r_{sQ} = \sqrt{\frac{\sum\left(Y_{m_i} - \overline{Y}_m\right)\left(Y_{est_i} - \overline{Y}_{est}\right)}{\sqrt{\left[\sum\left(Y_{m_i} - \overline{Y}_m\right)^2\right]\left[\sum\left(Y_{est_i} - \overline{Y}_{est}\right)^2\right]}}} \cdot \sqrt{\frac{\sum\left(R_{Y_{m_i}} - \overline{R}_{Y_m}\right)\left(R_{Y_{est_i}} - \overline{R}_{Y_{est}}\right)}{\sqrt{\left[\sum\left(R_{Y_{m_i}} - \overline{R}_{Y_m}\right)^2\right]\left[\sum\left(R_{Y_{est_i}} - \overline{R}_{Y_{est}}\right)^2\right]}}} \quad (6)$$

÷ Kendall's tau rank correlation coefficient ($\tau_{Ken,a}$, $\tau_{Ken,b}$, $\tau_{Ken,c}$):

$$\tau_{Ken,a} = \frac{2(C - D)}{n(n-1)} \quad (7)$$

where $\tau_{Ken,a}$ = Kedall tau-a correlation coefficient, C = number of concordant pairs (measured activity and estimated activity), and D = number of discordant pairs.

$$\tau_{Ken,b} = \frac{C - D}{\sqrt{\left[n(n-1) - 2t\right]\left[n(n-1) - 2u\right]}} \quad (8)$$

where $\tau_{Ken,b}$ = Kedall tau-b correlation coefficient, t = number of tied of the measured values and u = number of tied of the estimated values.

$$\tau_{Ken,c} = \frac{2(C-D)}{n^2} \quad (9)$$

where $\tau_{Ken,c}$ = Kedall tau-c correlation coefficient, n = sample size.

÷ Goodman - Kruskal Gamma correlation coefficient (Γ):

$$\Gamma = \frac{C-D}{C+D} \quad (10)$$

where Γ = Goodman - Kruskal Gamma correlation coefficient.

RESULTS & DISCUSSION

The equations (4) - (10) were applied on previous obtained models (Eq.(1) - Eq.(3)) and the obtained results, express as squared correlation coefficients and associated Student's test (t), respectively Z parameter (Z) (for a significance level of 5%) are presented (see Table 2) and were analyzed.

Table 2. Triazine analogues: squared correlation coefficients and associated significance

	Eq.(1)		Eq.(2)		Eq.(3)	
Coefficient name	r^2	t	r^2	t	r^2	t
Pearson	0.9694	29.77†	0.9712	30.71†	0.9850	42.86†
Spearman	0.6648	7.45†	0.7485	9.13†	0.8163	11.15†
Semi-Quantitative	0.8028	10.68†	0.8526	12.73†	0.8967	15.59†
	r^2	Z	r^2	Z	r^2	Z
Kendall τ_a	0.4084	4.96†	0.5079	5.53†	0.5931	5.98†
Kendall τ_b	0.4299	4.98†	0.5174	5.54†	0.6028	5.98†
Kendall τ_c	0.3816	4.81†	0.4746	5.35†	0.5542	5.78†
Gamma	0.4423	3.43†	0.5246	4.07†	0.6098	4.73†

Eq.(1) = model reported by Diudea (2002); Eq.(2) = MDF-SAR model with two descriptors (Bolboacă 2006); Eq.(3) = MDF-SAR model with four descriptors (idem); † $p < 0.001$
r^2 = squared correlation coefficient; t = Student t parameter; Z = Z test parameter

The following can be observed analyzing the data presented in Table 2:

÷ The correlation coefficients obtained by applying of all seven methods on the investigated models (Eq.(1) - Eq.(3)) were statistical significant ($p < 0.0001$).

÷ The greater value of the squared correlation coefficients was obtained, without any exception, by Eq.(3) (seven cases out of seven possibilities). This observation sustained the goodness-of-fit of the MDF-SAR model with four descriptors (Eq.(4)).

÷ The smaller value of the difference between the squared correlation coefficient obtained by Eq.(3) and by Eq.(1) was observed for Pearson method ($r^2_{Prs-Eq.(3)} - r^2_{Prs-Eq.(1)} = 0.0156$); the greatest value of difference was obtained for Kendal tau-a method ($r^2_{Ken,a-Eq(3)} - r^2_{Ken,a-Eq.(1)} = 0.1847$).

÷ The lowest value of squared correlation coefficient was obtained, without any exception, by the model from Eq.(1). The lowest value of squared correlation coefficient was obtained by Eq.(1) with Kendall τ_c method. Remark that the Kendall τ_c method obtained the lowest value for all models (Eq.(1) - Eq(3), see Table 3). Looking at the Kendall's τ_c method it can be observed that the absolute difference between squared correlation coefficient obtained by Eq.(1) and by Eq.(3) ($|r^2_{Ken,c-Eq.(1)} - r^2_{Ken,c-Eq.(3)}| = 0.1725$) is almost two times greater comparing with the absolute difference between squared correlation coefficient obtained by Eq.(1) and by Eq.(2) ($|r^2_{Ken,c-Eq.(1)} - r^2_{Ken,c-Eq.(2)}| = 0.0929$).

÷ The values of Pearson squared correlation coefficients were very close to each other for all investigated models (Eq.(1) - Eq.(3)). These differences varied from 0.0018 ($r^2_{Prs-Eq.(2)} - r^2_{Prs-Eq.(1)}$) to 0.0156 ($r^2_{Prs-Eq.(3)} - r^2_{Prs-Eq.(1)}$). Analyzing the squared correlation coefficients obtained by Eq.(1) and Eq.(2) it can be say that these models are similar to each other in terms of the quality of the fit between model-calculated and measured herbicidal activity of studied triazine analogues.

÷ The values obtained by applying the Spearman method were not as closed to each other as the values obtained by applying the Pearson method. For Spearman squared correlation coefficient, the differences varied from 0.0678 ($\rho^2_{Spm-Eq.(3)} - \rho^2_{Spm-Eq.(2)}$) to 0.1515 ($\rho^2_{Spm-Eq.(3)} - \rho^2_{Spm-Eq.(1)}$). If the Spearman squared correlation coefficients are analyzed it cannot be say that the models from Eq.(1) and Eq.(2) are as similar as resulted when the Pearson squared correlation coefficients are analyzed.

÷ The values of Semi-Quantitative squared correlation coefficient, as was expected, had values somewhere between the values obtained with Pearson and Spearman methods.

÷ For all three Kendall tau methods, the smallest difference was obtained between squared correlation coefficient obtained by Eq.(3) and the squared correlation coefficient obtained by Eq.(2) ($\tau_{Ken,a} = 0.0852$; $\tau_{Ken,b} = 0.0854$; $\tau_{Ken,c} = 0.0796$); the greatest difference between squared correlation coefficient obtained by Eq.(3) and by Eq.(1) ($\tau_{Ken,a} = 0.1847$; $\tau_{Ken,b} = 0.1728$; $\tau_{Ken,c} = 0.1725$).

÷ Regarding the Gamma squared correlation coefficients, the results are similar with those obtained for Pearson squared correlation coefficient; the smallest value is obtained for the difference between squared correlation coefficient obtained by Eq.(3) and that obtained by Eq.(1) ($r^2_{\Gamma-Eq.(2)} - r^2_{\Gamma-Eq.(1)} = 0.0823$). The greatest value is obtained for the difference between squared correlation coefficient obtained by Eq.(3) and by Eq.(1) ($r^2_{\Gamma-Eq.(3)} - r^2_{\Gamma-Eq.(1)} = 0.1675$).

The plots of the correlation coefficients and 95%CI for the investigated models (Eq.(1)-Eq.(3)) according with the computational method (Eq.(4)-Eq.(10)) are presented in figures:

÷ 1 - Pearson's correlation coefficient (r_{Prs});

÷ 2 - Spearman's rank correlation coefficient (ρ_{Spm});

÷ 3 - Semi-Quantitative correlation coefficient, (r_{sQ});

÷ 4 - Gamma correlation coefficient (Γ);
÷ 5 - Kendall's tau correlation coefficients ($\tau_{Ken,a}$, $\tau_{Ken,b}$, $\tau_{Ken,c}$).

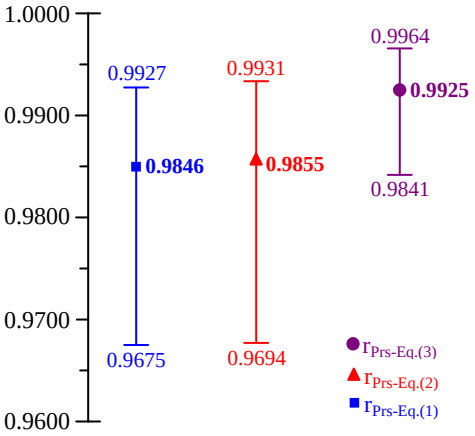

Figure 1. Pearson correlation coefficients and 95%CI for studied models

The values obtained with Pearson method are similar to each other for the models obtained by Eq.(1) and Eq.(2) regarding the correlation coefficient as well as the 95%CI (see Figure 1). The MDF-SAR model with four descriptors (Eq.(3)) had the greatest value of correlation coefficient and a narrow width of 95%CI comparing with the MDF-SAR model with two descriptors (Eq.(2)) and/or with the model obtained by Eq.(1) (see Diudea 2002).

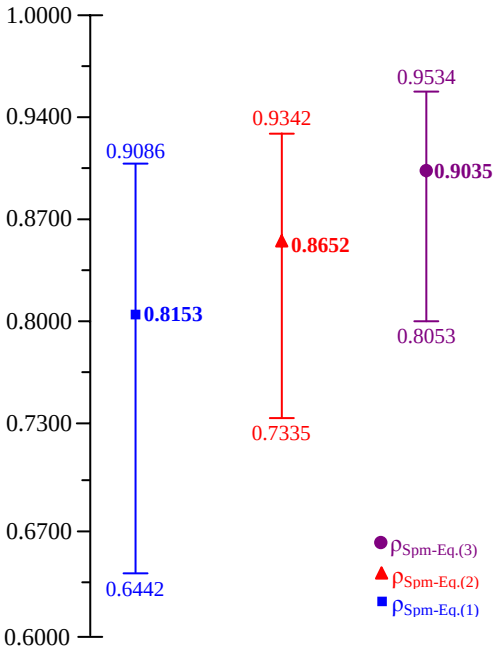

Figure 2. Spearman correlation coefficients and 95%CI for studied models

Analyzing the values obtained by Spearman method it cannot be say that the model obtained by Eq.(2) is similar with that obtained by Eq.(1). It is well known that Spearman's rank correlation is satisfactory for testing the null hypothesis of no relationship (which was not the aim of our research), and it is difficult to be interpreted as a measure of the strength of the relationship (Bland, 1995). The MDF-SAR model with four descriptors had a narrow 95%CI width, followed by the MDF-SAR model with descriptors (see Figure 2). Analyzing the values obtained with the Semi-Quantitative method (see Figure 3) an identical behavior can be observed.

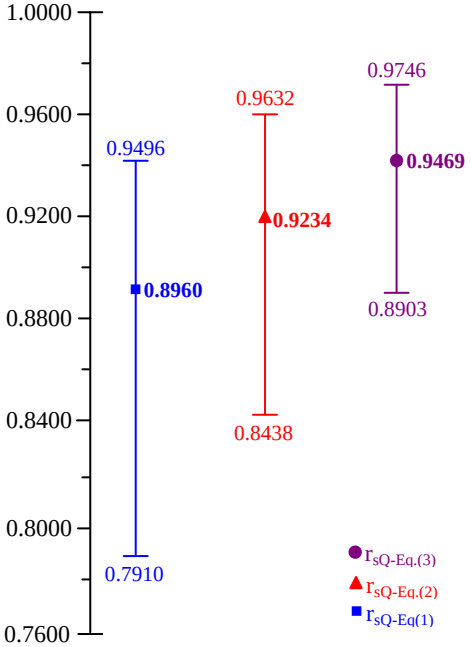

Figure 3. Semi-Quantitative correlation coefficients and 95%CI for studied models

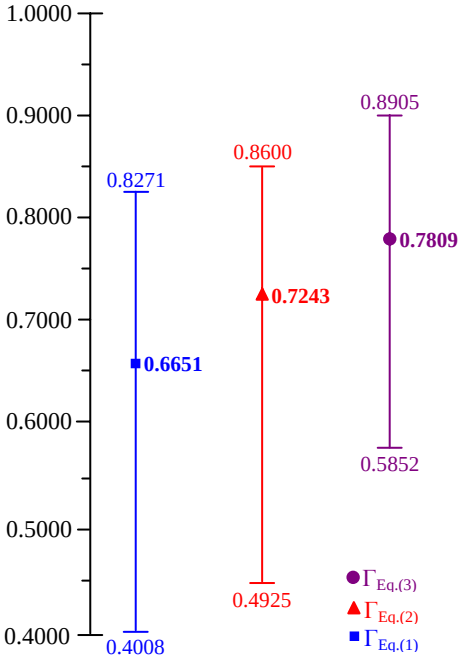

Figure 4. Gamma correlation coefficients and 95%CI for studied models

Analyzing the results obtained with Kendall's and Gamma methods the observation remain the same as were resulted from analyzing of the Spearman method: the MDF-SAR model with four descriptors obtained the greatest value of the correlation coefficient and the lowest width of 95%CI.

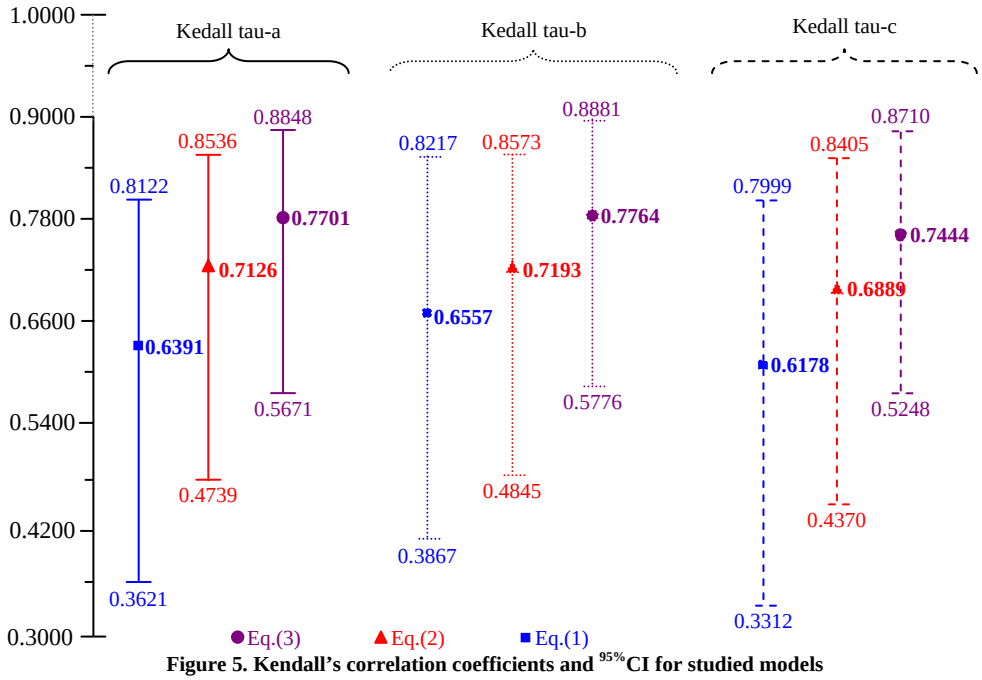

Figure 5. Kendall's correlation coefficients and $^{95\%}$CI for studied models

Analyzing the width of the 95%CI the following can be noticed:

÷ The narrowest $^{95\%}$CI widths are obtained with Pearson method (the difference between the upper and lower boundary were: 0.0252 for Eq.(1), 0.0237 for Eq.(2), and 0.0123 for Eq.(3));

÷ The widest $^{95\%}$CI widths are obtained with Kendall tau-c method for all three studied models (the difference between the upper and lower boundary were: 0.4687 for Eq.(1), 0.4035 for Eq.(2), and 0.3462 for Eq.(3));

÷ The classification of the $^{95\%}$CI width in ascending order according with the computational method is: (1) Pearson, (2) Semi-Quantitative, (3) Spearman, (4) Gamma, (5) Kendall tau-b, (6) Kendall tau-a, (7) Kendall tau-c;

÷ Analyzing the Pearson correlation coefficient and associated $^{95\%}$CI it can be say that the model obtained by Eq.(1) is similar with the model obtained by Eq.(2). The other computational methods (Eq.(5) - Eq.(10)) did not sustain this observation.

CONCLUSIONS

The correlation coefficients and associated squared values are simple measures able to quantify the relationships between dependent variable (in our case herbicidal activity) and one or more independent variables (in our case topological descriptors - Eq.(1) and molecular descriptors - Eq.(2), and Eq.(3)), being used as measures of the statistical fit of models.

All seven computational methods used to evaluate the relationships between measured and estimated herbicidal activity obtained with Eq.(1), Eq.(2), and Eq.(3) were statistical significant, sustaining the goodness-fit-of the models.

The method of computing the correlation coefficient should be carefully chooses according with the type of variable. Because herbicidal activity of triazine analogues is rather a semi-quantitative variable (the measurement of the herbicidal activity depend on quality of Chlorella culture, quality of material used in experiment, researcher experimental abilities, experience and expertise) then a quantitative variable, it can be consider that the Semi-Quantitative proposed method (Eq.(6)) is the best evaluator for the models from Eq.(1), Eq.(2), and Eq.(3).

ACKNOWLEDGEMENTS

The research was partly supported by UEFISCSU Romania through projects ET36/2005 & ET108/2006.

REZUMAT

Coeficienţii de corelaţie şi pătratul valorilor lor sunt folosiţi ca parametrii de evaluare în validarea relaţiilor structură chimică - activitate biologică. Şapte coeficienţi de corelaţie au fost calculaţi pe modelele care caracterizează prin intermediul relaţiei cu structura chimică activitatea erbicidă a unei serii de derivaţi ai triazinei.

Valorile determinate experimental ale activităţii erbicide şi valorile estimate de trei modele raportate anterior în literatură au fost comparate utilizând următorii coeficienţi de corelaţie: Pearson (r_{Prs}), Spearman (ρ_{Spm}), Kendall (τ_a, τ_b, τ_c), Gamma (Γ), şi Semi-Cantitativ (r_{sQ}), acesta din urmă fiind nou propus. Rezultatele studiului cumulează coeficienţii de corelaţie, intervalele de încredere cu o probabilitate de 95% ale acestora, coeficienţii de determinare, parametrul t (Student), şi parametrul Z (Fisher). Demersul ştiinţific al lucrării permite o fină evaluare a relaţiilor semicantitative structură chimică - activitate biologică.

REFERENCES

1. Afantitis A., Melagraki G., Sarimveis H., Koutentis P. A., Markopoulos J., Igglessi-Markopoulou O., 2006 - *A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes*, Bioorg. Med. Chem., 14(19), p. 6686-6694, MEDLINE PMID: 16782350.
2. Bland M., 1995 - *An Introduction to Medical Statistics*, Oxford University Press; Oxford, New York & Tokyo, p. 205-225.
3. Bolboacă S., Jäntschi L., 2006 - *Molecular Descriptors Family on Structure-Activity Relationships: Modeling Herbicidal Activity of Substituted Triazines Class*, Bulletin of University of Agricultural Sciences and Veterinary Medicine - Agriculture (AcademicPres, Cluj-Napoca), 62, p. 35-40.
4. Deconinck E., Coomans D., Vander Heyden Y., 2007 - *Exploration of linear modelling techniques and their combination with multivariate adaptive regression splines to predict gastro-intestinal absorption of drugs*, J. Pharm. Biomed. Anal., 43(1), p. 119-130, MEDLINE PMID: 16859855.
5. Diudea M. V., Gutman I., Jäntschi L., 2001 - *Molecular Topology*, Nova Science Publishers, New York, USA, p. 1-332, AGRICOLA CNB: QD461 .D55 2001.
6. Diudea M. V., Jäntschi L., Pejov L., 2002 - *Topological Substituent Descriptors*, LEJPT, 1(1), p. 1-18 SCIENTIFIC COMMONS: 15879909.
7. Hansch C., Leo A., 1995 - *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, American Chemical Society, Washington DC, p. 1-543, AGRICOLA CNB: QP517.S85E98 1995.
8. Gramatica P., 2004 - *Evaluation of Different Statistical Approaches for the Validation of Quantitative Structure - Activity Relationships*, Joint Research Centre, The European Commission, Final raport of JRC Contract ECVA-CCR.496576-Z.
9. Isayev O., Rasulev B., Gorb L., Leszczynski J., 2006 - *Structure-toxicity relationships of nitroaromatic compounds*, Mol. Divers., 10(2), p. 233-245, PUBMED PMID: 16710810.
10. Jäntschi L., 2005 - *Molecular Descriptors Family on Structure Activity Relationships 1. Review of the Methodology*, LEJPT, 4(6), p. 76-98, SCIENTIFIC COMMONS: 18187794.
11. Lu A., Zhang J., Yin X., Luo X., Jiang H., 2007 - *Farnesyltransferase pharmacophore model derived from diverse classes of inhibitors*, Bioorg. Med. Chem. Lett., 17(1), p. 243-249, MEDLINE PMID: 17049856.
12. Ma W., Luan F., Zhao C., Zhang X., Liu M., Hu Z., Fan B., 2006 - *QSAR Prediction of the Penetration of Drugs Across a Polydimethylsiloxane Membrane*, QSAR & Combinatorial Sciences, 25(10), p. 895-904.
13. Rogers D., Hopfinger A. J., 1994 - *Application of genetic function approximation to quantitative structure-activity relationships and quantitative structure-property relationships*, J. Chem. Inf. Comput. Sci., 34, p. 854-866, AGRICOLA CNB: 241.64 J82.
14. Rosner B., 2000 - *Fundamentals of Biostatistics*, 5th Edition, Duxbury, Pacific Grove, California, USA, p. 1-792, AGRICOLA CNB: QH323.5 .R674 2000.
15. Si H. Z., Wang T., Zhang K. J., Hu Z. D., Fan B. T., 2006 - *QSAR study of 1,4-dihydropyridine calcium channel antagonists based on gene expression programming*, Bioorg. Med. Chem., 14(14), p. 4834-4841, MEDLINE PMID: 16580211.
16. Worth P. A., Cronin M. T. D., 2004 - *Report of the Workshop on the Validation of QSARs and Other Computational Prediction Models*, ATLA (Alternatives To Laboratory Animals), 32(1B), 703-706.

AUTHOR'S ADDRESS

Lorentz JÄNTSCHI: *Technical University of Cluj-Napoca, Romania, 103-105 Muncii Bvd., 400641 Cluj-Napoca, Cluj, Romania,*
http://lori.academicdirect.org,
lori@academicdirect.org

Sorana Daniela BOLBOACĂ: *"Iuliu Hatieganu" University of Medicine and Pharmacy, 6 Louis Pasteur Str., 400349 Cluj-Napoca, Cluj, Romania,*
http://sorana.academicdirect.ro,
sbolboaca@umfcluj.ro