

پیش‌بینی Log P (ضریب توزیع آب-اکتانل) آفت‌کش‌ها به کمک روش رگرسیون خطی چندگانه

عصمت محمدی‌نسب

استادیار، دانشگاه آزاد اسلامی، واحد اراک، گروه شیمی، اراک

چکیده

آفت‌کش‌ها از مهم‌ترین آلاینده‌های محیط می‌باشند و انتقال آن‌ها از محیط زیست به داخل بدن موجود زنده و رسیدن به سطحی بالاتر از حد طبیعی آن یکی از مهم‌ترین نگرانی‌های محیط زیستی می‌باشد. آفت‌کش‌ها به علت کاربردهای مختلف در زمینه بیوشیمی و محیط زیست و کشاورزی مورد توجه بسیاری قرار گرفته‌اند. برای شناخت آفت‌کش‌ها باید خصوصیات فیزیکی و شیمیایی آن‌ها به‌خصوص میزان سمیت آن‌ها را مورد توجه قرار داد. این مطالعه با هدف بررسی ارتباط Log P (ضریب توزیع آب - اکتانل) برخی از آفت‌کش‌ها با برخی توصیف‌گرهای توپولوژیکی به کمک نظریه گراف و روش رگرسیون خطی چندگانه انجام گرفته است. مطالعه کمی ساختار-خاصیت جهت بررسی این ارتباط به‌کار گرفته شد و نتایج نشان داد که از میان توصیف‌گرهای مورد مطالعه برای پیشگویی Log P آفت‌کش‌های مورد نظر ترکیبی از چهار توصیف‌گر توپولوژیکی: پلت (P)، هاراری (H)، راندیک (X) و سگد (Sz) مناسب است. بهترین مدل در این پژوهش نشان داد که توصیف‌گرهای ساختاری به‌کار رفته نقش مهمی را در تعیین ضریب توزیع آب - اکتانل آفت‌کش‌ها ایفا می‌نمایند. این تحقیق برای اولین بار رابطه میان ضریب پخش فاز آبی- آلی را با برخی شاخص‌های ساختاری به کمک نرم‌افزار SPSS و روش رگرسیون خطی چندگانه در آفت‌کش‌ها ارائه نموده است.

واژه‌های کلیدی: آفت‌کش‌ها، ضریب توزیع آب-اکتانل، رگرسیون خطی چندگانه

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

تاریخ دریافت مقاله (۹۵/۳/۵) - تاریخ پذیرش مقاله (۹۵/۷/۱۰)



مقدمه

قدمت مبارزه شیمیایی با آفات و بیماری‌ها به بیش از ۱۰۰۰ سال قبل از میلاد که هومر مورخ یونانی مطلبی در رابطه با کنه‌کش‌های گوگردی نوشت می‌رسد. همچنین در قرن ۱۶ میلادی مخلوطی از ترکیبات آرسنیک با عسل برای از بین بردن مورچه‌ها ابداع گشت. برای از بین بردن آفت گیلاس ترکیباتی شامل نیکوتین در قرن ۱۷ میلادی به‌کار برده شد. د.د.ت به‌عنوان اولین ترکیب شیمیایی در سال ۱۸۷۴ ساخته شد ولی در سال ۱۹۳۹ کاربرد آن برای از بین بردن حشرات معلوم گردید. سمیت آفت‌کش‌ها از طریق پوستی، تنفسی و دهانی است و نیز می‌تواند با ذخیره در بافت‌های بدن مسمومیت مزمن ایجاد کنند. معمولاً میزان پایداری سمیت آفت‌کش‌ها در خاک از یک ماه تا بیش از دو سال می‌باشد. به‌طور کلی یک آفت‌کش خوب ضمن عدم تاثیر نامناسب بر محصول سم‌پاشی شده باید برای انسان و دام بی‌خطر یا کم‌خطر باشد. و در اکوسیستم‌های کشاورزی و غیرکشاورزی و نیز در زنجیره غذایی ایجاد اختلال نکند (Gilden et al., 2010; Chiou et al., 1977; 1982).

آفت‌کش‌ها وقتی وارد محیط زیست می‌شوند، در اثر تجزیه شیمیایی، میکروبی، فتولیز، تبخیر، جذب توسط موجودات زنده و یا گیاهان و ... به تدریج از بین می‌روند. مسئله تجمع زیستی یا جمع شدن یک ماده شیمیایی در بدن موجود زنده و انتقال آفت‌کش از محیط زیست به داخل بدن موجود زنده و رسیدن به سطح بالاتر از حد طبیعی آن همیشه مورد توجه پژوهشگران محیط زیست بوده است. برای بیان میزان سمیت یک ماده معیارهایی در نظر گرفته می‌شود که این معیارها میزان خطرناک بودن یک آفت‌کش برای انسان، موجودات زنده و محیط زیست را بیان می‌کند. این معیارها نشان می‌دهد که هر ماده‌ای می‌تواند بسته به مقدار دریافت سمی باشد. بدیهی است سم علاوه بر اثرات فوری می‌تواند اثرات مزمن نیز داشته باشد یعنی در اثر ورود تدریجی مقداری کمی از سم به بدن، ایجاد بیماری یا نارسایی در ارگان‌های مختلف نماید. همچنین هر آفت‌کشی به تدریج تحت تاثیر عوامل زنده و غیرزنده به مواد دیگری تبدیل می‌شود. این مواد ثانوی را متابولیت می‌نامند. تاکسیکولوژی این مواد ضرورت دارد (Braumann, 1986; Biagi et al., 1991; Lailong, Hongmei 2011).

یکی از خصوصیت‌های مورد توجه آفت‌کش‌ها ضریب تعادل مولکول‌ها در بین اکتانول و آب است. این ضریب عددی است که نشان دهنده چگونگی توزیع آفت‌کش بین آب و اکتانول در حالت تعادل می‌باشد. این عدد یک پارامتر مهم بوده و اغلب در بررسی‌های مربوط به سرانجام و جابجایی مواد شیمیایی آلی مورد استفاده قرار می‌گیرد. گاهی آن را ضریب تجمع زیستی نیز می‌نامند و پارامتری برای تعیین توان جمع شونده یک ماده در بدن است و به‌صورت لگاریتم پی ($Log P$) نشان داده می‌شود. ضریب توزیع، نسبت غلظت یک ترکیب غیر یونیزه را در دو فاز غیرقابل امتزاج (مانند اکتانول و آب) بیان می‌دارد (Braumann, 1986).

$[X]_{Oct}$: غلظت جزء غیر یونیزه حل شده در فاز اکتانول

$[X]_{H_2O}$: غلظت جزء غیر یونیزه حل شده در فاز آب

$$\log p_{oct/wat} = \log \frac{[x]_{Oct}}{[x]_{H_2O}} \quad (1)$$

با توجه به این‌که ساختار آفت‌کش‌ها ترکیبات شیمیایی هستند و از طرق گوناگون قادر به عبور از سیستم‌های قطبی نظیر پلاسمای خون و نیز بافت‌های چربی مثل چربی کبد و مغز و غیره می‌باشند و سمیت این مولکول‌ها قابل توجه بوده و تاثیر مخرب آن‌ها بر بدن آشکار می‌باشد. تعیین مقدار $Log P$ و یا $Log Kow$ در آفت‌کش‌ها از روش‌های

آزمایشگاهی مانند: shake-flask و RP-HPLC فاز معکوس و... با خطرات و مشکلات زیادی همراه است؛ لذا در این پژوهش سعی شده است که با کمک روش‌های محاسباتی با توجه به نظریه گراف (Lailong & Hongmei, 2011; Shafiei, 2015; Basak, 2008; Bertineto, 2013; Mohammadinasab & Goodarzi, 2011). توصیف‌گرهای ساختاری در این ترکیبات پرداخته شود و سپس از طریق بررسی ارتباط ضرایب توزیع آب - اکتانل تجربی این ترکیبات با توصیف‌گرهای به‌کار رفته، یک روش محاسباتی مناسب برای تعیین $\log P$ در آفت‌کش‌های مورد مطالعه ارایه گردد. نظریه گراف یکی از زیرشاخه‌های علم ریاضیات گسسته است که علاوه بر علوم جامعه‌شناسی و روان‌شناسی و علوم کامپیوتر و مهندسی و... در فیزیک و شیمی و زیست‌شناسی نیز کاربرد فراوان دارد و امروزه زمینه‌های تحقیقاتی پیشرفته‌ای در عرصه گرایش‌های مختلف شیمی و بیولوژی را به خود اختصاص داده است (Gutman, 1994; Gakh, et al., 1994). گراف‌هایی که در شیمی به‌کار می‌روند، به‌طور کلی، به یکی از دو دسته بزرگ زیر تعلق دارند: گراف‌های ساختاری که متناظر با ساختار شیمیایی معینی است، گراف‌های واکنشی که به مجموعه‌ای از واکنش‌های شیمیایی مربوطه اطلاق می‌شود. (Gakh et al., 1994).

نظریه گراف‌ها روش‌های متفاوت فراوانی برای مشخص کردن عددی ساختارهای شیمیایی به‌دست می‌دهد. شیمی دان‌ها، هنگامی که می‌خواهند درباره مواد شیمیایی و ساختار آن‌ها اطلاعاتی منتقل کنند، از نام‌ها و فرمول‌های گوناگون استفاده می‌کنند. اما در اغلب موارد، نام‌ها و فرمول‌ها معنی مستقیم، فوری، یا صریح ریاضی ندارند. از سوی دیگر، نظریه گراف‌ها روش‌های متفاوت فراوانی برای مشخص کردن عددی ساختارهای شیمیایی به‌دست می‌دهد که در اینجا به معرفی شاخص‌های ساختاری مورد مطالعه در این پژوهش پرداخته شده است.

شاخص‌های توپولوژی

شاخص وینر

شاخص وینر از کالج برکلین آمریکا ارایه شد که در سال‌های اخیر برای تبدیل ساختار توپولوژیکی یک مولکول به یک عدد مشخص، به‌عنوان پایه پیشگویی رفتار شیمیایی مورد استفاده قرار می‌گیرد. شاخص وینر بر مفهوم فاصله در نظریه گراف استوار است. فاصله میان هر دو راس در یک گراف برابر است با تعداد لبه‌هایی که برای طی کوتاه‌ترین مسیر میان آن دو راس بایستی در گراف مولکول طی شود.

فرمول محاسبه شاخص وینر:

$$W = \frac{1}{2} \sum_{i,j} D_{ij} \quad (2)$$

عدد وینر، با نصف مجموع عناصر ماتریس فاصله برابر است، علت اینکه با نصف آن برابر است، این است که این ماتریس شامل فاصله اتم i تا اتم j و همچنین اتم j تا اتم i است (Wiener, ۱۹۴۷)

شاخص بالابان

شاخص بالابان ($J=J(G)$)، توسط بالابان در سال ۱۹۸۲ به‌صورت زیر تعریف گردید (Balaban, 1982):

$$J = \frac{M}{\mu + 1} \sum [(D_i)(D_j)]^{0.5} \quad (3)$$

M تعداد پیوندها و μ تعداد حلقه‌ها در G بوده و D_i مجموع عناصر ماتریس فاصله و i برابر است با
 $i = 1, 2, 3, \dots, N$

شاخص سگد

شاخص مشابه شاخص وینر، شاخص سگد است که توسط گوتمن و خادیکار در سال ۱۹۹۵ به صورت فرمول زیر ارائه شده است (Klavzar et al., 1996; Khadikar, 1995).

$$(4) \quad Sz_v(G) = \sum_e n_u(e|G) \cdot n_v(e|G)$$

که در آن $n_v(e|G)$ عبارت است از تعداد رئوس گراف G که فاصله آن‌ها تا راس v از فاصله‌شان تا راس u بیشتر است.

برتری اصلی شاخص سگد اصلاح شاخص وینر برای گراف‌های حلقوی است و شاخص سگد منطبق بر شاخص وینر است.

شاخص هایپر وینر

شاخص هایپر وینر WW گراف G به صورت (Klein et al., 1995; Cash et al., 2002):

$$(5) \quad WW(G) = \frac{1}{2} \sum d(u, v) + \frac{1}{2} \sum d(u, v)^2$$

تعریف می‌شود که در آن $d(u, v)$ نشان‌دهنده فاصله بین رئوس u و v در گراف G است و مجموع‌ها بر روی همه جفت‌های (نامرتب) از رئوس G اجرا می‌شوند.

شاخص راندیک

راندیک در سال ۱۹۷۵، شاخصی توپولوژیک پیشنهاد کرد که در مطالعات QSAR و QSPR به صورت یکی از پر مصرف‌ترین شاخص‌ها درآمد. شاخص شاخه‌زنی راندیک، R، به صورت زیر تعریف می‌شود (Randic, 1993).

$$(6) \quad R = \sum \left(\frac{1}{mn} \right)^{\frac{1}{2}}$$

این مجموع‌بندی به ازای هر یال گراف ساختاری هیدروژن زدوده شامل یک جمله است، متغیرهای m و n ظرفیت نقاط مجاوری هستند که با هر یال به هم وصل می‌شوند.

شاخص پلت

شاخص پلت $F(G)$ از گراف G یک مولکول به صورت مجموع کل درجه پیوندها در آن گراف تعریف می‌شود که $D(e_i)$ درجه لبه e_i (تعداد پیوندهای متصل به e_i) را نشان می‌دهد. M تعداد پیوندها را نشان می‌دهد (Babujee, 2012).

$$F(G) = \sum_{i=1}^M D(e_i) \quad (7)$$

شاخص هاراری

شاخص هاراری یک گراف G به صورت $H(G)$ به صورت جداگانه توسط افرادی به نام Plavsic و همکاران وی و همچنین به وسیله Ivanciuc و همکاران او در سال ۱۹۹۳ معرفی گردید و به مناسبت تولد ۷۰ سالگی Frank. Harary به نام هاراری نام گذاری شد (Li & Fan, 2015; Xu et al., 2015). شاخص هاراری یک گراف ساده G به صورت:

$$H(G) = \sum_{\{u,v\} \in V(G)} \frac{1}{dG(u,v)} \quad (8)$$

معرفی شد که عبارت است از مجموعه‌ای از معکوس همه زوج رئوس $\{u,v\}$.

شاخص وینرپلاریتی

$Wp(G)$ برای یک گراف مولکولی G از مرتبه n ، تعداد جفت‌های نامرتب از رئوس u و v از گراف G می‌باشد، به طوری که فاصله بین u, v یعنی $dG(u, v)$ معادل ۳ در نظر گرفته می‌شود (Behmaram & yousefi Azari, 2011; Liu & Liu 2011; Deng, 2011).

$$Wp(G) = |\{\{u,v\} | dG(u,v) = 3, u, v \in V\}| \quad (9)$$

روش کار

روش کار

در این پژوهش، جهت بررسی ارتباط شاخص‌های توپولوژیکی با ضریب توزیع آب-اکتانل ($\log P$) در برخی آفت‌کش‌ها و پیش‌بینی این شاخص مهم در مولکول‌های مورد مطالعه از دو روش اقدام به جمع‌آوری اطلاعات مورد نیاز گردید. در بخش اول، تحقیق با استفاده از تئوری گراف به محاسبه برخی شاخص‌های توپولوژیکی نظیر وینر-پلاریتی (Wp)، هایپر-وینر (WW)، هاراری (H)، راندیک (X)، بالابان (J)، سگد (Sz)، وینر (W) و پلت (P) به کمک فرمول‌های ذکر شده برای تمام مولکول‌های مورد نظر پرداخته شد که نتایج در جدول ۱ آورده شده است. در بخش دوم، مقادیر تجربی $\log P$ یا $\log Kow$ مربوط به ۲۶ آفت‌کش که از طریق روش RP-HPLC محاسبه شده‌اند، جمع‌آوری گردیدند که نتایج به همراه منابع مورد استفاده در جدول ۳ ثبت شده‌اند. از این تعداد مولکول تحت مطالعه، موارد ۴ و ۷ و ۸ و ۹ و ۱۲ و ۱۳ و ۱۴ و ۱۵ و ۲۴ و ۲۵ از دسته حشره‌کش‌ها و بقیه علف‌کش محسوب می‌شوند. در بخش سوم، همبستگی شاخص $\log P$ مولکول‌های مورد نظر با شاخص‌های ساختاری محاسبه شده مورد بررسی قرار گرفت. بدین منظور به کمک نرم‌افزار اکسل، منحنی ارتباط خاصیت در مقابل توصیف‌گرهای ساختاری رسم گردید. در بخش چهارم، توسط نرم‌افزار SPSS با روش رگرسیون خطی چندگانه و با کمک روش برگشتی با توجه به ضرایب تعیین کننده مهم به دست آمده از این روش که عبارتند از ضریب همبستگی، مجذور ضریب همبستگی، ضریب فیشر، فاکتور دوربین واتسون و... بهترین توصیف‌گرها برای پیش‌گویی شاخص $\log P$ آفت‌کش‌ها مشخص گردید.

نتایج

همان‌گونه که گفته شد مقادیر مربوط به تمام شاخص‌های ساختاری مورد مطالعه به ترتیب: سگد، وینر-پلاریتی، هایپر-وینر، وینر، هاراری، بالابان، راندیک و پلت برای تمام مولکول‌های تحت بررسی مطابق فرمول‌های ۲ الی ۹ محاسبه شده و در جدول ۱ آورده شده‌اند.

جدول ۱- مقادیر شاخص‌های سگد، وینر-پلاریتی، هایپر-وینر، وینر، هاراری، بالابان، راندیک و پلت

Table 1- The values of topological indices: Sz, Wp, WW, W, H, J, X, P

No	Compound	Sz	Wp	WW	W	H	J	X	P
1	Ametryn	544	18	1077	384	41.11	2.27	7.15	38
2	Atratone	544	18	1077	384	41.11	2.27	7.15	38
3	Atrazine	456	16	870	319	37.08	2.37	6.61	36
4	Chlorfenvinphos	1021	29	2631	818	64.31	2.69	9.45	54
5	Cyanazine	627	20	1332	456	45.73	2.47	7.46	44
6	2,4-Dichlorophenoxy Acetic acid	378	15	723	266	33.06	2.08	6.09	34
7	p,p'-DDD	939	26	1820	603	56.39	1.85	8.58	52
8	4,4-DDE	939	26	1820	603	56.39	1.85	8.58	52
9	4,4-DDT	1038	28	2040	678	61.98	1.83	8.88	58
10	Desethyl atrazine	301	13	489	202	29.73	2.34	5.58	32
11	Desisopropyl atrazine	244	12	363	158	26	2.09	5.22	28
12	Dichlorvos	168	12	405	168	24.77	3.69	5.04	26
13	dieldrin	879	46	1203	520	72.53	1.38	8.78	88
14	Heptachlor	579	42	842	388	59.87	1.87	7.7	70
15	Methoxychlor	1412	32	3156	932	70.78	1.9	9.95	62
16	Prometon	639	19	1331	458	45.38	2.47	7.51	42
17	Prometryn	639	19	1331	458	45.38	2.47	7.51	42
18	Propazine	542	17	1096	386	41.2	2.24	6.97	40
19	Simazine	380	15	680	260	33.08	2.15	6.26	32
20	Simetryn	459	17	859	318	36.96	2.39	6.8	34
21	Terbumeton	631	19	1298	452	45.75	2.5	7.44	44
22	Terbuthylazine	534	17	1063	380	41.58	2.28	6.9	42
23	Terbutryn	631	19	1298	452	45.75	2.5	7.44	44
24	Tetrachlorvinphos	900	30	2090	690	60.6	2.59	8.86	54
25	Trichlorphon	185	18	400	185	30.2	4.53	5.28	34
26	Trietazine	528	20	1022	374	41.57	2.33	7.21	38

مقادیر تجربی $Log P$ مربوط به آفت‌کش‌های تحت مطالعه در جدول ۳ آورده شده است. پس از بررسی ارتباط میان مقادیر تجربی $Log P$ با شاخص‌های توپولوژی مورد مطالعه، معادلات خطی و ضرایب رگرسیون مربوطه تعیین شدند که (معادلات ۱۰-۱۷) حاصل این بررسی می‌باشند:

معادلات خطی (۱۷-۱۰) نشان می‌دهند که با توجه به مقادیر ضرایب رگرسیون (R^2)، میان مقادیر $\log P$ با مقادیر هیچ یک از شاخص‌ها در مولکول‌های تحت بررسی، همبستگی خوبی وجود ندارد:

$$\log P = 0.0857 P - 0.628 \quad R^2 = 0.6 \quad (10)$$

431

$$\log P = 0.9602 X - 3.8429 \quad R^2 = 0.7 \quad (11)$$

104

$$\log P = -1.5522 J + 6.854 \quad R^2 = 0.4 \quad (12)$$

011

$$\log P = 0.0958 H - 1.1879 \quad R^2 = 0.7 \quad (13)$$

496

$$\log P = 0.0057 W + 0.7064 \quad R^2 = 0.5 \quad (14)$$

645

$$\log P = 0.0015 WW + 1.3378 \quad R^2 = 0.4 \quad (15)$$

666

$$\log P = 0.1358 Wp + 0.2905 \quad R^2 = 0.5 \quad (16)$$

974

$$\log P = 0.0042 Sz + 0.5989 \quad R^2 = 0.6 \quad (17)$$

812

بنابر این مشخص گردید که هیچ‌یک از توصیف‌گرهای ساختاری مورد مطالعه را به‌تنهایی نمی‌توان جهت پیش‌بینی $\log P$ آفت‌کش‌های مورد نظر به‌کار برد، به همین دلیل روش رگرسیون خطی چندگانه به‌کار برده شد و تاثیر همه توصیف‌گرها باهم جهت پیش‌بینی شاخص $\log P$ آفت‌کش‌ها مورد بررسی قرار گرفت. در این روش به کمک نرم‌افزار SPSS و روش آماری MLR (رگرسیون خطی چندگانه) با انتخاب روش برگشتی و در نظر گرفتن خطای ۰/۰۵ اقدام به انجام محاسبات گردید.

نتایج این بررسی در جدول ۲ آمده است. نتایج این محاسبات نشان می‌دهد که ۳ مدل برای تعیین شاخص $\log P$ آفت‌کش‌های مورد مطالعه پیش‌بینی شده است. در هر مدل تعدادی از توصیف‌گرها به‌کار گرفته شدند و ضرایب تعیین‌کننده مهم مانند: ضرایب رگرسیون- مجذور ضرایب رگرسیون- ضرایب رگرسیون تنظیم شده- مقدار انحراف استاندارد- ضریب فیشر- مجذور میانگین و سطح معنی‌داری در آن‌ها مشخص گردیدند.

جدول ۲- مدل‌های پیش‌بینی - شاخص‌های پیش‌بینی کننده- مقادیر ضریب رگرسیون - مجذور ضریب رگرسیون -

ضریب رگرسیون Adjust- انحراف استاندارد- ضریب فیشر- مجذور میانگین- سطح معنی‌داری

Table 2- Predictive Models, Predictor Indices, The values of R, R², R²_{Adj}, σ, F Statistic, Mean, Significance level

Model	Predictors	R	R ²	R ² _{Adj}	σ	Fisher Statistic	Mean Square	significance level
1	X, H, P, J, Wp, WW, Sz	0.951	0.905	0.868	0.539 ^٦	24.436	7.114	0.000
2	X, H, P, Wp, WW, Sz	0.950	0.903	0.873	0.5291	29.599	8.286	0.000
3	X, H, P, WW, Sz	0.949	0.901	0.877	0.520 ^٧	36.603	9.923	0.000

به‌طور کلی در میان همه مدل‌های پیشنهادی برای پیش‌بینی هر خاصیت، مدلی که از ضریب فیشر بالاتری برخوردار است اهمیت بیشتری داراست، چون شامل تعداد بیشتری از داده‌ها و متغیرهای به‌کار رفته است. صفر بودن Sig نشان دهنده آن است که سطح معنی‌داری کوچک‌تر و بنابر این مدل خطی به‌کار رفته معنادار است. ضمناً ضرایب همبستگی R, R² نشان می‌دهند که چند درصد تغییرات متغیر وابسته توسط متغیرهای مستقل تبیین می‌شوند و هر چقدر تفاوت میان R², R²_{Adj} کمتر باشد نشانگر آن است که متغیرهای مستقل که به مدل اضافه شده‌اند به درستی انتخاب گردیده‌اند. همچنین در صورتی که مقدار ضریب دوربین واتسون بین صفر تا چهار باشد نشان‌دهنده مناسب بودن مدل پیشنهادی است که با توجه به مقدار دوربین واتسون برابر با ۲/۵۱۱ مدل نهایی مورد تایید قرار می‌گیرد.

معادلات تعیین‌کننده شاخص $Log P$ بر اساس ضرایب به‌دست آمده در هر سه مدل به شرح زیر می‌باشند:

$$Log P = -3.379 X + 0.948 H - 0.469 P - 0.172 J - 0.060 Wp - 0.00 WW + 0.012 Sz + 8.618 \quad (18)$$

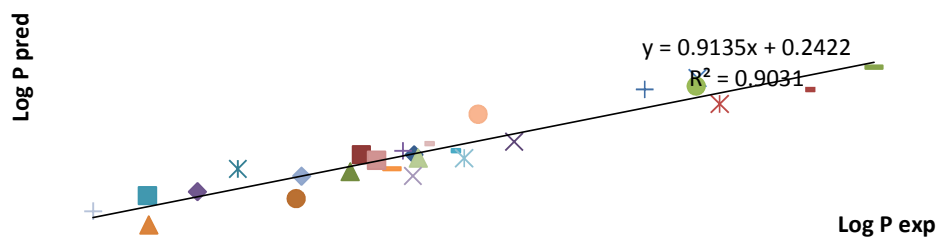
$$Log P = -3.069 X + 0.892 H - 0.450 P - 0.054 Wp - 0.008 WW + 0.013 Sz + 7.211 \quad (19)$$

$$Log P = -2.494 X + 0.749 H - 0.411 P - 0.007 WW + 0.013 Sz + 6.019 \quad (20)$$

نتایج این بررسی‌ها نشان می‌دهد که مدل سوم با فاکتور دوربین واتسون ۲/۵۱۱ و با بیشترین مقدار ضریب فیشر برابر با ۳۶/۶۰۳ و نیز کمترین انحراف از استاندارد برابر با ۰/۵۲۰۶۷ و بالاترین مقدار مجذور میانگین برابر با ۹/۹۲۳ از سایر مدل‌های پیشگویی بهتر است ولی با توجه به مقدار بسیار کم ضریب شاخص هایپر-وینر در معادله سوم که برابر با ۰,۰۰۷ می‌باشد، می‌توان از تاثیر آن صرف نظر نمود. بنابراین معادله نهایی برای تعیین $Log P$ مولکول‌های تحت بررسی به‌صورت زیر در خواهد آمد:

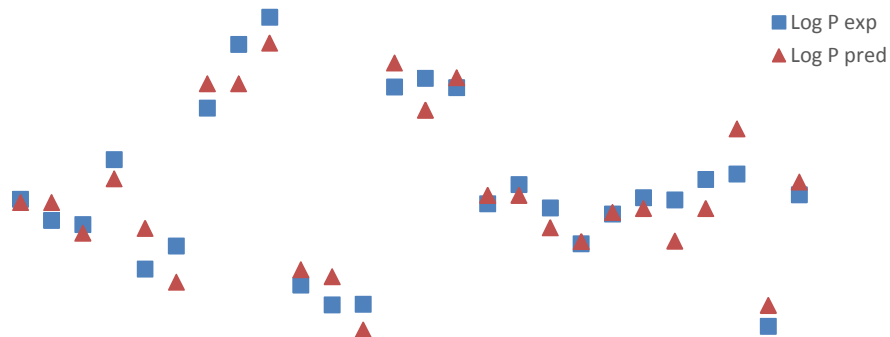
$$Log P = -2.494 X + 0.749 H - 0.411 P + 0.013 Sz + 6.019 \quad (21)$$

در جدول ۳ آمده است. حال برای حصول اطمینان از نتیجه حاصل، نمودارهای نشان‌دهنده ارتباط میان مقادیر تجربی $Log P$ با مقادیر پیش‌بینی شده $Log P$ با توجه به مدل رگرسیون پیشنهادی (21)، به شرح زیر رسم شدند:



شکل ۱- نمودار ارتباط میان مقادیر تجربی $Log P$ با مقادیر پیش‌بینی شده $Log P$ با توجه به (معادله 21)

با توجه به مقدار ضریب رگرسیون به دست آمده در این نمودار ($R^2=0.9015$) معلوم می‌شود که روش به کار رفته برای پیشگویی $Log P$ آفت‌کش‌ها از دقت بسیار خوبی برخوردار می‌باشد. همچنین برای بررسی نحوه توزیع یکسان خطاها همان‌گونه که در شکل ۲ آمده است، مقادیر اختلاف $Log P$ تجربی با مقادیر محاسبه شده از طریق معادله 21 یک الگوی نسبتاً تصادفی را نشان می‌دهد و مشخص می‌شود که خطاها در اطراف محور X ها از توزیع تقریباً یکسانی برخوردارند (جدول 2 مقادیر باقیمانده‌ها را نشان می‌دهد). این مدل تصادفی نشان می‌دهد که مدل خطی برای پیش‌بینی شاخص $Log P$ مولکول‌های مورد مطالعه مناسب و معقول به نظر می‌رسند و این نمایانگر مناسب بودن الگوهای انتخابی برای پیش‌بینی ضریب توزیع آب-اکتانل آفت‌کش‌ها می‌باشد.



شکل ۲- نمودار مقایسه مقادیر تجربی با مقادیر پیش‌بینی شده مربوط به $Log P$

جدول ۳- آفت‌کش‌های مورد مطالعه، مقادیر تجربی و پیش‌بینی $Log P$ رفرنس‌ها، خطاها

Table 3- Studied Pesticides, $LogP_{exp}$, References, $Log P_{Pred}$, Residuals

No	Pesticides	$LogP_{exp}$	Reference	$Log P_{Pred}$	Residuals
1	Ametryn	3.07	Felsot & Dahm, 1979	3.01	0.05
2	Atraton	2.69	Felsot & Dahm, 1979	3.01	-0.32
3	Atrazine	2.61	Felsot & Dahm, 1979	2.45	0.16
4	Chlorfenvinphos	3.79	Dubelman & Bremer, 1983	3.44	0.35
5	Cyanazine	1.8	Felsot & Dahm, 1979	2.54	-0.74
6	2,4-Dichlorophenoxy Acetic acid	2.22	Braumann & Grimme, 1981	1.56	0.66
7	p,p'-DDD	4.73	Mc Duffie, 1981	5.17	-0.44
8	4,4-DDE	5.89	Mc Duffie, 1981	5.17	0.72
9	4,4-DDT	6.38	Gustafson, 1989	5.91	0.47
10	Desethylatrazine	1.51	Felsot & Dahm, 1979	1.79	-0.28
11	desisopropylatrazine	1.15	Felsot & Dahm, 1979	1.66	-0.51
12	Dichlorvos	1.16	Karikoff, 1981	0.69	0.47
13	dieldrin	5.11	Gustafson, 1989	5.55	-0.44
14	Heptachlor	5.27	Mc Duffie, 1981	4.69	0.58
15	Methoxychlor	5.1	Briggs & Bromilov, 1982	5.28	-0.18
16	Prometon	2.99	Felsot & Dahm, 1979	3.14	-0.15
17	Prometryn	3.34	Felsot & Dahm, 1979	3.14	0.19
18	Propazine	2.91	Felsot & Dahm, 1979	2.55	-0.54
19	Simazine	2.26	Felsot & Dahm, 1979	2.30	-0.04
20	Simetryn	2.8	Felsot & Dahm, 1979	2.83	-0.03
21	Terbumetron	3.1	Felsot & Dahm, 1979	2.90	0.19
22	Terbutylazine	3.06	Felsot & Dahm, 1979	2.31	0.74
23	Terbutryn	3.43	Felsot & Dahm, 1979	2.90	0.53
24	Tetrachlorvinphos	3.53	Dubelman & Bremer, 1983	4.35	-0.82
25	Trichlorphon	0.76	Karikoff, 1981	1.14	-0.38
26	Trietazine	3.15	Felsot & Dahm, 1979	3.38	-0.23

بحث و نتیجه گیری

با توجه به این که تعیین خواص فیزیکی-شیمیایی آفت‌کش‌ها من جمله مقدار ضریب توزیع آب-اکتانال از طریق روش‌های آزمایشگاهی با اثرات سمی بالا بر بافت‌های بدن همراه است، لذا تعیین مقدار دقیق این ضرایب در آن‌ها از روش‌های غیر آزمایشگاهی، ضروری است. نتایج این تحقیق نشان داد که شاخص‌های توپولوژیکی مورد استفاده در این تحقیق در مدل‌سازی با روش رگرسیون خطی و تعیین شاخص $\log P$ مربوط به آفت‌کش‌ها نقش بسیار مهمی را ایفا می‌کنند. به کمک روش رگرسیون خطی چندگانه و مدل‌سازی توسط متد QSPR معلوم گردید که از میان شاخص‌های وینر، هایپر-وینر، وینرپولاریتی، هاراری، بالابان، راندیک، سگد و پلت برای پیش‌بینی $\log P$ مناسب‌ترین توصیف‌گرهای ساختاری، شاخص‌های پلت، هاراری، سگد و راندیک می‌باشند.

Reference

- Babujee, B. J. 2012.** Topological Indices and New Graph Structures Journal of Applied Mathematical Sciences, 6 (108): 5383-5401 .
- Balaban, A. T. 1982.** Highly Discriminating Distance-Based Topological Index. Chemical Physics Letters, 89 (5): 399-404.
- 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, pp: 479-4987.
- Behmaram, A. and Yousefi Azari, H. 2011.** Further Results on Wiener Polarity Index of Graphs. Iranian Journal of Mathematical Chemistry, 2: 67-70.
- 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, MATH Communications in Mathematical and in Computer Chemistry, 70: 1005-1021.
- Biagi, G. L., Guerra, M. C., Barbaro, A. M., Recanatini, M., Borea, T. A. and Sapone, A. 1991.** Lipophilicity indices of triazine herbicides, Elsevier, pp: 109: 33-40.
- Braumann, T. and Grimme, L. H. 1981.** Determination of Hydrophobic Parameters for pyridazinone herbicides by liquid-liquid partition and reversed-phase high-performance liquid chromatography Journal Chromatography A, 206(1): 7-15.
- Braumann, T. 1986.** Determination of Hydrophobic Parameters by Reversed-Phase Liquid Chromatography: Journal Chromatography A, 373: 191-225.
- Briggs, G. G. and Bromilov, R. H. 1982.** Relationships between lipophilicity and root uptake and translocation of non-ionised chemicals by barley. Pest Management Science, 13(5): 495-504.
- Cash, G., Klavzar, S. and Petkovsek, M. 2002.** Three Methods for Calculation of the Hyper-Wiener Index of Molecular Graphs, J. Chemical Information and Computer Sciences, 42: 571-576.
- Chiou, C. T., Freed, V. H., Schmedding, D. W. and Kohnert, R. L. 1977.** Partition coefficient and bioaccumulation of selected organic chemicals, Environmental Science & Technology, 11:475-478.
- Chiou, C. T., Schmedding, D. W. and Manes, M. 1982.** Partitioning of organic compound in octanol water systems. Environmental Science & Technology, 16(1): 4-10.
- Deng, H. 2011.** On the Extremal Wiener Polarity Index of Chemical Trees. MATCH, Communications in Mathematical and in Computer Chemistry, 66: 305-314.
- Dubelman, S. and Bremer, M. J. 1983.** Determination of octanol/water partition coefficient of MAPC products. Rept. No MSL-3219, Monsanto Company, Agricultural Research. Division, Street. Louis, 799pp.

- Felsot, A. and Dahm, P. A. 1979.** Sorption of organophosphorus and carbamate insecticides by soil. *Journal of Agricultural and Food Chemistry*, 27: 557-563.
- 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.
- Gilden, R. C., Huffling, K. and Sattler, B. 2010.** Pesticides and health risks. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 39(1): 103-110.
- Gustafson, D. I. 1989.** Groundwater ubiquity score: A simple method for assessing pesticide leachability. *Environmental Toxicology and Chemistry*, 8: 339-357.
- Gutman, I. 1994.** *Graph Theory Notes*, New York, 27: 9-15.
- Isnard, P. and Lambert, S. 1989.** Aqueous solubility and octanol/water coefficient correlation, *Chemosphere*, 18: 1837-1853.
- Khadikar, P. V., Deshpande, N. V., Kale, P. P., Dobrynin, A., Gutman, I and Domotor, G. 1995.** The Szeged index and an analogy with the Wiener index. *Journal of Chemical Information and Computer Sciences*, 35: 547-550.
- Karikoff, S. W. 1981.** Semi-empirical estimation of sorption of hydrophobic pollutants on natural sediments and soil. *Chemosphere*, 10: 833-846.
- Klavzar, S., Rajapakse, A. and Gutman, I. 1996.** The Szeged and the Wiener index of graphs. *Applied Mathematics Letters*, 9(5): 45-49.
- Klein, D. J., Lukovits, I. and Gutman, I. 1995.** On the definition of the hyperWiener index for cycle-containing structures, *Chemical Information and Computer Sciences*, 35: 50-52.
- Lailong, M. and Hongmei, H. 2011.** Quantitative Structure–Property Relations (QSPRs) for Predicting the Standard Absolute Entropy (S_{298K}°) of Gaseous Organic Compounds. *Industrial & Engineering Chemistry Research*, 50(14): 8764-8772.
- Li, X. X. and Fan, Y. Z. 2015.** The connectivity and the Harary index of a graph, *Discrete Applied Mathematics*, 181:167-173.
- Liu, M. and Liu, B. 2011.** On the Wiener Polarity Index. *MATH. Communications in Mathematical and in Computer Chemistry*, 66: 293-304.
- Lyman, W. J., Reehl, W. F. and Rosenblatt, D. H. 1982.** *Handbook of chemical property estimation methods*, McGraw-Hill Book Company, New York, pp. 9-1 to 9-85.
- Mc Duffie, B. 1981.** Estimation of octanol/water coefficients for organic pollutants using reversed-phase HPLC. *Chemosphere*, 10: 73-83.
- Mohammadinasab, E. and Goodarzi, M. 2011.** Theoretical Study of the Relationship Between Wiener, Padmakar-Ivan, and Szeged Topological Indices in Contrast to Energy, Electric Moments and Partition Coefficient of Armchair Polyhex Carbon Nanotubes with Various Circumference and Fixed Lengths, *Fullerenes, Nanotubes and Carbon Nanostructures*, 19: 550-563.
- Randic, M. 1993.** Novel molecular descriptor for structure-property studies, *Chemical. Physical. Letter*, 211: 478-483.
- Shafiei, F. 2015.** Relationship between Topological Indices and Thermodynamic Properties and Monocarboxylic Acids Applications in QSAR. *Iranian Journal of Mathematical Chemistry*, 6: 15-28.
- Wiener, H. 1947.** Structural Determination of Paraffin Boiling Points. *J. American. Chemistry. Society*, 69: 17-20.
- Xu, K., Das, K. C. and Trinajstić, N. 2015.** *The Harary Index of a Graph*, Springer-Verlag Berlin Heidelberg, 74 pp.

Prediction of the Log P of pesticides using multiple linear regression

E. Mohammadinab

Assistant Professor, Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran

Abstract

Pesticides are considered as one of the most significant environmental pollutants. The transfer of pesticides into the living organism and the maximum residue limit of pesticides is one of the main environmental concerns. Pesticides have got the attention due to their various applications in biochemistry, environment and agriculture. Consequently, the physical and chemical characteristics of pesticides particularly the pesticides toxicity needs to be taken into consideration. The aim of this research is investigation the relationship between Log P of some pesticides with topological descriptors by graph theory and multiple linear regression methods. Quantitative structure-property relationship study was used and results was showed that Platt, Harary, Randic, Szeged indices are suitable for prediction of Log P the pesticides than the other topological indices. The best model in this study indicated that those structural descriptors, play an important role in effect on Log P of pesticides. For the first time, the relationship between Log P of pesticides and some topological indices using SPSS and multiple linear regression model is investigated.

Keywords: Pesticides, Log P, Multiple linear regression

* Corresponding Author, E-mail: e-mohammadinab@iau-arak.ac.ir
Received: 25 May 2016- Accepted: 1 Oct. 2016