

HUBUNGAN KUANTITATIF STRUKTUR DAN INDEKS BIAS DARI SENYAWA ORGANIK BERDASARKAN DESKRIPTOR MOLEKULAR

Iqmal Tahir, Karna Wijaya dan Nelly Nuroniah

Pusat Kimia Komputasi Indonesia Austria, Jurusan Kimia Fakultas MIPA UGM Jogjakarta

INTISARI

Telah dilakukan analisis Hubungan Kuantitatif Struktur–Sifat fisika (HKSS) dari senyawa organik menggunakan deskriptor molekular. Deskriptor dari setiap senyawa dijabarkan dalam bentuk jumlah dan atom serta jumlah dan jenis ikatan kimia senyawa. Indeks bias (n_D^{20}) dipelajari sebagai fungsi linear dari seluruh deskriptor dengan menggunakan analisis regresi multilinear metode backward.

Model alternatif persamaan HKSS terbaik untuk indeks bias :

$$\begin{aligned} n_D^{20} = & 1,375 - 5,804 \cdot 10^{-3} \cdot n(\mathbf{C}_{\text{prim alf}}) + 1,904 \cdot 10^{-3} \cdot n(\mathbf{C}_{\text{sek alf}}) - 4,697 \cdot 10^{-3} \cdot n(\mathbf{C}_{\text{ter alf}}) \\ & - 1,061 \cdot 10^{-2} \cdot n(\mathbf{C}_{\text{kwar alf}}) + 2,209 \cdot 10^{-2} \cdot n(\mathbf{C}_{\text{sek arm}}) + 1,675 \cdot 10^{-2} \cdot n(\mathbf{O}) \\ & + 4,708 \cdot 10^{-2} \cdot n(\mathbf{Cl}) + 9,762 \cdot 10^{-3} \cdot n(\mathbf{C-C}_{\text{sikl}}) + 1,117 \cdot 10^{-2} \cdot n(\mathbf{C}_2=\mathbf{C}_3) \\ & + 1,045 \cdot 10^{-2} \cdot n(\mathbf{C}_3=\mathbf{C}_4) + 1,536 \cdot 10^{-2} \cdot n(\mathbf{C}_4=\mathbf{C}_5) + 1,483 \cdot 10^{-2} \cdot n(\mathbf{C}_1=\mathbf{C}_2) \\ & + 2,316 \cdot 10^{-2} \cdot n(\mathbf{C}_2=\mathbf{C}_3) + 2,402 \cdot 10^{-2} \cdot n(\mathbf{C}_3=\mathbf{C}_4) + 1,983 \cdot 10^{-2} \cdot n(\mathbf{C}_4=\mathbf{C}_5) \\ & + 2,453 \cdot 10^{-2} \cdot n(\mathbf{C}_5=\mathbf{C}_6) + 2,748 \cdot 10^{-2} \cdot n(\mathbf{C}_6=\mathbf{C}_7) - 2,075 \cdot 10^{-2} \cdot n(\mathbf{C-O}) \\ & - 3,606 \cdot 10^{-2} \cdot n(\mathbf{C-F}) - 3,847 \cdot 10^{-2} \cdot n(\mathbf{C-Cl}) + 4,057 \cdot 10^{-2} \cdot n(\mathbf{C-Br}) \\ & - 7,240 \cdot 10^{-2} \cdot n(\mathbf{C=O}) + 2,619 \cdot 10^{-2} \cdot n(\mathbf{N-H}) + 4,389 \cdot 10^{-2} \cdot n(\mathbf{S-H}) \\ & + 2,503 \cdot 10^{-2} \cdot n(\mathbf{ortho}) + 1,861 \cdot 10^{-2} \cdot n(\mathbf{meta}) - 2,271 \cdot 10^{-6} \cdot n(\mathbf{HxBM}) \\ & + 1,287 \cdot 10^{-4} \cdot n(\mathbf{C}_{\text{prim}} \times \mathbf{BM}) \end{aligned}$$

dengan parameter statistik : $n = 340$; $r^2 = 0,956$; $SD = 1,020 \cdot 10^{-2}$; $F_{\text{hitung}} = 243,306$.

Kata kunci : HKSS, analisis regresi multilinear, indeks bias

QUANTITATIVE STRUCTURE AND REFRACTIVE INDEX RELATIONSHIPS OF ORGANIC COMPOUNDS USING FRAGMENTAL DESCRIPTOR

ABSTRACT

The Quantitative Structure-Property Relationships (QSPR) of organic compounds using fragmental descriptor has been studied. Fragmental description of every compound is explained in the form of atom and it's bonding types. The study was made in the form of refractive index, as linear function of all fragmental descriptor. The best equation model was determined by regression multilinear analysis. The alternative model of the best QSPR equation for refractive index properties showed that refractive index linearly correlated and is expressed by:

$$\begin{aligned} n_D^{20} = & 1,375 - 5,804 \times 10^{-3} \cdot n(\mathbf{C}_{\text{prim alf}}) + 1,904 \times 10^{-3} \cdot n(\mathbf{C}_{\text{sek alf}}) - 4,697 \times 10^{-3} \cdot n(\mathbf{C}_{\text{ter alf}}) \\ & - 1,061 \times 10^{-2} \cdot n(\mathbf{C}_{\text{kwar alf}}) + 2,209 \times 10^{-2} \cdot n(\mathbf{C}_{\text{sek arm}}) + 1,675 \times 10^{-2} \cdot n(\mathbf{O}) \\ & + 4,708 \times 10^{-2} \cdot n(\mathbf{Cl}) + 9,762 \times 10^{-3} \cdot n(\mathbf{C-C}_{\text{sikl}}) + 1,117 \times 10^{-2} \cdot n(\mathbf{C}_2=\mathbf{C}_3) \\ & + 1,045 \times 10^{-2} \cdot n(\mathbf{C}_3=\mathbf{C}_4) + 1,536 \times 10^{-2} \cdot n(\mathbf{C}_4=\mathbf{C}_5) + 1,483 \times 10^{-2} \cdot n(\mathbf{C}_1=\mathbf{C}_2) \\ & + 2,316 \times 10^{-2} \cdot n(\mathbf{C}_2=\mathbf{C}_3) + 2,402 \times 10^{-2} \cdot n(\mathbf{C}_3=\mathbf{C}_4) + 1,983 \times 10^{-2} \cdot n(\mathbf{C}_4=\mathbf{C}_5) \\ & + 2,453 \times 10^{-2} \cdot n(\mathbf{C}_5=\mathbf{C}_6) + 2,748 \times 10^{-2} \cdot n(\mathbf{C}_6=\mathbf{C}_7) - 2,075 \times 10^{-2} \cdot n(\mathbf{C-O}) \\ & - 3,606 \times 10^{-2} \cdot n(\mathbf{C-F}) - 3,847 \times 10^{-2} \cdot n(\mathbf{C-Cl}) + 4,057 \times 10^{-2} \cdot n(\mathbf{C-Br}) \\ & - 7,240 \times 10^{-2} \cdot n(\mathbf{C=O}) + 2,619 \times 10^{-2} \cdot n(\mathbf{N-H}) + 4,389 \times 10^{-2} \cdot n(\mathbf{S-H}) \\ & + 2,503 \times 10^{-2} \cdot n(\mathbf{ortho}) + 1,861 \times 10^{-2} \cdot n(\mathbf{meta}) - 2,271 \times 10^{-6} \cdot n(\mathbf{HxBM}) \\ & + 1,287 \times 10^{-4} \cdot n(\mathbf{C}_{\text{prim}} \times \mathbf{BM}) \end{aligned}$$

with statistical parameters: $n = 340$, $r^2 = 0,956$, $SD = 1,020 \times 10^{-2}$, $F_{\text{calc}} = 243,306$

Keywords : QSPR, refractive index, multiple linear regression

PENDAHULUAN

Rumus struktur diketahui dapat menetapkan gambaran sifat dan aktivitas senyawa secara tidak langsung, seperti kecepatan oksidasi senyawa, konstanta kesetimbangan, kecepatan adsorpsi pada beberapa permukaan, dan juga untuk sifat-sifat fisika antara lain titik didih, titik leleh, indeks bias, kelarutan, kerapatan dan lain-lain (Katritzky dan Labanov, 1995). Pada prinsipnya rumus struktur molekul dari setiap senyawa kimia mengandung gambaran yang mencerminkan informasi tentang sifat kimia, sifat fisika atau aktivitas biologi dari senyawa tersebut (Kubinyi, 1993). Untuk dapat memprediksi antara gambaran sifat dan aktivitas molekul tersebut terdapat dua pendekatan teoritik yang banyak digunakan oleh para ahli kimia sekarang. Alternatif pertama berupa suatu perhitungan teoritik yang berdasarkan mekanika kuantum dan mekanika statistik (Leach, 1996). Perhitungan dengan berdasarkan persamaan Schrödinger akan dapat menghasilkan gambaran keadaan elektronik dari seluruh faktor molekul dalam berbagai keadaan. Saat ini telah banyak sumbangan kemajuan berupa aplikasi dari penggunaan metoda tersebut untuk berbagai sistem kimia baik yang sederhana maupun yang kompleks. Gambaran sifat dan aktivitas tersebut dipelajari berdasarkan data-data yang terkait pada struktur geometri teroptimasi dan struktur elektronik yang diperoleh. Alternatif kedua adalah suatu metode yang disebut dengan *Quantitative Structure-Activity Relationships/Quantitative Structure-Property Relationships* (QSAR/QSPR) (Kubinyi, 1993). Analisis ini didasarkan pada hubungan matematika antara aktivitas atau sifat dengan satu atau lebih parameter deskriptif (deskriptor) yang merepresentasikan struktur molekul tersebut. Dengan deskriptor-deskriptor yang didapatkan dari sifat eksperimental senyawa, maka akan dapat diperoleh persamaan QSAR/QSPR yang diturunkan untuk memprediksi aktivitas atau sifat dari senyawa yang tersedia. Apabila sebuah model dapat diajukan, maka model tersebut dapat dipakai sebagai pedoman sintesis kimia dalam pilihan alternatif struktur hipotesis.

Secara umum analisis QSAR/QSPR dapat menjelaskan sifat yang dihubungkan dengan struktur molekul (Katritzky dan Labanov, 1995). Estimasi sifat fisik senyawa seperti titik didih, titik leleh, sifat lipofilitas, indeks bias, kelarutan dan kerapatan telah banyak dilaporkan berupa analisis dengan menggunakan deskriptor geometri, parameter sterik atau struktur elektronik dari molekul. Salah satu deskriptor yang mudah dijabarkan dari struktur molekul adalah deskriptor struktural berupa deskriptor atomik dan fragmental dari senyawa. Deskriptor ini bisa dijabarkan secara uraian analitik biasa yakni dengan melihat jenis dan jumlah atom, jenis dan jumlah ikatan, berat molekul dan lain-lain. Penggunaan deskriptor ini telah digunakan oleh Carlton (1998) untuk analisis QSPR titik didih dari 52 senyawa turunan senyawa kloroflouroetana. Duchowicz dan Castro (2000) juga telah menggunakan teknik yang sama untuk perhitungan nilai koefisien partisi senyawa dalam dua pelarut tak bercampur (oktanol dan air) dalam bentuk parameter log P.

Salah satu sifat fisik yang sering digunakan oleh para ahli kimia untuk karakterisasi senyawa adalah parameter indeks bias. Indeks bias (n) adalah perbandingan kecepatan rambatan cahaya di ruang hampa (c) dengan kecepatan cahaya dalam medium yang diamati (v), dimana harga n tidak pernah lebih kecil dari 1. Nilai indeks bias dihitung dari persamaan berikut :

$$n = \frac{c}{v}$$

Reinhard dan Drefahl (1999) mengemukakan bahwa indeks bias dapat digunakan untuk:

1. Ukuran kemurnian suatu senyawa
2. Estimasi titik didih
3. Estimasi viskositas larutan
4. Menghitung kepolaran molekul elektronik

Analisis QSPR tentang indeks bias berbagai senyawa organik dengan menggunakan deskriptor struktural masih perlu dilakukan. Indeks bias suatu senyawa dipengaruhi oleh beberapa faktor meliputi: ukuran molekul, jenis isomer dan ikatan van der Waals. Kebanyakan perlakuan teoritis untuk memprediksi indeks bias telah dikembangkan sampai indeks molar, dengan mengukur kekuatan intrinsik bias dari unit dasar struktur material. Berdasarkan asumsi ini maka diduga deskriptor molekular akan berpengaruh dan memiliki korelasi dengan nilai indeks bias suatu senyawa. Tujuan dari penulisan penelitian ini adalah untuk mengkaji korelasi sifat fisika senyawa berupa indeks bias dari berbagai senyawa organik dengan menggunakan deskriptor molekular.

METODA

Peralatan

Peralatan yang digunakan meliputi perangkat lunak statistik berupa program *SPSS for Windows release 10.0*. Perangkat keras yang digunakan berupa 1 unit komputer dengan prosesor Intel Pentium 4 dengan kecepatan processor 1,8 GHz dan RAM 128 MB.

Bahan

Data sifat fisik senyawa organik yang digunakan dalam penelitian ini sebanyak 340 senyawa, meliputi golongan-golongan alkana alifatik, alkana siklis, alkena, alkuna, ester, benzena, alkohol, eter, aldehid, keton, asam karboksilat, ester, anilin, halobenzena dan beberapa senyawa lain.

Prosedur Penelitian

a. Analisis deskriptor molekular

Data struktur geometri bisa diperoleh dari analisis struktur kimia dari setiap senyawa, sedangkan data-data indeks bias eksperimen dikompilasi dari literatur (Lide, 2000). Data indeks bias tersebut dikompilasi berupa data indeks bias pada kondisi pengukuran yang sama

yakni dengan pengukuran menggunakan lampu Deuterium dan pada temperatur 20°C. Hal ini dilakukan untuk memperoleh data komparatif sehingga diharapkan dalam proses estimasi akan mendapatkan kesahihan yang tinggi. Selanjutnya indeks bias ini dituliskan sebagai n_D^{20} eksperimen. Data struktur molekul yang digunakan sebagai prediktor berupa analisis jumlah atom, ikatan, berat molekul dan beberapa kombinasi antar variabel. Jenis dan keterangan tiap variabel yang digunakan pada analisis ini disajikan pada tabel 1.

b. Analisis Persamaan QSPR

1. Data disajikan dalam tabel meliputi data indeks bias (n_D^{20}) senyawa organik sebagai variabel tak bebas dan semua deskriptor struktural (tabel 1) sebagai variabel bebas.
2. Dihitung korelasi semua variabel bebas dan dikaitkan dengan harga variabel tak bebas. Perhitungan menggunakan analisis regresi linear program SPSS metode *backward*, hasil perhitungan digunakan untuk menunjukkan urutan variabel bebas penting yang berfungsi sebagai deskriptor.
3. Pemilihan model hubungan QSPR dilakukan secara intensif berdasarkan kriteria nilai korelasi dan beberapa kemungkinan yang tepat.
4. Setiap alternatif model dilakukan evaluasi meliputi dengan parameter statistik r^2 , SD, dan F. Analisis dilakukan pada tingkat kepercayaan 95 %.
5. Persamaan QSPR diperoleh dalam bentuk:

$$Y = \sum_{i=1}^n k_i \cdot x_i + C$$

dengan Y adalah variabel tidak bebas yakni data n_D^{20} , k_i adalah konstanta yang dicari dari *fitting* data, x_i adalah variabel bebas dari tabel 1, n adalah banyaknya data dan C adalah konstanta (Kubinyi, 1993).

Tabel 1. Lambang dan arti deskriptor struktural yang digunakan pada estimasi kerapatan

No.	Deskriptor	Arti deskriptor	No.	Deskriptor	Arti deskriptor
1	n(H)	Jumlah atom H	28	n(C₆≡C₇)	Jumlah ikatan rangkap tiga C ₆ ≡C ₇
2	n(C_{prim})	Jumlah atom C primer	29	n(C-O)	Jumlah ikatan C-O
3	n(C_{sek alf})	Jumlah atom C sekunder alifatik	30	n(C-N)	Jumlah ikatan C-N
4	n(C_{ter alf})	Jumlah atom C tersier alifatik	31	n(C-S)	Jumlah ikatan C-S
5	n(C_{kwar})	Jumlah atom C kuartar	32	n(C-F)	Jumlah ikatan C-F
6	n(C_{sek sik})	Jumlah atom C sekunder siklik	33	n(C-Cl)	Jumlah ikatan C-Cl
7	n(C_{ter sik})	Jumlah atom C tersier siklik	34	n(C-Br)	Jumlah ikatan C-Br
8	n(C_{sek arm})	Jumlah atom C sekunder aromatik	35	n(C-I)	Jumlah ikatan C-I
9	n(C_{ter arm})	Jumlah atom C tersier aromatik	36	n(C=O)	Jumlah ikatan C=O

10	n(O)	Jumlah atom O	37	n(O-H)	Jumlah ikatan O-H
11	n(N)	Jumlah atom N	38	n(N-H)	Jumlah ikatan N-H
12	n(F)	Jumlah atom F	39	n(S-H)	Jumlah ikatan S-H
13	n(Cl)	Jumlah atom Cl	40	n(E)	Jumlah senyawa berstruktur E
14	n(C-H)	Jumlah ikatan C-H	41	n(Z)	Jumlah senyawa berstruktur Z
15	n(C-C_{alf})	Jumlah ikatan C-C alifatik	42	n(ortho)	Jumlah gugus yang terikat pada struktur <i>ortho</i>
16	n(C-C_{sik})	Jumlah ikatan C-C siklik	43	n(meta)	Jumlah gugus yang terikat pada struktur <i>meta</i>
17	n(C=C_{arm})	Jumlah ikatan rangkap dua C=C pada cincin aromatik	44	n(para)	Jumlah gugus yang terikat pada struktur <i>para</i>
18	n(C₁=C₂)	Jumlah ikatan rangkap dua C ₁ =C ₂	45	BM	berat molekul
19	n(C₂=C₃)	Jumlah ikatan rangkap dua C ₂ =C ₃	46	n(H)²	Jumlah atom H kuadrat
20	n(C₃=C₄)	Jumlah ikatan rangkap dua C ₃ =C ₄	47	(BM)²	Berat molekul kuadrat
21	n(C₄=C₅)	Jumlah ikatan rangkap dua C ₄ =C ₅	48	n(C_{prim})²	Jumlah atom C sekunder kuadrat
22	n(C₅=C₆)	Jumlah ikatan rangkap dua C ₅ =C ₆	49	n(C_{sek})²	Jumlah atom C primer kuadrat
23	n(C₁≡C₂)	Jumlah ikatan rangkap tiga C ₁ ≡C ₂	50	n(C-C_{alf})²	Jumlah ikatan C-C alifatik kuadrat
24	n(C₂≡C₃)	Jumlah ikatan rangkap tiga C ₂ ≡C ₃	51	n(C-H)²	Jumlah ikatan C-H kuadrat
25	n(C₃≡C₄)	Jumlah ikatan rangkap tiga C ₃ ≡C ₄	52	n(H x BM)	Jumlah atom H kali berat molekul
26	n(C₄≡C₅)	Jumlah ikatan rangkap tiga C ₄ ≡C ₅	53	n(C_{prim} x BM)	Jumlah atom C primer kali berat molekul
27	n(C₅≡C₆)	Jumlah ikatan rangkap tiga C ₅ ≡C ₆			

6. Selanjutnya dengan menggunakan persamaan QSPR terpilih digunakan untuk menghitung indeks bias teoritis (n_D^{20} prediksi) dan ditentukan parameter baru PRESS (*Prediction of Sum of Square*). PRESS dihitung dengan menggunakan :

$$\text{PRESS} = \sum_{i=1}^n (Y_{pred} - Y_{obs})^2$$

HASIL DAN PEMBAHASAN

Evaluasi persamaan QSPR

Penelitian ini menggunakan total 340 senyawa organik yang bervariasi terdiri dari senyawa-senyawa hidrokarbon, senyawa aromatik dan senyawa organik lain yang mengandung F, Cl, Br, I,

S, O, N, senyawa berstruktur E dan Z, serta senyawa yang mengandung gugus yang terikat pada *ortho*, *meta* dan *para*.

Analisis regresi multilinear antara deskriptor molekular dengan indeks bias memberikan 13 model terbaik berdasarkan metoda *backward*. Hasil ringkasan parameter statistik seluruh model secara lengkap diberikan pada tabel 2.

Tabel 2. Parameter statistik indeks bias hasil regresi multilinear

Model	n	m	r	r ²	SD	F _{hitung}	F _{tabel}	F _{hit/tab}
1	340	40	0,979	0,958	1,023.10 ⁻²	169,499	1,435	118,117
2	340	39	0,979	0,958	1,022.10 ⁻²	174,269	1,439	121,025
3	340	38	0,979	0,958	1,020.10 ⁻²	179,388	1,445	124,136
4	340	37	0,979	0,958	1,019.10 ⁻²	184,589	1,450	127,261
5	340	36	0,979	0,958	1,018.10 ⁻²	190,253	1,456	130,658
6	340	35	0,979	0,958	1,017.10 ⁻²	196,058	1,462	134,102
7	340	34	0,979	0,958	1,016.10 ⁻²	202,306	1,468	137,793
8	340	33	0,979	0,957	1,015.10 ⁻²	208,787	1,474	141,582
9	340	32	0,978	0,957	1,013.10 ⁻²	215,880	1,481	145,719
10	340	31	0,978	0,957	1,014.10 ⁻²	222,624	1,488	149,547
11	340	30	0,978	0,957	1,015.10 ⁻²	229,540	1,496	153,413
12	340	29	0,978	0,957	1,017.10 ⁻²	236,223	1,504	157,042
13	340	28	0,978	0,956	1,020.10 ⁻²	243,306	1,512	160,847

Keterangan : n = jumlah data, m = jumlah variabel bebas.

Jika dilihat pada beberapa parameter statistik seperti yang disajikan pada tabel 2 maka parameter yang dapat untuk membedakan adalah parameter rasio F_{hit}/F_{tab} saja. Untuk parameter koefisien determinasi (r^2) dari model 1 sampai model 13 relatif tidak begitu berbeda yakni berkisar antara 0,956 sampai 0,958. Koefisien determinasi ini merupakan ukuran linearistik dari model persamaan QSPR yang dihasilkan. Hal tersebut menunjukkan pengertian bahwa seluruh model relatif memiliki korelasi yang cukup kuat antara indeks bias dan deskriptor molekular. Untuk parameter SD dari seluruh model juga relatif tidak begitu berbeda yakni berkisar antara $1,013.10^{-2}$ sampai $1,023.10^{-2}$. Parameter SD ini menunjukkan tingkat penyimpangan model yang dihasilkan. Parameter SD ini relatif lebih kecil dibandingkan dengan parameter SD untuk seluruh variabel tidak bebas yang digunakan.

Dari nilai F pada tabel 2 tersebut terlihat bahwa seluruh model relatif memenuhi syarat signifikansi pada tingkat kepercayaan 95 %. Hal ini ditunjukkan dengan rasio F_{hit}/F_{tab} yang lebih besar dari 1. Data tersebut menunjukkan bahwa seluruh model cukup kuat untuk membuktikan korelasi antara nilai indeks bias dengan deskriptor molekular yang diteliti. Nilai rasio F_{hit}/F_{tab} tersebut seluruhnya lebih dari 100 dan tertinggi adalah untuk model 13 yakni 160,847. Dari uraian ini dapat diambil suatu kesimpulan bahwa model ke 13 dapat digunakan sebagai persamaan QSPR untuk menyatakan hubungan indeks bias dan deskriptor molekular.

Tabel 3 Koefisien deskriptor pada persamaan estimasi indeks bias

No	Deskriptor	Koefisien	Standar deviasi
	Intersep	1,375	0,006
1	$n(C_{prim\ alf})$	$-5,804 \cdot 10^{-3}$	0,003
2	$n(C_{sek\ alf})$	$1,904 \cdot 10^{-3}$	0,000
3	$n(C_{ter\ alf})$	$-4,697 \cdot 10^{-3}$	0,002
4	$n(C_{kwar\ alf})$	$-1,061 \cdot 10^{-2}$	0,004
5	$n(C_{sek\ arm})$	$2,209 \cdot 10^{-2}$	0,001
6	$n(O)$	$1,675 \cdot 10^{-2}$	0,003
7	$n(Cl)$	$4,708 \cdot 10^{-2}$	0,007
8	$n(C-C_{sikl})$	$9,762 \cdot 10^{-3}$	0,001
9	$n(C_2=C_3)$	$1,117 \cdot 10^{-2}$	0,004
10	$n(C_3=C_4)$	$1,045 \cdot 10^{-2}$	0,004
11	$n(C_4=C_5)$	$1,536 \cdot 10^{-2}$	0,005
12	$n(C_1=C_2)$	$1,483 \cdot 10^{-2}$	0,003
13	$n(C_2=C_3)$	$2,316 \cdot 10^{-2}$	0,005
14	$n(C_3=C_4)$	$2,402 \cdot 10^{-2}$	0,005
15	$n(C_4=C_5)$	$1,983 \cdot 10^{-2}$	0,006
16	$n(C_5=C_6)$	$2,453 \cdot 10^{-2}$	0,007
17	$n(C_6=C_7)$	$2,748 \cdot 10^{-2}$	0,010
18	$n(C-O)$	$-2,075 \cdot 10^{-2}$	0,003
19	$n(C-F)$	$-3,606 \cdot 10^{-2}$	0,002
20	$n(C-Cl)$	$-3,847 \cdot 10^{-2}$	0,006
21	$n(C-Br)$	$4,057 \cdot 10^{-2}$	0,003
21	$n(C=O)$	$-7,240 \cdot 10^{-2}$	0,004
23	$n(N-H)$	$2,619 \cdot 10^{-2}$	0,002
24	$n(S-H)$	$4,389 \cdot 10^{-2}$	0,005
25	$n(Ortho)$	$2,503 \cdot 10^{-2}$	0,003
26	$n(Meta)$	$1,861 \cdot 10^{-2}$	0,003
27	$n(H \times BM)$	$-2,271 \cdot 10^{-6}$	0,000
28	$n(C_{prim} \times BM)$	$1,287 \cdot 10^{-4}$	0,000

Model ke 13 tersebut melibatkan 28 deskriptor molekular yang berpengaruh. Deskriptor tersebut secara lengkap beserta nilai parameter-parameter yang terkait disajikan pada tabel 3. Pada tabel 3 tersebut juga disajikan nilai standar deviasi dari setiap parameter. Untuk tiap parameter seperti yang tercantum pada tabel 3 relatif memiliki simpangan yang relatif kecil.

Bentuk persamaan dengan menggunakan parameter seperti yang tercantum dalam tabel 3 tersebut dapat dituliskan dalam bentuk persamaan regresi sebagai berikut :

$$\begin{aligned}
 n_D^{20} = & 1,375 - 5,804 \cdot 10^{-3} \cdot n(C_{prim\ alf}) + 1,904 \cdot 10^{-3} \cdot n(C_{sek\ alf}) - 4,697 \cdot 10^{-3} \cdot n(C_{ter\ alf}) \\
 & - 1,061 \cdot 10^{-2} \cdot n(C_{kwar\ alf}) + 2,209 \cdot 10^{-2} \cdot n(C_{sek\ arm}) + 1,675 \cdot 10^{-2} \cdot n(O) \\
 & + 4,708 \cdot 10^{-2} \cdot n(Cl) + 9,762 \cdot 10^{-3} \cdot n(C-C_{sikl}) + 1,117 \cdot 10^{-2} \cdot n(C_2=C_3) \\
 & + 1,045 \cdot 10^{-2} \cdot n(C_3=C_4) + 1,536 \cdot 10^{-2} \cdot n(C_4=C_5) + 1,483 \cdot 10^{-2} \cdot n(C_1=C_2) \\
 & + 2,316 \cdot 10^{-2} \cdot n(C_2=C_3) + 2,402 \cdot 10^{-2} \cdot n(C_3=C_4) + 1,983 \cdot 10^{-2} \cdot n(C_4=C_5) \\
 & + 2,453 \cdot 10^{-2} \cdot n(C_5=C_6) + 2,748 \cdot 10^{-2} \cdot n(C_6=C_7) - 2,075 \cdot 10^{-2} \cdot n(C-O) \\
 & - 3,606 \cdot 10^{-2} \cdot n(C-F) - 3,847 \cdot 10^{-2} \cdot n(C-Cl) + 4,057 \cdot 10^{-2} \cdot n(C-Br) \\
 & - 7,240 \cdot 10^{-2} \cdot n(C=O) + 2,619 \cdot 10^{-2} \cdot n(N-H) + 4,389 \cdot 10^{-2} \cdot n(S-H) \\
 & + 2,503 \cdot 10^{-2} \cdot n(ortho) + 1,861 \cdot 10^{-2} \cdot n(meta) - 2,271 \cdot 10^{-6} \cdot n(H \times BM)
 \end{aligned}$$

$$+ 1,287 \cdot 10^{-4} \cdot n(C_{prim} \times BM)$$

Parameter statistik yang terkait dengan model 13 adalah:

$$n = 340; \quad r = 0,956; \quad SD = 1,020 \cdot 10^{-2}; \quad F_{hitung} = 243,306; \quad F_{hitung}/F_{tabel} = 160,847$$

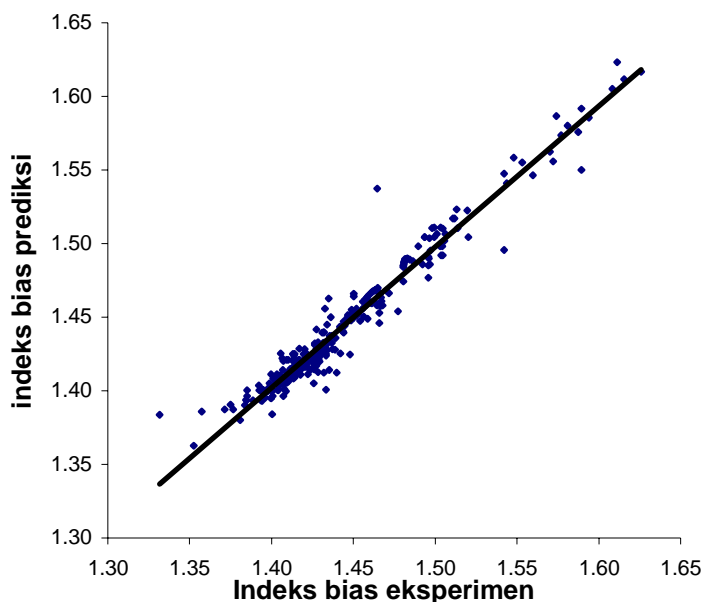
Uji validasi persamaan QSPR

Dari persamaan QSPR yang telah diperoleh tersebut selanjutnya harus dilakukan pengujian atau uji validasi. Teknik pengujian yang dilakukan adalah secara internal yakni dengan menggunakan data uji berupa senyawa asal sejumlah 340 senyawa. Data deskriptor molekular dimasukkan ke dalam persamaan QSPR dan dilakukan perhitungan dengan masing-masing parameter terkait, kemudian hasil akhirnya menunjukkan nilai $(n_D^{20})_{prediksi}$. Hasil perhitungan $(n_D^{20})_{prediksi}$ untuk 340 senyawa tersebut selanjutnya dikomparasikan dengan nilai $(n_D^{20})_{eksperimen}$. Komparasi ini secara grafik disajikan pada gambar 1.

Garis pada gambar 1 menunjukkan grafik antara indeks bias $(n_D^{20})_{eksperimen}$ dengan indeks bias prediksi $(n_D^{20})_{prediksi}$, ditunjukkan hasil alternatif yang ideal dimana $y = x$ atau $(n_D^{20})_{eksperimen} = (n_D^{20})_{prediksi}$. Garis imajiner untuk data yang diperoleh seharusnya mengikuti persamaan :

$$(n_D^{20})_{prediksi} = 0,0628 + 0,9566 (n_D^{20})_{eksperimen}$$

Pemodelan hasil analisis regresi multilinear memberikan estimasi untuk mengetahui besarnya indeks bias suatu senyawa yang belum diketahui indeks biasnya dari suatu basis data. Akurasi tingkat prediksi indeks bias tersebut relatif cukup baik terutama jika senyawa yang akan dihitung diperkirakan memiliki indeks bias antara 1,35 sampai 1,60. Di luar kisaran tersebut diduga akan dihasilkan tingkat prediksi indeks bias yang mengalami penyimpangan.



Gambar 1. Grafik hubungan antara $(n_D^{20})_{prediksi}$ dengan $(n_D^{20})_{eksperimen}$

Terkait dengan uji validasi maka dibahas analisis populasi data dan nilai PRESS dari model persamaan QSPR. Untuk analisis populasi tersebut digunakan perbandingan antara data eksperimen dan data prediksi meliputi nilai indeks bias tertinggi dan terendah serta nilai indeks bias rata-rata. Data yang dihasilkan relatif tidak menghasilkan kesalahan yang berarti. Semakin kecil nilai PRESS, maka semakin baik hubungan yang terjadi. Di antara 13 model persamaan, model no 13 memiliki harga PRESS yang paling minimal yaitu 0,0467. Hal ini menunjukkan bahwa nilai indeks bias prediksi mempunyai selisih yang kecil dengan nilai indeks bias eksperimen.

Tabel 4 Analisis populasi indeks bias

	Indeks bias, n_D^{20}	
	Eksperimen	Prediksi
n_D^{20} tertinggi	1,6260	1,6232
n_D^{20} rata-rata	1,4408	1,4411
n_D^{20} terendah	1,3316	1,3625

Analisis Molekular Berdasarkan Deskriptor Berpengaruh

Persamaan QSPR model no 13 merupakan persamaan hubungan kuantitatif antara indeks bias dan deskriptor kimia senyawa organik. Dari persamaan tersebut dikaji deskriptor-deskriptor yang berpengaruh terhadap indeks bias yang terdapat pada tabel 2. Secara lengkap deskriptor molekular tersebut dapat digolongkan sebagai berikut :

a. Atom karbon

Deskriptor $n(C_{prim\ alif})$ dengan nilai $-5,804 \cdot 10^{-3}$ yang berarti berpengaruh negatif terhadap nilai indeks bias atau menyebabkan penurunan nilai indeks bias. Secara tidak langsung nilai deskriptor ini juga merujuk pada faktor percabangan untuk senyawa yang memiliki deskriptor ini lebih dari 2. Untuk senyawa alifatik tak bercabang maka deskriptor ini bernilai tidak lebih dari 2, sedangkan untuk senyawa bercabang maka deskriptor akan bernilai lebih dari 2. Dengan demikian percabangan akan menyebabkan indeks bias berkurang.

Nilai deskriptor $n(C_{sek\ alif})$ sebesar $1,904 \cdot 10^{-3}$ menunjukkan arti bahwa semakin panjang rantai atom C pada alkana alifatik maka nilai indeks bias akan semakin tinggi. Hal ini diperkirakan karena struktur semakin panjang maka kerapatan senyawa semakin besar. Dengan demikian cahaya yang dibiaskan akan semakin besar. Demikian juga untuk faktor atom sekunder pada senyawa aromatik yang ditunjukkan dengan deskriptor $n(C_{sek\ arm})$ dengan nilai parameter sebesar $2,209 \cdot 10^{-2}$.

Adanya atom karbon tersier pada senyawa alifatik (parameter $n(C_{ter\ alif})$) dan atom karbon kwarternar alifatik (parameter $n(C_{kwar\ alif})$) memberi pengaruh negatif dengan nilai parameter

masing-masing $-4,697 \cdot 10^{-2}$ dan $1,061 \cdot 10^{-2}$. Hal ini berarti percabangan untuk molekul yang sama akan memperkecil nilai indeks bias.

Koefisien deskriptor $n(\mathbf{C-C}_{\text{sikl}})$ sebesar $9,762 \cdot 10^{-2}$ menunjukkan arti bahwa adanya ikatan C-C siklik akan menambah besar indeks bias senyawa.

b. Ikatan rangkap

Senyawa alkena dan alkuna dengan karakteristik memiliki ikatan rangkap akan berpengaruh terhadap nilai indeks bias. Posisi ikatan rangkap tersebut juga akan berpengaruh. Untuk senyawa alkena yakni senyawa dengan ikatan rangkap dua, kenaikan estimasi indeks bias ditunjukkan dengan nilai deskriptor untuk $n(\mathbf{C}_2=\mathbf{C}_3)$, $n(\mathbf{C}_3=\mathbf{C}_4)$ dan $n(\mathbf{C}_4=\mathbf{C}_5)$ berturut-turut adalah $1,117 \cdot 10^{-2}$, $1,045 \cdot 10^{-2}$ dan $1,536 \cdot 10^{-2}$. Untuk senyawa alkuna yakni senyawa dengan ikatan rangkap tiga, kenaikan estimasi indeks bias ditunjukkan dengan nilai deskriptor untuk $n(\mathbf{C}_1\equiv\mathbf{C}_2)$, $n(\mathbf{C}_2\equiv\mathbf{C}_3)$, $n(\mathbf{C}_3\equiv\mathbf{C}_4)$, $n(\mathbf{C}_4\equiv\mathbf{C}_5)$, $n(\mathbf{C}_5\equiv\mathbf{C}_6)$ dan $n(\mathbf{C}_6\equiv\mathbf{C}_7)$ berturut-turut adalah $1,483 \cdot 10^{-2}$, $2,316 \cdot 10^{-2}$, $2,402 \cdot 10^{-2}$, $1,983 \cdot 10^{-2}$, $2,453 \cdot 10^{-2}$ dan $2,748 \cdot 10^{-2}$. Ikatan rangkap tersebut merupakan ikatan tidak jenuh dan secara fisik berakibat terhadap kerapatan antar molekul akan menjadi lebih besar dan dengan demikian cahaya datang pada bahan akan lebih dibiaskan.

c. Pengaruh atom hetero

Adanya atom hetero oksigen dan khlorida memberikan deskriptor yang bernilai positif yakni $n(\mathbf{O})$ dan $n(\mathbf{Cl})$ masing-masing sebesar $1,675 \cdot 10^{-2}$ dan $4,708 \cdot 10^{-2}$. Hal ini menunjukkan bahwa dengan adanya atom O dan Cl pada suatu senyawa akan menambah harga estimasi indeks bias suatu senyawa. Fenomena ini dapat dipahami karena munculnya atom berukuran besar akan memperbesar kerapatan molekul dan dengan demikian interaksi antar molekul akan bertambah. Faktor ini akan menyebabkan cahaya yang dibiaskan semakin besar dan nilai indeks bias bertambah besar.

Sebaliknya pada sisi lain, ikatan yang terjadi dari adanya atom hetero dengan elektronegativitas rendah yakni atom-atom oksigen, fluor dan khlor akan menyebabkan pengaruh negatif terhadap penurunan nilai kerapatan. Efek dari hal tersebut adalah terjadinya penurunan nilai indeks bias. Fenomena ini direpresentasikan pada persamaan QSPR yang diperoleh yakni untuk deskriptor-deskriptor $n(\mathbf{C-O})$, $n(\mathbf{C-F})$ dan $n(\mathbf{C-Cl})$ dengan nilai masing-masing sebesar $-2,075 \cdot 10^{-2}$, $-3,606 \cdot 10^{-2}$ dan $-3,847 \cdot 10^{-2}$. Ikatan C dengan atom yang memiliki elektronegativitas tinggi akan menaikkan estimasi indeks bias seperti yang telah diterangkan di atas. Oleh karena itu untuk ikatan atom karbon dengan atom brom akan menghasilkan efek yang berlawanan. Untuk deskriptor $n(\mathbf{C-Br})$ ini memiliki nilai positif yakni $4,057 \cdot 10^{-2}$.

Deskriptor untuk ikatan N-H dan S-H dari persamaan QSPR memiliki harga $2,619 \cdot 10^{-2}$ dan $4,389 \cdot 10^{-2}$ berturut-turut untuk $n(\mathbf{N-H})$ dan $n(\mathbf{S-H})$. Untuk ikatan yang terbentuk dari ikatan antar

atom N-H dan S-H tersebut memiliki karakteristik untuk menghasilkan fenomena ikatan hidrogen antar molekul dan inter molekul yang cukup kuat. Hal ini akan dapat mengurangi laju kecepatan cahaya dalam senyawa dan dengan demikian akan terjadi kenaikan estimasi indeks bias untuk senyawa-senyawa organik yang mengandung atom N (golongan amina) dan S (golongan sulfida)

d. Ikatan karbonil

Dari persamaan QSPR ditunjukkan bahwa adanya ikatan karbonil dalam suatu senyawa akan menyebabkan penurunan estimasi indeks bias. Koefisien deskriptor $n(\text{C=O})$ sebesar $-7,240 \cdot 10^{-3}$. Akan tetapi ikatan ini relatif tidak begitu berpengaruh karena nilai parameternya kecil sedangkan jumlah ikatan karbonil pada suatu senyawa biasanya juga kecil (satu atau dua saja). Adanya ikatan karbonil sebenarnya juga akan mengakibatkan terbentuknya ikatan hidrogen tetapi berdasarkan asumsi jumlah ikatan yang ada maka diduga efek ini menyebabkan terjadinya penurunan nilai indeks bias.

e. Posisi gugus

Untuk senyawa-senyawa aromatis tersubstitusi, posisi gugus akan berpengaruh terhadap nilai indeks bias. Dari model 13 posisi *para* relatif tidak berpengaruh secara statistik terhadap perubahan nilai indeks bias. Posisi yang relatif berpengaruh adalah posisi *ortho* dan *meta*. Koefisien deskriptor $n(\text{ortho})$ dan $n(\text{meta})$ masing-masing adalah $2,503 \cdot 10^{-2}$ dan $1,861 \cdot 10^{-2}$. Hal itu menunjukkan arti bahwa pada senyawa aromatik bila terdapat gugus pada posisi *ortho* dan *meta* maka indeks bias senyawa akan bertambah besar.

f. Pengaruh deskriptor lain

Selain deskriptor jenis atom dan ikatan kimia maka dalam penelitian ini perlu ditambahkan satu parameter penting lain yakni berat molekul $n(\text{BM})$ dan beberapa kombinasinya. Dari penelitian dibuktikan bahwa parameter ini hanya berperan dalam bentuk deskriptor turunan yakni dalam bentuk $n(\text{HxBM})$ dan $n(\text{C}_{\text{prim}}\text{xBM})$ yang ditunjukkan dengan nilai parameter sebesar $-2,271 \cdot 10^{-6}$ dan $-1,287 \cdot 10^{-4}$. Pengaruh kedua deskriptor tidak dapat dijelaskan secara fisik namun hanya memberikan arti secara statistik saja.

Jadi dari persamaan QSPR ini memang terlihat bahwa nilai indeks bias memiliki hubungan yang erat dengan struktur molekul senyawa dan telah dibuktikan hubungan tersebut secara kuantitatif. Tindak lanjut dari penelitian ini adalah pemanfaatan persamaan QSPR yang telah diperoleh untuk suatu aplikasi program komputer untuk estimasi nilai indeks bias dari suatu senyawa. Dalam hal ini sangat diperlukan kontribusi pihak lain yang tertarik untuk pembuatan dan pengembangan program tersebut.

KESIMPULAN

Model alternatif persamaan HKSS terbaik untuk indeks bias menunjukkan bahwa indeks bias berhubungan secara linear dengan deskripto molekular meliputi : $n(C_{prim\text{ alf}})$, $n(C_{sek\text{ alf}})$, $n(C_{ter\text{ alf}})$, $n(C_{kwar\text{ alf}})$, $n(C_{sek\text{ arm}})$, $n(O)$, $n(Cl)$, $n(C-C_{siki})$, $n(C_2=C_3)$, $n(C_3=C_4)$, $n(C_4=C_5)$, $n(C_1=C_2)$, $n(C_2=C_3)$, $n(C_3=C_4)$, $n(C_4=C_5)$, $n(C_5=C_6)$, $n(C_6=C_7)$, $n(C-O)$, $n(C-F)$, $n(C-Cl)$, $n(C-Br)$, $n(C=O)$, $n(N-H)$, $n(S-H)$, $n(ortho)$, $n(meta)$, $n(HxBM)$ dan $n(C_{prim}xBM)$.

DAFTAR PUSTAKA

- Carlton, T. S., 1998, Correlation of Boiling Point with Molecular Structure for Chlorofluoroethanes, *J.Chem.Inf.Comput.Sci.*, 38, 158 - 164.
- Duchowicz, P., and Castro, E.A., 2000, A Rather Simple Method to Calculate log P Values in QSAR/QSPR Studies, *Acta Chim.Slov.*, 47, 281 – 292.
- Katritzky, A.R., and Labanov, V.S., 1995, QSPR: The Correlation and Quantitative Prediction of Chemical and Physical Property from Structure, *Chem.Soc.Rev.*,28, 279 – 287.
- Kubinyi, H., 1993, *QSAR: Hansch Analysis and Related Approached*, VCH, Weinheim.
- Leach, A.R., 1996, *Molecular Modelling : Principles and Application*, Adison-Wesley, Longman
- Lide, R., 2000, *CRC Hand Book of Chemistry and Physics*, 81ed, CRC Press, Clarendon
- Reinhard M, dan Drefahl A., 1999, *Handbook for Estimating Physicochemical Properties of Organic Compounds*, John Wiley & Sons, Inc. New York