

مطالعه ارتباط کمی ساختار- فعالیت برای پیش‌بینی شاخص LD₅₀ در آفت‌کش‌های ارگانوفسفات

مینا کیانپور^۱، عصمت محمدی نسب^{۱*}

۱- به‌ترتیب دانشجوی دکتری و استادیار، دانشگاه آزاد اسلامی، واحد اراک، گروه شیمی، اراک، ایران

چکیده

سموم ارگانوفسفات از جمله سموم شیمیایی خطرناک برای سلامت انسان محسوب می‌شوند. بسیاری از محققین، با وجود رعایت اصول ایمنی و جلوگیری از مواجهه با خطرات ناشی از استفاده از ترکیبات شیمیایی، به‌منظور بررسی میزان سمیت ترکیبات ارگانوفسفات در تماس با این سموم قرار دارند و احتمال جذب این سموم از طریق پوست وجود دارد. مطالعه ارتباط ساختار- فعالیت با کمک روش‌ها و مدل‌های تئوری پیش‌بینی کننده، با صرف حداقل وقت و هزینه، امکان دست‌یابی به داده‌ها، اطلاعات و خواص فیزیکی- شیمیایی ترکیبات مورد نظر را فراهم می‌نماید. در این مطالعه، روش‌های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی با مدل پرسپترون چند لایه با هدف بررسی ارتباط کمی شاخص سمیت LD₅₀ با برخی توصیف‌گرهای مولکولی، در برخی ترکیبات ارگانوفسفات به‌کار گرفته شد. بررسی مقادیر ضرایب همبستگی و میزان جذر خطای مجذور میانگین مدل‌های پیشنهادی در این مطالعه نشان داد که روش شبکه عصبی مصنوعی در پیش‌بینی شاخص سمیت LD₅₀ در ترکیبات ارگانوفسفات نسبت به روش رگرسیون خطی چندگانه، از برتری بسیار بالایی برخوردار می‌باشد.

واژه‌های کلیدی: ترکیبات ارگانوفسفات؛ متوسط دوز کشنده؛ رابطه کمی ساختار-فعالیت؛ شبکه عصبی مصنوعی؛ روش رگرسیون خطی چندگانه

* نویسنده رابط، پست الکترونیکی: e-mohammadinasab@iau-arak.ac.ir

تاریخ دریافت مقاله: ۹۷/۸/۲۶- تاریخ پذیرش مقاله: ۹۸/۱/۲۰



مقدمه

اغلب ترکیبات ارگانوفسفرفه در زمان جنگ جهانی دوم توسط دانشمند آلمانی به نام گرهارد شرادر کشف گردیده‌اند. ترکیب شرادان اولین ترکیب فسفره است که در سال ۱۹۴۱ کشف شد. سموم فسفره یکی از گروه‌های سموم هستند که تنوع و گروه‌های متعدد بسیاری دارند. بیشتر ترکیبات ارگانوفسففات به‌عنوان حشره‌کش و آفت‌کش در کشاورزی و دامپروری به‌کار می‌روند و دارای زمان ماندگاری کوتاه در خاک، هوا و غذای حیوانات هستند (Ebrahimzadeh, 2005). همچنین برخی آن‌ها به‌عنوان مرگ موش در مصارف خانگی مورد استفاده قرار می‌گیرند (Sattler, 2007). این ترکیبات به سرعت از راه پوست، مخاط، دستگاه گوارش و تنفس جذب شده و اثر آن‌ها طی ۱-۰/۵ ساعت شروع و پس از ۸-۲ ساعت به حداکثر می‌رسد. زمان ماندگاری آن‌ها از چند روز تا چند هفته متفاوت است (Rodgers, 1997) و طیف تاثیر گسترده آن‌ها باعث ظهور آفات ثانوی می‌شود. ضمن داشتن تاثیر سریع و مقاومت در آفات، به‌ویژه آفات چند نسلی، برای انسان و جانوران خون‌گرم، با مسمومیت شدید همراه هستند، سیستم عصبی را مختل می‌نمایند (Tafari, 1987). در صورت مواجهه، سبب مهار آنزیم بوتیریل کولین استراز می‌گردند. با توجه به نقش این آنزیم در رشد و نمو سیستم عصبی، مهار و کاهش فعالیت این آنزیم می‌تواند سبب تجمع انتقال‌دهنده کولین استراز در سیستم عصبی شده و در نتیجه سیستم عصبی مرکزی و محیطی را تخریب کند (Lopes Soares, 2012; Barkhordri, 2012). مکانیسم مهار این آنزیم به اثبات رسیده و برخی مطالعات، مهار برگشت ناپذیر فعالیت آنزیم کولین استراز ناشی از مسمومیت سموم ارگانوفسفرفه که موجب اختلال در عملکرد اعضای مختلف بدن می‌شود را تایید نموده‌اند (De Silva, 2006; Ranjbar, 2005). مهار آنزیم کولین استراز توسط سموم ارگانوفسفرفه در ابتدا به وسیله پیوند یونی صورت می‌گیرد که در نهایت به سمت پیوند کووالانسی غیر قابل برگشت پیشرفت می‌کند. این فرآیند که ۲۴ الی ۴۸ ساعت به طول می‌انجامد باعث برگشت فرآیند مهار آنزیم می‌گردد (Bakand, 2012; Maroni, 2000). مطالعات صورت گرفته نشان داده است، در کشورهای در حال توسعه که ۸۵ درصد از تولید جهانی آفت‌کش‌ها را مصرف می‌کنند، میزان وقوع مسمومیت ۱۳ برابر کشورهای صنعتی است (Mahmoudi, 2008). تنها در کشور چین تخمین زده می‌شود که مسمومیت با آفت‌کش‌ها غالباً از نوع سموم ارگانوفسفرفه بوده و سالانه سبب مرگ حدود ۱۷۵۰۰۰ نفر می‌شود (Eddleston, 2004; Moghaddamnia, 1999). طبق آمارهای منتشر شده تقریباً از ۳ میلیون مورد در سال که دچار مسمومیت با ترکیبات ارگانوفسفر شده‌اند نزدیک به ۳۰۰ هزار نفر دچار مرگ یا صدمات جدی به‌ویژه در کشورهای در حال توسعه می‌شوند (Aghilinejad, 2007). تاکنون مطالعات زیادی در زمینه مسمومیت با این ترکیبات در انسان و در ایران انجام یافته است (Razzaghi, 2016). تعیین مقادیر شاخص سمیت در آفت‌کش‌ها از روش‌های آزمایشگاهی با توجه به امکان جذب این سموم از طریق پوست همراه با خطرات و مشکلات جدی است. تاکنون مطالعات بسیاری برای پیش‌بینی خواص ترکیبات شیمیایی به کمک توصیف گره‌های مولکولی صورت گرفته است (Mohammadinasab, 2017; Bertinetto, 2013; Lailong, 2011; Basak, 2008) برای پی بردن به رابطه کمی ساختار-فعالیت^۱ (QSAR) در ترکیبات شیمیایی می‌توان از روش‌های محاسباتی گوناگونی استفاده نمود (Kunal, 2015; Narasimhan, 2007). مطالعه QSAR اولین بار در سال ۱۹۶۴ توسط هانش و فوجیتا مطرح گردید (Hansch, 1964). این پژوهشگران برای برقراری این ارتباط از روش‌های محاسباتی بهره جستند. یکی از مراحل مهم در این مطالعه، بعد از انتخاب تعداد مناسب از ترکیبات مورد نظر و مجموعه داده‌ها،

^۱ Quantitative Structure Activity Relationship

انتخاب و محاسبه توصیف‌گرهای مولکولی می‌باشد. این توصیف‌گرها نتیجه انجام یک سری عملیات ریاضی هستند که اطلاعات شیمیایی یک مولکول را به یک کمیت تبدیل می‌کنند. پس از انتخاب توصیف‌کننده مناسب بایستی مدل QSAR با استفاده از الگوریتم‌هایی مانند رگرسیون خطی چندگانه^۱ (MLR) و یا شبکه‌های عصبی مصنوعی^۲ (ANN) ساخته شود (Roy, 2015; Saxena, 2003; Balaban, 1999; Gakh, 1994). در صورتی که رابطه میان متغیر وابسته و متغیرهای مستقل از الگوی خطی مانند روش MLR پیروی نکند، از روش‌های غیر خطی مانند شبکه عصبی ANN بر پایه نورون‌های عصبی استفاده می‌شود (Tang, 1998). یکی از مزایای مهم ANN توانایی برقراری ارتباط غیرخطی بین متغیرها از یک دسته و افزایش ظرفیت و تنوع مدل‌های خروجی چندگانه به‌طور همزمان است. یکی از مهم‌ترین مدل‌های عصبی، مدل پرسپترون چندلایه^۳ (MLP) است که شامل همان لایه‌های موجود در روش کلی ANN است به‌طوری‌که در آن سعی شده است که خطای تخمین و میزان اشتباه‌ها در ساختار شبکه عصبی کمینه گردد (Webb, 2002; Bishop, 1997). از نکات مهم در عملکرد شبکه عصبی MLP، بهینه‌سازی تعداد لایه‌های ورودی، میانی، خروجی، نوع تابع تحریک نورون‌ها و شیوه آموزش شبکه است (Nazzal, 2008; Malik, 2005; Haykin, 1999). در این مطالعه، سعی شده است که با کمک روش‌های محاسباتی مختلف، ابتدا به استخراج برخی از توصیف‌گرهای ساختاری در سموم ارگانوفسفات پرداخته شود و سپس از طریق بررسی ارتباط شاخص سمیت LD₅₀ تجربی این ترکیبات با توصیف‌گرهای به‌کار رفته، با کمک روش‌های رگرسیون خطی چندگانه MLR و شبکه عصبی مصنوعی ANN با مدل پرسپترون چند لایه MLP، روش محاسباتی مناسبی برای پیش‌بینی LD₅₀ در آفت‌کش‌های مورد مطالعه ارائه گردد.

مواد و روش‌ها

در این مطالعه، جهت بررسی ارتباط شاخص سمیت LD₅₀ با برخی توصیف‌گرهای مولکولی در برخی آفت‌کش‌ها و پیش‌بینی این شاخص مهم در ترکیبات ارگانوفسفات در سه بخش اقدام به جمع‌آوری اطلاعات مورد نیاز گردید. در بخش اول، مقادیر تجربی LD₅₀ مربوط به ۵۶ نوع از ترکیبات ارگانوفسفات با استفاده از وب سایت ChemIDplus در جدول (۱) جمع‌آوری شدند (WWW.ChemIDplus, A Toxnet Database). (لازم به ذکر است، تمامی مقادیر گزارش شده برای LD₅₀ در این تحقیق، مربوط به شاخص سمیت تجویز خوراکی LD₅₀ می‌باشد). در بخش دوم، با استفاده از برنامه محاسباتی chemicalize علاوه بر مقادیر logP ترکیبات ارگانوفسفره، اقدام به استخراج ۹ توصیف‌گر مولکولی برای ترکیبات فوق گردید (WWW. Chemicalize). تاکنون تعداد زیادی از شاخص‌های مولکولی برای ترکیبات مختلف شیمیایی محاسبه شده‌اند. توصیف‌گرهای مورد آزمون در این مطالعه که با کمک محاسبات نرم‌افزار chemaxon به‌دست آمده‌اند عبارتند از وزن مولکولی^۴، برخی شاخص‌های هندسی مانند: حجم واندروالسی^۵، مساحت سطح واندروالسی^۶، مساحت سطح حلال در دسترس^۷، مساحت سطح قطبی ساختاری^۸، حداقل مساحت تصور شده^۹، حداکثر مساحت تصور

¹ Multiple Linear Regression

² Artificial Neural Network

³ Multiple-Layer Perceptron

⁴ Molecular Weight

⁵ Van der Waals volume

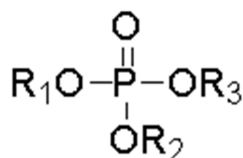
⁶ Van der Waals surface area

⁷ Surface area solvent accessible

⁸ Topological polar surface area

⁹ Minimum projection area

شده^۱، حداقل شعاع تصور شده^۲ و حداکثر شعاع تصور شده^۳. در بخش سوم، به منظور انجام محاسبات آماری از نرم‌افزار Statistica V. 12 استفاده شد و به منظور بررسی ارتباط بین شاخص سمیت LD₅₀ این ترکیبات با توصیف‌گرهای مولکولی محاسبه شده، روش‌های ANN و MLP با مدل MLR، به کار برده شدند.



شکل ۱- ساختار نمونه ارگانوفسفاته‌های به کار رفته در این مطالعه

Figure 1- The template structure of used organophosphates in present study

جدول ۲- ساختار ترکیبات ارگانوفسفاته با توجه به موقعیت گروه‌های عاملی

Table 1- The structure of organophosphate compounds according to the position of functional groups

No	Name	R ₁	R ₂	R ₃	LD ₅₀ (mgkg ⁻¹)
1	Tris(2-chloropropyl) phosphate	C ₃ H ₆ Cl	C ₃ H ₆ Cl	C ₃ H ₆ Cl	1017
2	Triethyl phosphate	C ₂ H ₅	C ₂ H ₅	C ₂ H ₅	1165
3	Tris(1-chloropropan-2-yl) phosphate	C ₃ H ₆ Cl	C ₃ H ₆ Cl	C ₃ H ₆ Cl	1500
4	Tris(1,3-dichloroisopropyl)phosphate	C ₃ H ₅ Cl ₂	C ₃ H ₅ Cl ₂	C ₃ H ₅ Cl ₂	1850
5	Coralox	C ₂ H ₅	C ₂ H ₅	C ₁₀ H ₆ ClO ₂	10
*6	Phosphoric acid, tris(2,3-dibromopropyl) ester	C ₃ H ₅ Br ₂	C ₃ H ₅ Br ₂	C ₃ H ₅ Br ₂	810
7	Phosphoric acid, diethyl p-(methylsulfonyl)Phenyl ester	C ₇ H ₇ SO ₂	C ₂ H ₅	C ₂ H ₅	5.012
8	Monocrotophos	C ₃ H ₈ NO	CH ₃	CH ₃	8
*9	Phosphoric acid, dibutyl phenyl ester	C ₆ H ₅	C ₄ H ₉	C ₄ H ₉	2140
*10	naled	C ₂ HBr ₂ Cl ₂	CH ₃	CH ₃	92
11	cis-Chlorfenvinphos	C ₂ H ₅	C ₂ H ₅	C ₈ H ₄ Cl ₃	10
12	[2-Bromo-1-(2,4-dichlorophenyl) vinyl] diethyl phosphate	C ₈ H ₄ BrCl ₂	C ₂ H ₅	C ₂ H ₅	21.4
13	heptenophos	C ₇ H ₆ Cl	CH ₃	CH ₃	96
*14	Mevinphos	CH ₃	CH ₃	C ₅ H ₇ O ₂	3
15	Phosphoric acid, tri-2-chloroethyl ester	C ₂ H ₄ Cl	C ₂ H ₄ Cl	C ₂ H ₄ Cl	1230
*16	Phosphoric acid, tri-(2-butoxyethyl) ester	C ₂ H ₄ OC ₄ H ₉	C ₂ H ₄ OC ₄ H ₉	C ₂ H ₄ OC ₄ H ₉	3000
17	Triphenylphosphate	C ₆ H ₅	C ₆ H ₅	C ₆ H ₅	3500
18	Tri-n-butylphosphate	C ₄ H ₉	C ₄ H ₉	C ₄ H ₉	3000
19	Tricresyl phosphate	C ₇ H ₇	C ₇ H ₇	C ₇ H ₇	1160
20	Chloropyrifos oxon	C ₂ H ₅	C ₂ H ₅	C ₅ HCl ₃ N	135
21	Diethyl p-nitrophenyl phosphate	C ₂ H ₅	C ₂ H ₅	C ₆ H ₅ NO ₂	1.8
22	Dicrotophos	CH ₃	C ₆ H ₁₀ NO	CH ₃	13
23	Tetrachlorvinphos	CH ₃	CH ₃	C ₈ H ₃ Cl ₄	480
*24	Phosphoric acid, tris-(2,3-dichloropropyl ester)	C ₃ H ₅ Cl ₂	C ₃ H ₅ Cl ₂	C ₃ H ₅ Cl ₂	2830
25	Chlorfenvinphos	C ₈ H ₄ Cl ₃	C ₂ H ₅	C ₂ H ₅	10
26	Phosphoric acid, trimethyl ester	CH ₃	CH ₃	CH ₃	840
27	cis-Mevinphos	C ₅ H ₇ O ₂	CH ₃	CH ₃	3
28	Diphenyl o-xenyl phosphate	C ₆ H ₅	C ₆ H ₅	C ₁₂ H ₉	15800
*29	Crotoxyphos	C ₁₂ H ₁₃ O ₂	CH ₃	CH ₃	38.4
30	Phosphoric acid, tris(2-ethylhexyl) ester	C ₈ H ₁₇	C ₈ H ₁₇	C ₈ H ₁₇	37000
31	Triisobutyl phosphate	C ₄ H ₉	C ₄ H ₉	C ₄ H ₉	5000

¹ Maximum projection area

² Minimum projection radius

³ Maximum projection radius

No	Name	R ₁	R ₂	R ₃	LD ₅₀ (mgkg ⁻¹)
32	Dichlorvos	C ₂ HCl ₂	CH ₃	CH ₃	17
33	Phosphoric acid, 2,4-dichlorophenyl diethyl ester	C ₆ H ₃ Cl ₂	C ₂ H ₅	C ₂ H ₅	100
34	Phosphamidon	CH ₃	CH ₃	C ₈ H ₁₃ CINO	8
35	Triphenyl phosphate	C ₅ H ₁₁	C ₅ H ₁₁	C ₅ H ₁₁	23400
*36	Tris(2,6-xylenyl)phosphate	C ₈ H ₉	C ₈ H ₉	C ₈ H ₉	15800
37	Tris(2,4-xylenyl)phosphate	C ₈ H ₉	C ₈ H ₉	C ₈ H ₉	15800
38	Methylparaoxon	CH ₃	C ₆ H ₅ NO ₂	CH ₃	3.27
*39	Phosphoric acid, 2,2-dibromo-1,2-dichloroethyl dimethyl	CH ₃	C ₂ HBr ₂ Cl ₂	CH ₃	49
40	Diphenyl cresyl phosphate	C ₇ H ₇	C ₆ H ₅	C ₆ H ₅	1420
41	Bomyl	C ₇ H ₉ O ₄	CH ₃	CH ₃	31
42	Phosphoric acid, dibutyl 2-hydroxyethyl ester	C ₄ H ₉	C ₄ H ₉	C ₂ H ₄ OH	915
43	Phosphoric acid, 2-chlorovinyl diethyl ester	C ₂ H ₅	C ₂ H ₅	C ₂ H ₂ Cl	7
44	Ethyl dichlorvos	C ₂ H ₅	C ₂ HCl ₂	C ₂ H ₅	2.5
45	Phosphoric acid, bis(3-chloropropyl)p-nitrophenyl ester	C ₃ H ₆ Cl	C ₆ H ₅ NO ₂	C ₃ H ₆ Cl	33
46	Phosphoric acid, bis (2-chloropropyl)p-nitrophenyl ester	C ₃ H ₆ Cl	C ₆ H ₅ NO ₂	C ₃ H ₆ Cl	173
*47	Ethanol, 2-methoxy-, 1,1',1''-phosphate	C ₂ H ₄ OCH ₃	C ₂ H ₄ OCH ₃	C ₂ H ₄ OCH ₃	17150
48	Isodecyl diphenyl phosphate	C ₆ H ₅	C ₁₀ H ₂₁	C ₆ H ₅	15800
49	Methylbromfenvinphos	C ₈ H ₄ BrCl ₂	CH ₃	CH ₃	225
50	Dimethyl p-(methylthio) phenyl phosphate	C ₇ H ₇ S	CH ₃	CH ₃	7
51	Fenthion oxon sulfone	CH ₃	CH ₃	C ₈ H ₉ SO ₂	30
52	Dimethylvinphos	C ₈ H ₄ Cl ₃	CH ₃	CH ₃	97.5
53	Fospirate	CH ₃	CH ₃	C ₃ HCl ₃ N	869
*54	Naftalofos	C ₁₂ H ₆ NO ₂	C ₂ H ₅	C ₂ H ₅	70
55	Propaphos	C ₃ H ₇	C ₃ H ₇	C ₇ H ₇ S	61
56	TEPP	C ₄ H ₁₀ O ₃ P	C ₂ H ₅	C ₂ H ₅	0.5

روش رگرسیون خطی چندگانه (MLR)

در آنالیز رگرسیون خطی چندگانه (MLR)، رابطه بین مقادیر سمیت LD₅₀ به عنوان متغیر وابسته (y_i) و شاخص‌های مولکولی به عنوان متغیرهای مستقل (x_i) به صورت کلی:

$$y_i = a_0 + a_1x_{i1} + a_2x_{i2} + \dots + \varepsilon_i \quad (1)$$

نشان داده می‌شود (Kutner, 2004; Randic, 2000) که در آن a₀ عرض از مبدا و a₁، a₂، a₃ و..... ضرایب رگرسیونی هستند. در این مدل‌سازی، از میان توصیف‌گرهای محاسبه شده، مهم‌ترین آن‌ها با استفاده از آنالیز رگرسیون خطی چندگانه و به کمک روش برگشتی^۱ مشخص گردیدند. برای اعمال روش MLR به منظور پیش‌بینی شاخص سمیت LD₅₀ در ترکیبات ارگانوفسفات از نرم‌افزار Statistica V ۱۲ استفاده شد. به همین منظور جهت انجام محاسبات، ترکیبات توسط نرم‌افزار مذکور و به شکل تصادفی، به دو دسته ۴۵ آموزش و ۱۱ تست (به نسبت تقریبی ۸۰٪ و ۲۰٪) تقسیم شدند که شماره مولکول‌های آموزش در جدول (۱) بدون علامت ستاره مشخص گردیده‌اند.

¹ backward

روش شبکه عصبی MLP

برای ایجاد ارتباط بین توصیف‌گرهای مولکولی و مقادیر LD₅₀ ترکیبات مورد مطالعه در روش شبکه عصبی، از تعدادی لایه‌های ورودی و میانی و خروجی استفاده شد و مشخص گردید که وقتی تعداد لایه‌های ورودی در مرحله آموزش برابر ۱۰ باشد، تعداد نورون‌های لایه میانی شبکه در مرحله آزمون به ۶ تغییر می‌یابد و نتایج مطلوب، زمانی حاصل می‌شود که تعداد نورون‌های لایه خارجی به ۱ کاهش یابد. این موضوع ضرورت بهینه‌سازی کلیه پارامترهای شبکه و مواجهه عالمانه با آن‌ها را بیش از پیش آشکار می‌سازد. به‌منظور دستیابی به میزان توانایی مدل MLP در شبکه عصبی، مقادیر سمیت ترکیبات ارگانوفسففات به‌عنوان داده‌های ورودی، توسط نرم‌افزار Statistica V. ۱۲ به‌صورت تصادفی، به ۳ دسته شامل ۷۰٪ آموزش و ۱۵٪ تست و ۱۵٪ اعتبارسنجی تقسیم‌بندی شدند و برای هر کدام ضرایب تعیین و همبستگی محاسبه گردید.

نتایج

رگرسیون خطی چندگانه: پارامترهای اعتبارسنجی مربوط به آنالیز رگرسیون خطی چندگانه شامل ضریب همبستگی^۱ (R)، ضریب تعیین^۲ (R^۲)، ضریب تعیین تعدیل یافته^۳ (R^۲_{Adjust})، آماره فیشرف، سطح معناداری^۴ (Sig)، خطای مربعات میانگین^۵ (MSE) و جذر خطای میانگین مربعات^۶ (RMSE) برای ۴۵ ترکیب آموزش، در جدول (۲) آورده شده است.

جدول ۲- نتایج مربوط به آنالیز رگرسیون خطی چندگانه برای پیش‌بینی LD₅₀

Model	R	R ²	R ² _{Adjust}	F	MSE	RMSE	Sig
1	0.783	0.612	0.537	8.078	24635362.08	4963.40	0.000
2	0.783	0.612	0.546	9.285	24111234.75	4910.32	0.000
3	0.782	0.612	0.556	10.823	23628168.78	4860.88	0.000
4	0.780	0.609	0.561	12.726	23326388.98	4829.74	0.000
5	0.778	0.605	0.565	15.312	23105014.17	4806.76	0.000
6	0.774	0.600	0.568	19.109	22945910.94	4790.18	0.000
7	0.768	0.590	0.566	24.945	23055099.11	4801.57	0.000

شبکه عصبی مصنوعی: در روش شبکه عصبی مصنوعی از طریق مدل پرسپترون چند لایه، مقادیر ضریب همبستگی R=0.9935 و ضریب تعیین R²=0.9870 و جذر خطای مجذور میانگین RMSE=68.658 برای مقادیر سمیت LD₅₀ مربوط به ۴۰ نوع ترکیب، به‌عنوان سری آموزش و نیز پارامترهای R=0.9898 و R²=0.9797 و RMSE=200.379 برای مقادیر سمیت LD₅₀ مربوط به ۸ ترکیب، به‌عنوان سری تست و همچنین مقادیر R=0.8837 و R²=0.7809 و RMSE=1.668 برای مقادیر سمیت LD₅₀ مربوط به ۸ ترکیب، از سری اعتبارسنجی داده‌ها به‌دست آمد. مقادیر تجربی،

¹ Correlation coefficient² Coefficient of determination³ Adjusted R-squared⁴ Significance level⁵ Mean Squared Error⁶ Root Mean Squared Error

پیش‌بینی و باقی‌مانده شاخص سمیت $LD_{50}(mgkg^{-1})$ برای دسته آموزش در جدول (۳) و برای سری‌های آزمون و اعتبارسنجی در جدول (۴) آورده شده است.

جدول ۳- مقادیر تجربی، پیش‌بینی و باقی‌مانده $LD_{50}(mgkg^{-1})$ برای سری ترکیبات آموزش در روش شبکه عصبی مصنوعی

Table 3- The values of experimental, predicted and residual LD_{50} of training set in ANN method

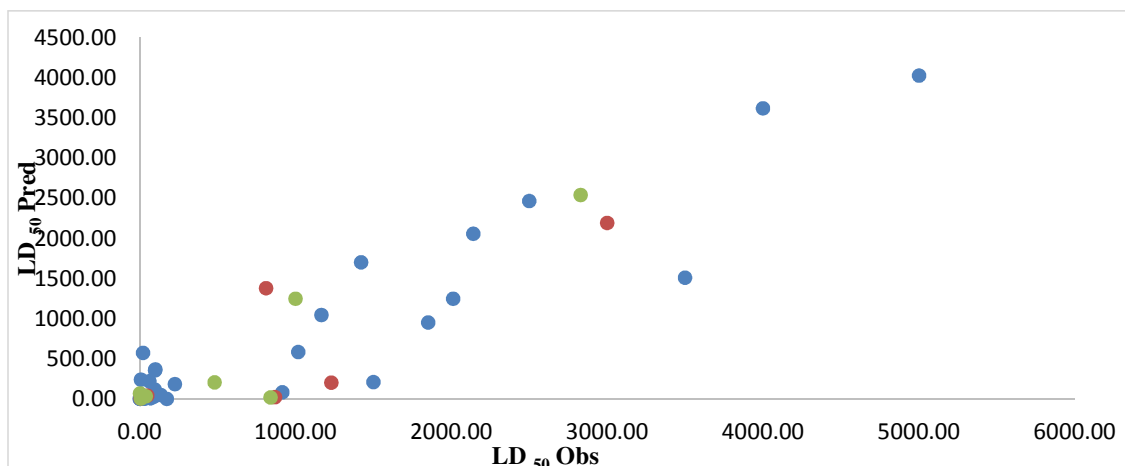
Compound	$LD_{50\text{ exp}}$	$LD_{50\text{ pred}}$	Residual
Tris(2-chloropropyl) phosphate	1017.00	582.77	434.23
Triethyl phosphate	1165.00	45.51	1119.49
Tris(1-chloropropan-2-yl) phosphate	1500.00	211.77	1288.23
Tris(1,3-dichloroisopropyl)phosphate	1850.00	951.43	898.57
Coralox	10.00	49.29	-39.29
Phosphoric acid, dibutyl phenyl ester	2140.00	2055.85	84.15
naled	92.00	29.31	62.69
[2-Bromo-1-(2,4-dichlorophenyl)vinyl] diethyl phosphate	21.40	573.05	-551.65
heptenophos	96.00	117.79	-21.79
cis-Mevinphos	3.00	5.34	-2.34
Triphenylphosphate	3500.00	1510.02	1989.98
Tri-n-butylphosphate	3000.00	4465.49	-1465.49
Tricresyl phosphate	1160.00	3619.25	-2459.25
Chloropyrifos oxon	135.00	49.15	85.85
Diethyl p-nitrophenyl phosphate	1.80	1.83	-0.03
Dicrotophos	13.00	9.83	3.17
Chlorfenvinphos	10.00	1248.82	-1238.82
cis-Mevinphos	3.00	5.34	-2.34
Triisobutyl phosphate	5000.00	1025.51	3974.49
Phosphoric acid, 2,4-dichlorophenyl diethyl ester	100.00	369.02	-269.02
Phosphamidon	8.00	20.31	-12.31
Tripentyl phosphate	23400.00	23464.69	-64.69
Phosphoric acid, tris(2-ethylhexyl) ester	37000.00	37029.40	-29.40
Tris(2,6-xylenyl)phosphate	15800.00	15868.56	-68.56
Tris(2,4-xylenyl)phosphate	15800.00	15377.77	422.23
Methylparaoxon	3.27	0.97	2.30
Diphenyl cresyl phosphate	1420.00	1699.15	-279.15
Bomyl	31.00	2.59	28.41
Phosphoric acid, dibutyl 2-hydroxyethyl	915.00	81.59	833.41
Phosphoric acid, bis (3-chloropropyl) p-nitrophenyl ester	33.00	23.41	9.59
Phosphoric acid, bis (2-chloropropyl) p-nitrophenyl ester	173.00	2.05	170.95
Ethanol, 2-methoxy-, 1,1',1''-phosphate	17150	17135.54	14.46
Isodecyl diphenyl phosphate	15800.00	15610.86	189.14
Methylbromfenvinphos	225.00	183.77	41.23
Dimethyl p-(methylthio) phenyl phosphate	7.00	238.12	-231.12
Fenthion oxon sulfone	30.00	3.42	26.58
Dimethylvinphos	97.50	355.58	-258.08
Naftalofos	70.00	5.98	64.02
Propaphos	61.00	220.48	-159.48
TEPP	0.50	4.74	-4.24

جدول ۴- مقادیر تجربی، پیش‌بینی و باقی‌مانده LD₅₀(mgkg⁻¹) برای سری‌های آزمون و اعتبارسنجی در روش شبکه عصبی مصنوعی

Table 4- The values of experimental, predicted and residual LD₅₀ of test and validation sets in ANN method

Compounds test	LD ₅₀ Obs (test)	LD ₅₀ pred (test)	LD ₅₀ Res (test)	Compounds test	LD ₅₀ Obs (valid)	LD ₅₀ pred (valid)	LD ₅₀ Res (valid)
Phosphoric acid, tris (2,3-dibromopropyl) ester	810.00	1376.76	-566.758	Monocrotophos	8.000	3.282	4.72
Phosphoric acid, diethyl p-(methylsulfonyl)	5.01	8.07	-3.057	cis-Chlorfenvinphos	10.000	1248.815	-1238.82
Phosphoric acid, tri-2-chloroethyl ester	1230.00	201.36	1028.639	Tetrachlorvinphos	480.000	204.961	275.04
Tri-n-butylphosphate	3000.00	1192.00	1807.998	Phosphoric acid, tris-(2,3-dichloropropyl ester)	2830.000	3939.046	-1109.05
Diphenyl oxenyl phosphate	15800.00	16479.83	-679.825	Phosphoric acid, trimethyl ester	840.000	17.870	822.13
Phosphoric acid, 2,2-dibromo-1, 2-dichloroethyl dimethyl	49.00	42.99	6.012	Crotoxyphos	38.400	33.968	4.43
Phosphoric acid, 2-chlorovinyl diethyl ester	7.00	48.22	-41.221	Dichlorvos	17.000	42.967	-25.97
Fospirate	869.00	21.28	847.717	Ethyl dichlorvos	2.500	68.713	-66.213

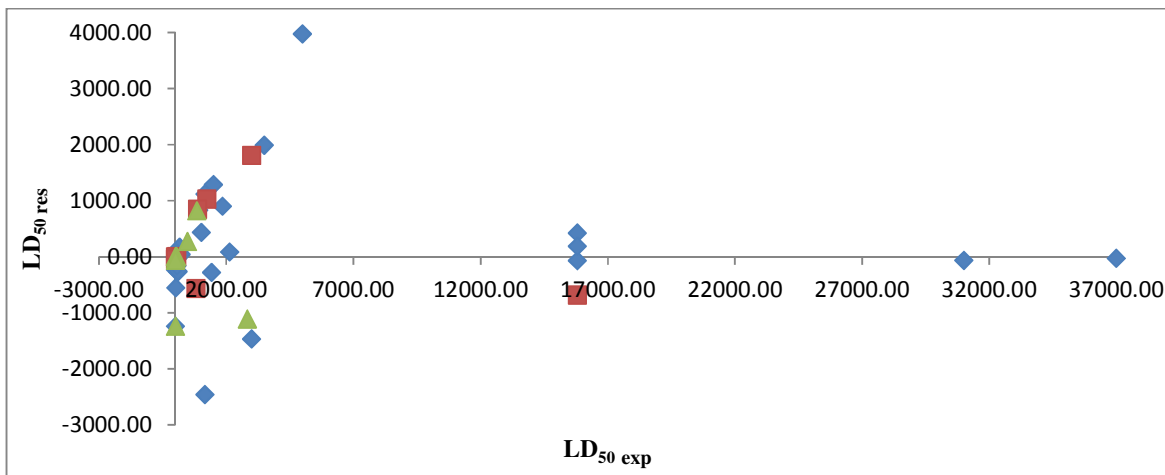
در شکل (۱) همبستگی میان مقادیر LD₅₀ تجربی با مقادیر پیش‌بینی شده در آنالیز ANN با مدل MLP برای هر سه سری داده‌ها، شامل آموزش و آزمایش و اعتبارسنجی نشان داده شده است.



شکل ۱- نمودار مقادیر پیش‌بینی برحسب تجربی LD₅₀ برای سری‌های آموزش، تست و اعتبارسنجی در روش شبکه عصبی مصنوعی

Fig. 1- The curve of the predicted versus experimental LD₅₀ values for training, test and validation sets on ANN method

نمودار تغییرات مقادیر باقی‌مانده برحسب مقادیر تجربی LD_{50} برای تعیین نرمال بودن مدل مورد استفاده قرار گرفت. در شکل (۲) مقادیر باقی‌مانده در اطراف محور x در یک الگوی تصادفی در آنالیز ANN با مدل MLP را نشان می‌دهد.



شکل ۲- نمودار مقادیر باقیمانده LD_{50} برحسب تجربی در روش شبکه عصبی مصنوعی

Fig. 2- The curve of the residual versus experimental LD_{50} values for training, test and validation sets on ANN method

تجزیه و تحلیل به روش رگرسیون خطی چندگانه (MLR): ضرایب تعیین کننده مهم در روش MLR مطابق با جدول (۲)، حاکی از آن است که نتایج مطلوب برای پیش‌بینی شاخص LD_{50} ترکیبات مورد مطالعه حاصل نشده است، به‌طوری‌که بر اساس بهترین الگوی به‌دست آمده در جدول (۲) (مدل (۷)) با بالاترین آماره فیشر و کمترین مقدار جذر خطای مجذور میانگین و مقدار ضریب تعیین در آن، فقط می‌توان ۵۹٪ از داده‌ها برای سمیت ترکیبات ارگانوفسفات را پیش‌بینی نمود، لذا جهت یافتن الگوی مناسب‌تر، مدل‌سازی از طریق روش غیرخطی ANN برای پیش‌بینی شاخص سمیت LD_{50} انجام گردید.

تجزیه و تحلیل به روش شبکه عصبی مصنوعی (ANN): در روش شبکه عصبی با مدل MLP، با در نظر گرفتن تعداد ۱۰ لایه ورودی، ۶ لایه پنهان و ۱ لایه خروجی در آزمون، مشخص گردید، پارامترهای مناسبی برای پیش‌بینی مقادیر سمیت ترکیبات مورد مطالعه، به‌دست آمده است، به‌طوری‌که با توجه به مقدار ضریب تعیین در رگرسیون مقادیر تجربی بر اساس پیش‌بینی، می‌توان حدوداً ۹۸٪ از داده‌های آموزش و یا تست را برای شاخص سمیت LD_{50} پیش‌بینی نمود. شکل (۱) همبستگی بالا میان مقادیر LD_{50} تجربی با پیش‌بینی شده در روش شبکه عصبی را برای هر سه سری داده‌ها، شامل آموزش و آزمایش و اعتبارسنجی تایید می‌نماید شکل (۲) نشان می‌دهند، خطاها در اطراف محور x از یک توزیع یکنواختی برخوردارند که این خود می‌تواند توانایی روش انتخاب شده برای پیش‌بینی مقادیر سمیت LD_{50} سموم ارگانو فسفات را به‌خوبی نشان دهد. نتایج این مطالعه نشان می‌دهند که در مدل‌سازی به روش QSAR، توصیف گرهایی مانند ضریب تقسیم ماده در فاز آلی به آبی، وزن مولکولی و نیز توصیف گرهایی هندسی ترکیبات به‌کار رفته که به‌راحتی با محاسبات نرم افزاری استخراج می‌شوند و برای اولین بار به‌طور همزمان به‌کار گرفته شده‌اند، نقش بسیار مهمی در یافتن روش مناسب برای پیش‌بینی مقادیر سمیت LD_{50} این دسته از ترکیبات شیمیایی ایفا نموده‌اند. در مطالعه انجام شده، سعی بر آن شده است که با کمک هر دو روش محاسباتی خطی و غیرخطی، به ارزیابی نتایج به‌دست آمده پرداخته شود و

بهترین الگو برای پیش‌بینی مقادیر سمیت ارگانوفسفاته‌های جدید ارائه گردد. نتایج این مطالعه نشان داد، الگوی پرسپترون چند لایه MLP در روش شبکه عصبی مصنوعی، که برای اولین بار به کار گرفته شده است، برای تخمین و تقریب سمیت LD₅₀ ترکیبات ارگانوفسففات، از کارایی بسیار بالایی برخوردار می‌باشد.

نتیجه گیری

با توجه به گستردگی استفاده از سموم ارگانو فسفات، احتمال آلودگی و مسمومیت با این ترکیبات خطری بالقوه برای سلامت انسانها به شمار می‌رود. مطالعات نشان می‌دهد که مرگ و میر ناشی از این مسمومیت چشمگیر است و بیشتر مطالعات نیز عوارضی شامل عوارض گوارشی، قلبی، تنفسی و عصبی را تایید کرده‌اند. از آنجا که تعیین میزان شاخص سمیت LD₅₀ در آفت‌کش‌های آلی به‌خصوص ترکیبات ارگانوفسففات به روش‌های تجربی در آزمایشگاه با آلودگی و نیز مسمومیت بالا بر بافت‌های بدن انسان همراه است، لذا در این مطالعه سعی شده است که به پیش‌بینی مقدار دقیق این پارامتر مهم بیولوژیکی از طریق روش‌های محاسباتی و بدون خطرات زیست محیطی و ضرر و زیان‌های حاصل از انجام آزمایش‌های تجربی، پرداخته شود. مدل‌سازی توسط نرم‌افزار Statistica، برتری روش غیرخطی را نسبت به روش خطی مشخص نمود. پارامترهای مهم اعتبارسنجی در این تحقیق نشان داد که توصیف‌گرهای مولکولی مورد استفاده در مدل‌سازی به روش شبکه عصبی با مدل پرسپترون چندلایه در تعیین شاخص سمیت LD₅₀ مربوط به آفت‌کش‌های ارگانوفسففات، نقش بسیار مهمی را ایفا می‌کنند. مدل‌سازی توسط آنالیز QSAR نشان داد که روش غیرخطی با بالاترین ضریب تعیین و کمترین خطا نسبت به روش خطی در پیش‌بینی شاخص سمیت LD₅₀ از برتری بسیار بالایی برخوردار می‌باشد.

Reference

- Aghilinejad, M., Mohamadi, S. and Farshad, A. 2007.** The effect of pesticides on farmers' health. *Research in Medicine*, 31(4): 327-31.
- Bakand, S., Dehghani, Y., Gohari, M., Mosadegh, M. and Mirmohammadi, S. 2012.** Exposure assessment of greenhouse workers with anticholinesterase pesticides by biological monitoring. *Iran Occup Health*, 9(3): 1-10.
- Balaban, A. T., Mills, D. and Basak, S. C. 1999.** Correlation between Structure and Normal Boiling Points of Acyclic Carbonyl Compounds. *Journal of chemical information and computer sciences*, 39(4): 758-764.
- Barkhordri, A., Raee Bandpey, T., Hekmati Moghaddam, S., Mosadegh, M., Fallahzadeh, H. 2012.** Evaluation of cholinesterase level in workers exposed to organophosphates. *Occupational Medicine Quarterly Journal*, 4(1-2): 9-16.
- Basak, S. C., Mills, D., Natarajan, R. and Gute, BD. 2008.** Predicting chemical reactivity and bioactivity from structure: A mathematical-cum-computational approach, in *Theory of Chemical Reactivity*, PK Chattaraj, Ed. Taylor & Francis. 304pp.
- Bertinetto, C., Duce, C., Solaro, R., Solaro, R. and Rosaria Tiné, M. 2013.** Modeling of the Acute Toxicity of Benzene Derivatives by Complementary QSAR Methods. *Communications in Mathematical and in Computer Chemistry*, 70: 1005-1021 .
- Bishop, C. M. 1997.** *Neural Networks for Pattern Recognition*. Oxford University Press. 480pp.
- De Silva, H. J., Samarawickrema, N. A. and Wickremasinghe, A. R. 2006.** Toxicity due to organophosphorus compounds: what about chronic exposure? *Trans R Soc Trop Med Hyg*, 100(9): 803-6.
- Ebrahimzadeh, M. A., Shokrzadeh, M. and Bioukabadi, M. 2005.** Effect of organophosphorous pesticides on acetyl cholinesterase activity in agricultural workers. *J Shahrekord Univ Med Sci*; 7(1): 1-7.
- Eddleston, M., Phillips, M. R. 2004.** Self poisoning with pesticides. *BMJ*, 328(7430): 42-4.
- Gakh, A. A., Gakh, E. G., Sumpter, B. G. and Noid, D. W. 1994.** Neural network-graph theory approach to the prediction of the physical properties of organic compounds. *Journal of chemical information and computer sciences*, 34: 832-839.
- Hansch, C. and Fujita, T. 1964.** ρ - σ - π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *Journal of the American Chemical Society*, 86(8): 1616-1626
- Kunal, R., Supratik, K. and Rudra Narayan, D. 2015.** A Primer on QSAR/QSPR Modeling. *Journal of Theoretical and Computational Chemistry*, 2: 37.121pp.
- Kutner, M. H., Nachtsheim, C. J. and Neter, J. 2004.** *Applied linear regression models*, McGraw-Hill/Irwin, Boston: Mass. 2714pp.
- Lailong, M. and Hongmei, H. 2011.** Quantitative Structure-Property Relations (QSPRs) for Predicting the Standard Absolute Entropy ($S_{298\text{ K}}$) of Gaseous Organic Compounds. *Journal of Industrial & Engineering Chemistry* 50 (14): 8764-8772.
- Lopes Soares, W. and Firpo de Souza Porto, M. 2009.** Estimating the social cost of pesticide use: An assessment from acute poisoning in Brazil. *Ecological Economics*, 68(10): 2721.
- Haykin, S., 1999.** *Neural networks and learning machines*. 3rd ed. New York Boston San Francisco. Prentice Hall. 122pp.
- Maroni, M., Colosio, C., Ferioli, A. and Fait, A. 2000.** Biological monitoring of pesticide exposure: A review. *Introduction. Toxicology*; 143(1): 1-118.
- Hagan, M. T., Demuth, H. B. and Jesus, O. D. 2002.** An introduction to the use of neural networks in control systems, *international journal of robust and nonlinear control*; 12: 959-985.
- Mahmoudi, G. and Asaee, R. 2008.** Epidemiologic study of organophosphate and organochlorate pesticides poisoning in hospitalized patients in khorramabad shohada ashayer hospital from march to august 2006; 10 (1):3-10.
- Malik, N. 2005.** *Artificial Neural Networks and their Applications*. National conference on Unear thing Technological Developments & their transfer for serving Masses, GLA ITM, Mathura, Mathura, India.

- Moghaddamnia, A. A. and Hallagisani, S. 1999.** Study of acute poisoning in Ramsar. The Journal of Qazvin University of Medical Sciences, 3(2):43-9.
- Mohammadinasab, E. 2017.** Modeling boiling points of alkane derivatives. Studia Universitatis Babeş-Bolyai Seria Chemia, 1: 143-152.
- Narasimhan, B., Judge, V., Narang, R., Ohlan, R. and Ohlan, S. 2007.** Quantitative structure activity relationship studies for prediction of antimicrobial activity of synthesized 2,4-hexadienoic acid derivatives, Bioorganic & Medicinal Chemistry Letters, 17: 5836-5845.
- Nazzal, J. M., El-Emary, I. M. and Najim, S. A. 2008.** Multi layer Perceptron Neural Network (MLPs) For Analyzing the Properties of Jordan Oil Shale. World Applied Sciences Journal 5 (5): 546-552.
- <https://chem.sis.nlm.nih.gov/chemidplus/> NIH, Bethesda, MD, United States.
- Pashm Forush, M., Shafiei, F. and Dialamehpour, F. 2016.** QSPR Study On Benzene Derivatives To Some Physico Chemical Properties By Using Topological Indices. Iranian Journal Of Mathematical Chemistry 7(1). 93-110.
- Randic, M., Basak, S. C. 1999.** Multiple regression analysis with optimal molecular descriptors. SAR and QSAR in Environmental Research, 11: 1-23.
- Ranjbar, A., Solhi, H., Mashayekhi, F. J., Susanabdi, A., Rezaie, A. and Abdollahi, M. 2005.** Oxidative stress in acute human poisoning with organophosphorus insecticides; a case control study. Environmental Toxicology and Pharmacology, 20(1): 88-91.
- Rodgers, G. C. 1997.** Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning. JAMA, 8: 278(14): 1201.
- Saxena, A. K., Prathipati, P. 2003.** Comparison of MLR, PLS and GA-MLR in QSAR analysis. SAR and QSAR in Environmental Research, 14: 433-445.
- Sattler, C., Kachele, H. and Verch, G. 2007.** Assessing the intensity of pesticide use in agriculture. Agric Ecosyst Environ; 119(3): 299-304.
- Tafari, J., Roberts, J. 1987.** Organophosphate poisoning. Annals of Emergency Medicine, 16(2):193-202.
- Tang, Y., Chen, K. X., Jiang, H. L. and Ji, R. Y. 1998.** QSAR/QSTR of fluoroquinolones: an example of simultaneous analysis of multiple biological activities using neural network method. European Journal of Medicinal Chemistry, 33(7-8): 647-658.
- Web search engine developed by ChemAxon;** software available at "http:// WWW. Chemicalize.org".
- Webb, A. R. 2002.** Statistical Pattern Recognition, Second Edition, John Wiley & Sons Ltd.

The study of relationship between molecular descriptors and LD₅₀ of organophosphate pesticides

*M. Kianpour*¹, *E. Mohammadinasab*^{1*}

1- Respectively P.Hd student and Assistant Professor, Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran

Abstract

The organophosphate pesticides are considered as dangerous chemical pesticides for human health. The possibility of absorbing these toxins through the skin is accessible to many researchers who have been studying the toxicity of organophosphate compounds. Experimental methods are time-consuming and high cost, and they come with a lot of dangers. Quantitative structure activity/property relationship studies provide the capability to access data, information and physico-chemical properties of chemical compounds, using the methods and modeling. In this study, the multiple regression linear method and the artificial neural network with multi-layer perceptron (MLP) model were used to investigate the quantitative relationship of LD₅₀ (mgkg⁻¹) toxicity index with some molecular descriptors of some organophosphate compounds. Investigation of correlation coefficients and root mean square errors values of final models in this study showed that ANN method using the MLP model was higher than the MLR method for prediction of LD₅₀(mgkg⁻¹) of organophosphates compounds.

Key words: Organophosphate compounds, LD₅₀, QSAR, Artificial Neural Network, Multiple Linear Regression Method.

* Corresponding Author, E-mail: e-mohammadinasab@iau-arak.ac.ir

Received: 17 Nov. 2018– Accepted: 9 Apr. 2019

