

Investigación

Estudio de ciertas modificaciones al índice de H. P. Schultz. Uso de hidrocarburos saturados como modelo

E. Cornwell

Departamento de Química Inorgánica y Analítica, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Casilla 233, Santiago, Chile.

Recibido el 25 de junio del 2001; aceptado el 13 de diciembre del 2001

Resumen. Se proponen tres modificaciones principales a la matriz de distancia correspondiente al índice topológico de H. P. Schultz. (2-MTI). Se evaluaron las diferencias entre los índices resultantes de estas modificaciones y el índice 2-MTI. La modificación consistente en el uso de distancias reales entre los átomos de carbono de la molécula para configurar la matriz de distancia, se caracteriza como $C_{g.d.A^0}$. La regresión lineal entre los logaritmos de los tiempos de retención relativos de 28 hidrocarburos saturados y el índice $C_{g.d.A^0}$, presenta un mayor coeficiente de determinación y de Fisher que el índice propuesto por H. P. Schultz. Sin embargo, en este caso, estas diferencias no son significativas, este hecho valida el uso de la matriz de distancia cuyos elementos pertenecen al conjunto de los números naturales en vez de números que representan las distancias reales entre los átomos.

Palabras claves: Índice de Schultz, matriz de distancia, índice topológico, hidrocarburos saturados.

Abstract. Three main modifications to the distance matrix corresponding to H. P. Schultz topological index (2-MTI) are proposed. The differences between the modifications of these resulting indices and 2-MTI index are evaluated. The modification is based in the use of actual distances between carbon atoms of the molecule to configure the distance matrix, which is characterized as topological index $C_{g.d.A^0}$. The lineal regression between the logarithms of the relative retention times of 28 saturated hydrocarbons and $C_{g.d.A^0}$, present greater determination and Fisher coefficients than that of the proposed H. P. Schultz index. Nevertheless, in this case these differences are not significant. This fact validates the original proposal of the use of the distance matrix, whose elements belong to the set of the natural numbers instead of numbers representing the real physical distances between atoms.

Keywords: Schultz index, the matrix distance, topological index, saturated hydrocarbons.

Introducción

Las propiedades de cada molécula pueden ser traducidas a un número escalar llamado índice. Si en la conceptualización de este escalar participan elementos topológicos, se está en presencia de un índice topológico.

La existencia de una relación inequívoca entre una propiedad de una molécula y la magnitud del índice topológico que se le atribuye, se denomina invariante, cuando se tiene un valor de índice topológico asociado a dos o mas moléculas, este hecho se describe como grados degenerativos propios del índice en cuestión [1].

El estudio de estas relaciones es tratada por la disciplina denominada, relación propiedad y estructura química (QSPR) Un subconjunto contenido en esta disciplina trata en forma específica de la relación entre los tiempos de retención cromatográficos y las estructuras químicas (QSRR) [2].

El índice topológico de H. P. Schultz [3] denominado 2-MTI, el cual es descriptivo de las estructuras moleculares solo para los hidrocarburos saturados, se define a través de tres elementos matemáticos: la matriz cuadrada de adyacencia denominada $[A]_{n \times n}$, la matriz cuadrada de distancia denominada $[B]_{n \times n}$, ambas constituidas por elementos provenientes de consideraciones topológicas descriptivas de las estructuras

moleculares y el vector multiplicativo anterior denominado $V_{1 \times n}$, cuyos elementos son equivalentes a los elementos a_{ii} de la matriz $[A]^2_{n \times n}$, esto último válido para el caso particular de los hidrocarburos saturados.

Los elementos a_{ij} de la matriz de adyacencia $[A]_{n \times n}$ son iguales a 0 para indicar que los átomos de carbono i, j de la molécula no son adyacentes, en caso contrario, a_{ij} es igual a 1. Los elementos b_{ij} de la matriz de distancia $[B]_{n \times n}$ indican distancias topológicas entre los átomos de carbono i, j de una molécula, Distancia topológica implica distancias medidas a través de los enlaces por el camino mas corto, utilizando elementos numéricos pertenecientes al subconjunto de los naturales, así, como ejemplo, la distancia entre dos átomos de carbono C_i y C_j separados por tres enlaces σ , presenta una distancia igual a 3, siendo el número de carbonos involucrados igual a 4.

El índice topológico de H. P. Schultz queda definido por un escalar denominado 2-MTI [3] su cálculo obedece a las siguientes ecuaciones (1), (2)

$$V_{1 \times n} * ([A]_{n \times n} + [B]_{n \times n}) = [e_1, e_2, e_3, \dots, e_n] \quad (1)$$

$$2 - \text{MTI} = \sum_{f=1}^N e_f \quad (2)$$

donde $f = 1, N$, significa, súmese los elementos desde e_1 hasta e_n

Para un grupo de moléculas, la relación entre sus propiedades y los índices topológicos se expresan generalmente mediante regresiones lineales asociadas a los parámetros estadísticos que las caracterizan [4-13].

El objetivo de este trabajo es: estudiar ciertas modificaciones en la matriz cuadrada de distancia $[B]_{n \times n}$ que define el índice topológico de H. P. Schultz (2-MTI). Estas modificaciones consisten en proponer elementos de esta matriz b_{ij} y b_{ii} distintos a los expresados en el índice topológico 2-MTI original, manteniendo las definiciones propias tanto del vector multiplicativo anterior $V_{1 \times n}$ como de la matriz de adyacencia cuadrada $[A]_{n \times n}$. Comparar la eficiencia de estos índices topológicos corregidos en base de las modificaciones anteriormente expuestas, versus el índice 2-MTI original. Estas comparaciones consisten en evaluar estadísticamente el grado de correlación entre la variable dependiente e independiente que se describen a continuación.

En forma anexa, se realizó una comparación entre el índice topológico resultante del valor propio mayor positivo [5] ($[\lambda^+ + [B]_{n \times n}]$) proveniente de la matriz resultante de la operación sobre la expresión siguiente: $([A]_{n \times n} + [B]_{n \times n})$ y mi proposición, donde se reemplaza la matriz de distancia $[B]_{n \times n}$ de la expresión anterior por la matriz de distancia $[B^A]_{n \times n}$ que se definirá a continuación.

Se proponen 3 modificaciones a la matriz cuadrada de distancia $[B]_{n \times n}$ original de H. P. Schultz, que implican los siguientes cambios:

1. Cambio de los elementos b_{ij} de la matriz de distancia $[B]_{n \times n}$ por los elementos $(b_{ij})^{-1}$, manteniendo $b_{ii} = 0$. Este cambio define la matriz cuadrada de distancia $[D']_{n \times n}$. Los demás elementos mantienen sus definiciones originales, esto se refiere a: $(V_{1 \times n}, [A]_{n \times n})$. Esto implica un índice topológico 2-MTI modificado con el elemento distinto $[D']_{n \times n}$ y $b_{ii} = 0$.
2. El mismo cambio mencionado anteriormente pero con elementos $b_{ii} = 1$, definiendo la matriz cuadrada de distancia $[D'']_{n \times n}$. Esto implica un índice topológico 2-MTI modificado con elemento distinto $[D'']_{n \times n}$.
3. Reemplazo de la matriz de distancia $[B]_{n \times n}$ del índice original 2-MTI por la matriz $[B^A]_{n \times n}$ cuyos elementos d_{ij} son distancias físicas reales entre los átomos de carbono i, j de la molécula, distancia medidas por el camino más corto y realizado a través de los enlaces σ .

Esta modificación origina el índice propuesto $C_{g.d.A^\circ}$, cuyos elementos d_{ii} son igual a 0. Todos los otros elementos que definen el índice 2-MTI se mantienen en sus definiciones originales, vale decir los elementos $V_{1 \times n}$ y $[A]_{n \times n}$. Los índices 2-MTI modificados fueron calculados a través de las mismas operaciones matemáticas que las aplicadas en el cálculo del 2-MTI original y fueron comparadas con el índice de conectividad químico [14] (χ_v). Se utilizó este índice topológico como elemento comparativo debido que es uno de los más eficientes [14].

Como modelo de trabajo para discriminar las diferencias entre el índice 2-MTI original y los modificados en este trabajo, se utilizaron 28 hidrocarburos saturados y se emplearon como variable dependiente los logaritmos de los tiempos de retención relativos al *n*-heptano ($\log t_{rr}$). Cabe señalar que, los datos de tiempos de retención relativos bibliográficos [15] (t_{rr}) están descritos paralelamente en el Guide to Analysis of Hydrocarbons by Gas Chromatography 2° Edition Poly Sciences Corporation (1968). Se utilizó como variable dependiente en las regresiones el logaritmo del tiempo de retención relativo, debido que, expresados en esta forma, presentan relaciones de correlación lineales con los parámetros fisicoquímicos de los hidrocarburos saturados.

Las diferencias existentes entre el índice de H. P. Schultz (2-MTI original) y los propuestos (modificaciones sobre el índice 2-MTI original) se evaluaron mediante la comparación de los parámetros estadísticos inherentes a las regresiones lineales [10,13] esto implica la correlación de $\log trr$ como variable dependiente versus cada índice tratado en este trabajo excepto el índice de conectividad químico que se utilizó como referente. Del mismo modo, se estudiaron las correlaciones lineales que responden al modelo $y = mx + n$ siendo la variable dependiente los logaritmos de retención relativos calculado versus los obtenidos a través de los datos experimentales. Estas relaciones, demuestran las bondades de las proposiciones como así mismo las diferencias con la proposición original.

Desarrollo

El cálculo del índice de conectividad químico de L. B. Kier (χ_v) para cada hidrocarburo saturado se realizó según bibliografía [16-17].

Los cálculos de los índices topológicos 2-MTI modificados y $C_{g.d.A^\circ}$ se obtienen mediante las siguientes ecuaciones: (3), (4), (5), (6), (7), (8).

matriz D'

$$V_{1 \times n} * \begin{pmatrix} [A]_{n \times n} + & \left| \begin{array}{ccc} 0 & (b_{12})^{-1}K & (b_{1j})^{-1} \\ M & 0 & M \\ (b_{12})^{-1} & K & 0 \end{array} \right| \end{pmatrix} = [p_1, p_2, p_3, K, p_N] \quad (3)$$

$$\text{2-MTI modificado con matriz } D' = \sum_{i=1}^N p_i \quad (4)$$

matriz D''

$$V_{1 \times n} * \begin{pmatrix} [A]_{n \times n} + & \left| \begin{array}{ccc} 1 & (b_{12})^{-1}K & (b_{1j})^{-1} \\ M & 1 & M \\ (b_{12})^{-1} & K & 1 \end{array} \right| \end{pmatrix} = [l_1, l_2, l_3, K, l_N] \quad (5)$$

$$\text{2-MTI modificado con matriz } D'' = \sum_{i=1}^N l_i \quad (6)$$

Matriz $[B^A]$

$$V_{lxn} \begin{pmatrix} [A]_{nxn} + \begin{vmatrix} 1 & d_{12}K & d_{1j} \\ M & 1 & M \\ d_{ii} & K & 1 \end{vmatrix} \end{pmatrix} = [m_1, m_2, m_3, K, m_N] \quad (7)$$

$$C_{g.d.A^\circ} = \sum_{i=1}^N m_i \quad (8)$$

Los valores propios para cada hidrocarburo se obtuvieron según las expresiones dadas en la introducción, para tal fin se utilizó el software incorporado en la calculadora Hewlett Packard 48GX.

La obtención de los elementos b_{ij} de la matriz $[D^A]_{n \times n}$ se realizó mediante el software Hyperchem [18] previa optimización de las estructuras moleculares de los hidrocarburos saturados correspondientes mediante el método semiempírico AM1, usando el algoritmo Steep-Descent con gradiente RMS de 1 Kcal / (A° mol) hasta convergencia y luego el algoritmo Polak-Ribiere a 0.01 Kcal / (A° mol) hasta convergencia.

En la Tabla 1 se presentan para los 28 hidrocarburos saturados sus tiempos de retención relativos al *n*-heptano [15] los logaritmos de estos tiempos de retención y los valores calculados mediante los diversos índices topológicos bajo estudio correspondientes a los hidrocarburos saturados.

Utilizando como variable dependiente $\log t_{rr}$, e independiente los índices topológicos señalados en la Tabla 1, se obtuvieron las regresiones lineales óptimas y sus parámetros estadísticos a través del software Origin 4.0 [19] el cual incluye un análisis ANOVA sobre la regresión, ver Tabla 2.

Los parámetros estadísticos asociados a las regresiones lineales [10-13] como son el coeficiente de correlación (*r*), el test de Fisher (*F*) propios del modelo $Y = a + bx$ y el coeficiente de determinación (R^2) propio del modelo $y = a + bx + cx^2$, permiten un análisis comparativos del grado de asociación entre las variables dependientes e independiente en los distintos casos estudiados. De esta forma se determinó la eficiencia de cada índice topológico en estudio, en cuanto al grado de ajuste al modelo de regresión lineal propuesto.

Tendencias al valor 1 del coeficiente de determinación (R^2) o valores -1 o 1 del índice de correlación (*r*), como altos valores de *F*, son consistentes con diferencias pequeñas entre la variable dependiente experimental y calculada por la ecuación de regresión respectiva.

En la Tabla 3 se señalan los estadígrafos correspondientes al índice de correlación (*r*) y de Fisher *F* correspondientes a las regresiones entre los valores de los tiempos de retención relativos calculados versus experimentales [15].

Cabe señalar que los valores calculados se obtuvieron mediante las ecuaciones de regresión anotadas en la Tabla 2 correspondiente a los índices en estudio como al 2-MTI original.

Discusión

De la Tabla 2 se deduce que los índices topológicos correspondientes a aquellos que contienen modificaciones consistentes en las permutaciones de la matriz cuadrada de distancia $[B]_{n \times n}$ por las matrices cuadradas $[D']_{n \times n}$ y $[D'']_{n \times n}$ respectivamente, presentan en forma relativa a los índices topológicos de H. P. Schultz (2-MTI corregido) y el índice $C_{g.d.A^\circ}$ propuesto por el autor en este trabajo, valores menores tanto del índice de correlación (*r*) como del índice de Fisher (*F*). Esto indica que, la utilización de los elementos $(b_{ij})^{-1}$ en vez de los elementos (b_{ij}) , en la matriz cuadrada de distancia $[B]_{n \times n}$ para los casos en los cuales $b_{ii} = 0$ (matriz $[D']_{n \times n}$) o bien $b_{ii} = 0$ (matriz $[D'']_{n \times n}$), producen un decremento en la correlación de las variables dependiente e independientes, lo anterior implica las relaciones entre los $\log t_{rr}$ vs. los índices topológicos que contienen los elementos $[D']_{n \times n}$ y $[D'']_{n \times n}$, respectivamente. Este decremento es relativo a la correlación en la relación $\log t_{rr}$ vs. el índice 2-MTI original.

Los índices topológicos definidos en base de los valores propios que son función de las matrices cuadradas de distancias $[B]_{n \times n}$ o $[B^A]_{n \times n}$ respectivamente, son equivalentes y producen idénticos estadígrafos al relacionarse con la variable dependiente ($\log t_{rr}$), presentando ambas valores de *r* y *F* bajos relativos tanto al índice 2-MTI original y el índice $C_{g.d.A^\circ}$, pero son mayores a los caracterizados anteriormente. Estos dos análisis para este modelo (hidrocarburos saturados), induce a pensar que la opción que implica utilizar las matrices cuadradas $[D']_{n \times n}$ y $[D'']_{n \times n}$ es de poca utilidad.

El índice topológico propuesto $C_{g.d.A^\circ}$, cuya matriz de distancia $[B^A]_{n \times n}$ contiene elementos $b_{ii} = 0$ y los elementos b_{ij} este último, significando la distancia a través de los enlaces σ por el camino mas corto entre los átomos de carbono del hidrocarburo, presenta respecto al índice de H. P. Schultz (2-MTI original) valores de índice de determinación R^2 e índice de Fisher (*F*) mayores que el índice 2-MTI original, esto implica que la ecuación $\log t_{rr}$ en función de $C_{g.d.A^\circ}$, presenta una mayor cohesión entre las variables, dependiente e independiente (mayor R^2) y mayor poder predictivo [20], esto es, menor diferencia entre los valores calculados de t_{rr} por la ecuación de regresión y los valores de t_{rr} experimentales (mayor *F*). Sin embargo, estas diferencias anteriormente anotadas, no son estadísticamente significativas bajo el modelo empleado.

Tanto el índice 2-MTI original como el propuesto $C_{g.d.A^\circ}$, responden para un R^2 y un *F* óptimo a un mismo modelo matemático de regresión. Los factores A1, B1, b2 del modelo de regresión para el índice topológico $C_{g.d.A^\circ}$, presentan menor error que los correspondientes al modelo 2-MTI original.

Todos los índices analizados en este trabajo presentan una menor correlación comparado con el índice de conectividad químico, este último se utilizó solo como referente. Cabe hacer notar que el índice 2-MTI original está en etapa de desarrollo.

La aceptación o exclusión del índice topológico $C_{g.d.A^\circ}$, no solo puede depender del grado de significancia estadística

En la página 187 se indica en forma detallada el cálculo del índice de H. P. Schultz para el hidrocarburo 2-metilbutano.

Tabla 1. Hidrocarburos saturados, tiempos de retención relativos en G.L.C. e índices topológicos.

Hidrocarburos saturados	t_{rr}	$\log t_{rr}$	$^1\chi_v$	2-MTI [B] _{n × n}	2-MTI [D'] _{n × n}	2-MTI [D''] _{n × n}	C _{g.d.A°}	$\lambda + ([A]_{n × n} + [B^A]_{n × n})$	$\lambda + ([A]_{n × n} + [B]_{n × n})$
<i>n</i> -heptano	1.0000	0.000	3.41221	204	50.9760	62.9760	297.2539	26.6847	18.1783
<i>n</i> -propano	0.0088	-2.055	1.41421	16	12.5000	16.5000	21.0720	5.3599	4.0000
2-metilpropano	0.0205	-1.6882	1.73205	36	25.5000	31.5000	48.3384	8.3322	6.0000
<i>n</i> -butano	0.0318	-1.4976	1.91421	38	21.7222	27.7222	52.2788	9.1396	6.5311
2,2-metilpropano	0.0366	-1.4365	2.00000	64	43.0000	51.0000	86.9262	11.3634	8.0000
<i>n</i> -pentano	0.104	-0.9830	2.41421	74	31.2914	39.2914	104.6782	13.9494	9.7297
2,2-dimetilbutano	0.155	-0.8097	2.56066	106	53.6666	63.6666	148.7814	16.0651	11.0769
2,3-dimetilbutano	0.220	-0.6576	2.64274	108	49.8888	59.8888	152.7942	16.5639	11.4031
2-metilpentano	0.229	-0.6402	2.77006	118	45.3610	55.3610	170.5999	18.3512	12.5154
3-metilpentano	0.268	-0.5719	2.80806	114	45.7360	55.7360	162.6154	17.6946	12.1719
<i>n</i> -hexano	0.326	-0.4868	2.91421	128	41.0658	51.0658	184.3201	19.7971	13.6116
2,2-dimetilpentano	0.428	-0.3686	3.06066	170	63.9302	75.9302	239.7235	21.9750	15.0970
2,4-dimetilpentano	0.452	-0.3449	3.12589	176	59.7776	71.7776	253.9900	22.9844	15.6472
2,2,3-trimetilbutano	0.489	-0.3107	2.94338	156	68.8332	80.8332	228.4056	20.6680	13.8023
3,3-dimetilpentano	0.608	-0.2161	3.12132	162	64.6804	76.6804	232.1378	21.2842	14.4942
2-metilhexano	0.685	-0.1643	3.27006	190	55.3404	67.3404	275.5430	24.8381	16.9205
2,3-dimetilpentano	0.712	-0.1475	3.18074	168	60.5276	72.5276	242.1665	22.1546	15.0731
3-metilhexano	0.758	-0.1203	3.30806	182	55.8576	67.8576	263.6887	24.0076	16.3497
3-etilpentano	0.843	-0.0742	3.34607	174	63.1244	68.3746	251.8643	23.2870	15.7492
2,2,4-trimetilpentano	0.866	-0.0625	3.41650	242	78.7414	92.7414	352.7112	27.3118	18.5101*
2,2-dimetilhexano	1.225	0.0881	3.84741	260	74.1148	88.1148	378.5460	29.4765	19.9327
2,2,3,3-tetrametilbutano	1.251	0.0973	3.25000	214	88.4996	102.4996	307.8792	24.2468	16.3459
2,5-dimetilhexano	1.362	0.1342	3.62590	270	69.8198	83.8198	394.3292	30.5214	20.6429
2,4-dimetilhexano	1.408	0.1486	3.66390	258	70.4792	84.4792	380.7699	29.3947	19.8985
2,2,3 trimetilpentano	1.451	0.1617	3.48138	230	79.8189	93.8189	332.6014	36.2401	17.7574*
3,3-dimetilhexano	1.572	0.1965	3.62132	244	75.1491	89.1491	354.5538	27.9713	18.9193
2,3,4-trimetilpentano	1.725	0.2368	3.55342	236	75.6664	89.6664	343.6628	27.0804	18.5658*
2,3,3-trimetilpentano	1.832	0.2629	3.50404	226	80.1942	94.1942	327.3512	25.9226	17.9552*

Los encabezados de la Tabla 1 correspondientes a las columnas de 2-10, representan: tiempos de retención relativos al *n*-heptano, logaritmo del tiempo de retención relativo presentado en la columna 2, índice de conectividad químico, índice de H. P. Schultz, índices 2-MTI corregidos, su explicación se encuentra en el texto. Las columnas 9,10 son los índices basados en los valores propios tomados de la literatura o evaluados según se indica en el texto. (*) Implica, valores calculados por el autor del presente trabajo.

Tabla 2. Ecuaciones de regresión entre los logaritmos de los tiempos de retención relativos ($\log t_{rr}$) y los distintos índices topológicos en estudio.

Variables independientes	Ecuación de regresión	Factores	Valor	Error	r o R^2	F
$^1\chi_v$	$y = A + BX$	A	-3.35638	0.10314	0.9851	853.78
		B	0.98527	0.03372		
2-MTI[B] _{n × n}	$y = A + B_1X + B_2X^2$	A	-2.2603	0.08224	0.9729	448.74
		B ₁	0.01771	0.0012		
		B ₂	-3.246 E-5	3.933 E-6		
2-MTI[D'] _{n × n}	$y = A + BX$	A	-2.09994	0.16669	0.9028	114.59
		B	0.02958	0.00276		
2-MTI[D''] _{n × n}	$y = A + BX$	A	-2.19764	0.16585	0.9119	128.37
		B	0.02623	0.00232		
C _{g.d.A}	$y = A + B_1X + B_2X^2$	A	-2.218	0.0773	0.9748	471.64
		B ₁	0.01205	7.794 E-4		
		B ₂	-1.533 E-5	1.762 E-6		
$\lambda + ([A]_{n × n} + [B]_{n × n})$	$y = A + B_1X + B_2X^2$	A	-3.11113	0.16458	0.9697	401.17
		B ₁	0.27086	0.02694		
		B ₂	-0.00529	0.00103		
$\lambda + ([A]_{n × n} + [B^A]_{n × n})$	$y = A + B_1X + B_2X^2$	A	-2.97847	0.15257	0.9697	401.12
		B ₁	0.1748	0.01721		
		B ₂	-0.00229	4.53 E-4		

En la primera columna se indican los índices de L.B. Kier ($^1\chi_v$), H.P. Schultz (2-MTI) y los propuestos en este trabajo. Factores: r , R y F implican: factor, coeficiente de correlación, de determinación, índice estadístico de Fisher.

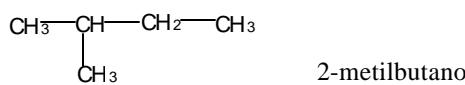
Tabla 3. Parámetros estadísticos de las correlaciones entre los tiempos de retención experimentales y calculados.

Índices topológicos	<i>r</i>	<i>F</i>
C_{g,d,A^o}	0.9359	183.54
2-MTI	0.9353	181.64
2-MTI $[D']_{n \times n}$	0.7492	33.26
2-MTI $[D'']_{n \times n}$	0.7764	39.46

respecto a su predecesor (2-MTI original) respecto al grado de correlación con la variable dependiente, en este caso el $\log t_{ri}$ para un modelo particular (hidrocarburos saturados), se hace necesario incluir en análisis futuros, otras proposiciones como modelos que contengan moléculas distintas a los hidrocarburos. Independientemente de lo anterior, se hace necesario además estudiar una nueva proposición consistente en el uso de la matriz de distancia cuadrada $[B^A]_{n \times n}$ en conjunto con el vector multiplicativo anterior modificado $V_{1 \times n}$. Una estructura nueva exitosa para este vector se propuso en un trabajo anterior [21].

Se debe considerar además que, hemos estudiado para ambos índices topológicos anteriormente citados su parte congruente, sin hacer un estudio de la parte no congruente, es decir, aquella zona que no correlaciona con el índice topológico 2-MTI. Esta zona no congruente puede producir, toda vez que el índice topológico C_{g,d,A^o} , participe en forma contributiva en una función de correlación multivariable, un grado de correlación significativo respecto a la variable dependiente. Un ejemplo de ello se describe y analiza en la literatura [22].

Evaluación del índice topológico de (2-MTI original):



$$[A]_{n \times n} = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{vmatrix} \quad [A]_{n \times n}^2 = \begin{vmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 3 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{vmatrix}$$

Los términos: a_{ii} de $[A]^2_{n \times n}$ son elementos del vector $V_{1 \times n} = [1\ 3\ 2\ 1\ 1]$.

$$[B]_{n \times n} = \begin{vmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{vmatrix} \quad \text{CH}_3-\underset{\text{CH}_3}{\text{CH}}-\text{CH}_2-\text{CH}_3$$

$$([A]_{n \times n} + [B]_{n \times n}) = \begin{vmatrix} 0 & 2 & 2 & 3 & 2 \\ 2 & 0 & 2 & 2 & 2 \\ 2 & 2 & 0 & 2 & 2 \\ 3 & 2 & 2 & 0 & 3 \\ 2 & 2 & 2 & 3 & 0 \end{vmatrix}$$

$$V_{1 \times n}([A]_{n \times n} + [B]_{n \times n}) = [15\ 10\ 12\ 16\ 15] = 68.$$

Luego el índice topológico de H. P. Schultz (2-MTI original) es 68.

Conclusiones

- El uso de la matriz de distancia $([B^A]_{n \times n})$ en el índice topológico 2-MTI, dando origen al índice topológico propuesto C_{g,d,A^o} , produce la mayor correlación respecto a todas las otras regresiones que se señalan en la Tabla 2, exceptuando al índice $^1\chi_V$, lo que indica que su incorporación en el índice 2-MTI original aumenta el grado de correlación entre las variables implicadas. Además, al tener un valor de (*F*) mayor, las correlaciones donde participa como variable independiente el índice propuesto C_{g,d,A^o} , implica que esas expresiones de correlación poseen un poder predictivo mayor.
- El uso de matrices de distancia en los índices que constituyen las variables independientes: $\lambda^+([A]_{n \times n} + [B]_{n \times n})$ y $\lambda^+([A]_{n \times n} + [B^A]_{n \times n})$, cuyos elementos respectivamente sean distancias topológicas representadas por el campo de los naturales o distancias reales entre los átomos de carbono, producen expresiones de regresión estadísticamente equivalentes. Lo indicado anteriormente, sugiere que sería válido integrar las modificaciones estudiadas en este trabajo con la modificación ya estudiada sobre el vector anterior multiplicativo [20] del índice 2-MTI original, con el propósito de evaluar ambas contribuciones en un nuevo índice.
- Mediante el análisis realizado en este trabajo se valida para el modelo 2-MTI original el uso de la matriz $[B]_{n \times n}$ constituida por elementos pertenecientes a los números naturales.

Referencias

- Randic, M.; Kleisner, A. F.; De Alba, L. M. *J. Chem. Inf. Comput. Sci.* **1994**, 34, 277-280.
- Kaliszan, R. *Anal. Chem.* **1992**, 64, 619A-631A.
- Schultz, H. P. *J. Chem. Inf. Comput. Sci.* **1989**, 29, 227-228.
- Schultz, H. P.; Schultz, E. B.; Schultz, T.P. *J. Chem. Inf. Comput. Sci.* **1990**, 30, 27-29.
- Schultz, H. P., Schultz, T.P. *J. Chem. Inf. Comput. Sci.* **1993**, 33, 240-244.
- Yang, Y.; Xu, L.; Hu, C. *J. Chem. Inf. Comput. Sci.* **1994**, 34, 1140-1145.
- Gálvez, J.; García, R.; Salabert, M.T., Soler, R. *J. Chem. Inf. Comput. Sci.* **1994**, 34, 520-525.
- Schultz, H. P.; Schultz, T. P. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 853-857.
- Randic, M. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 672-687.

10. Morgan, E. *Chemometric Experimental Designs*. John Wiley and Sons, London, **1991**.
11. Masart, D. L.; Vandeiginste, B. G. M., Deming, S. N. N.; Michotte, Y.; Kaufman, L. *Chemometrics A Textbook*. Elsevier, Amsterdam, **1988**.
12. Graham, R. C. *Data Analysis for the Chemical Sciences. A Guide to Statistical Techniques*. VCH Publisher, Inc. New York, **1993**.
13. Kateman, G.; Buydens, L. *Quality Control in Analytical Chemistry*, Vol. 60. John Wiley and Sons, Inc., New York, **1993**.
14. Katrizky, A. R., Gordeeva, E.V. *J. Chem. Inf. Comput. Sci.* **1993**, 33, 835-857.
15. Desty, D. H.; Goldup, A.; Swanton, W. T. *Gas Chromatography*. Academic Press, **1962**.
16. Sabljic, A.; Horvatic, D. *J. Chem. Inf. Comput. Sci.* **1993**, 33, 292-295.
17. Mihalic, Z.; Nikolic, S.; Trinajstic, N. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 28-37.
18. Hyperchem, Copyright 1993, Hypercube, Inc. and Autodesk, Inc. Developed by Hypercube, Inc.
19. Nicrocal Origin 4.00, Copyright 1991-1995. Microcal Software, Inc.
20. Muranaka, K. *J. Chem. Educ.* **2001**, 78, 1390-1393.
21. Randic, M. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 672-687.
22. Cornwell, E. *Bol. Soc. Chil. Quim.* **2000**, 45, 645-647.