



ZVEZA ZA TEHNIČNO KULTURO SLOVENIJE

54. srečanje mladih raziskovalcev Slovenije 2020

QSPR ali kvantitativna analiza razmerja med strukturo in lastnostmi izbranih organskih molekul

Raziskovalno področje: Interdisciplinarna področja (kemija, matematika)

Raziskovalna naloga

Avtorica: Ana Breznik

Mentorici: dr. Petra Žigert Pleteršek, Zdenka Keuc

Šola: II. gimnazija Maribor

Maribor, 2020

POVZETEK

Uporaba strukturnih molekulskih deskriptorjev za napoved fizikalno-kemijskih lastnosti ter biološke aktivnosti novih kemijskih entitet postaja nepogrešljiva metoda v moderni znanosti. V nalogi sem raziskovala Ohridski indeks (OI), topološki indeks, ki temelji na stopnji vozlišča. Kot skupino testnih molekul sem izbrala strukturne izomere oktana in 82 poliaromatskih ogljikovodikov (PAH). Izračunane vrednosti Ohridskega indeksa za posamezne molekule sem primerjala z vrednostmi drugih topoloških indeksov ter s fizikalno-kemijskimi lastnostmi. Korelacije z drugimi indeksi so visoke, najvišja je bila najdena z OI poliaromatskih ogljikovodikov in Poglianijevim indeksom ($r = 0.9968$). Kar se tiče fizikalnih lastnosti, Ohridski indeks odlično korelira s standardno izparilno entalpijo izomerov oktana (0.9607), vreliščem PAH (0.9758) ter logaritemsko vrednostjo koeficienta porazdelitve (0.9622). Na tej osnovi lahko zaključimo, da je Ohridski indeks 'uporaben topološki indeks', ki bi v prihodnosti lahko našel svoje mesto pri napovedovanju fizikalno-kemijskih lastnosti ter biološke aktivnosti novih kemijskih entitet.

Ključne besede: ohridski indeks, topološki indeks, napoved fizikalno-kemijskih lastnosti ter biološke aktivnosti novih kemijskih entitet

ZAHVALA

Iskreno se zahvaljujem obema mentoricama za vso podporo, tudi v času vikendov in dopustov, za zanesljivost, nasvete, predvsem pa za prejeto znanje, ki ga bom zagotovo pridoma uporabila.

Srčna hvala prof. dr. Gajserju za ponoven pregled ter svetovanje o postavitvi vejic.

Hvala tudi mojim staršem za potrpežljivost ob neprespanih nočeh ter boljše obvladovanje Microsoftovih programov.

KAZALO VSEBINE

1. UVOD.....	7
1.1 Raziskovalno vprašanje.....	7
1.2 Metoda dela.....	8
2 TEORETIČNO OZADJE.....	8
2.1 QSAR model.....	8
2.2 Teorija grafov.....	10
2.2.1 Primeri grafov.....	12
2.3 Molekulski graf.....	13
3 PRAKTIČNI DEL.....	15
3.1 Računanje Ohridskega indeksa.....	15
3.2 Statistična analiza.....	16
3.2.1 Osnovni statistični koncepti.....	16
3.3 Računanje Ohridskega indeksa za izomere oktana.....	18
3.3.1 Primerjava vrednosti Ohridskega indeksa z drugimi topološkimi deskriptorji za strukturne izomere oktana.....	24
3.3.2 Korelacija Ohridskega indeksa z izbranimi fizikalno-kemijskimi lastnostmi strukturnih izomerov oktana.....	27
3.4. Računanje vrednosti Ohridskega indeksa za poliaromatske ogljikovodike.....	33
3.4.1. Primerjava vrednosti Ohridskega indeksa z drugimi topološkimi indeksi za poliaromatske ogljikovodike.....	33
3.4.2. Korelacija Ohridskega indeksa z izbranimi fizikalno-kemijskimi lastnostmi poliaromatskih ogljikovodikov.....	40
4 ZAKLJUČEK.....	43
5 VIRI.....	45
6 DODATEK.....	45

KAZALO GRAFOV

Graf 1: Korelacija med Ohridskim indeksom (OI) in vrelščem.....	18
Graf 2: Korelacija med OI in ZM1.....	24
Graf 3: Korelacija med OI in Qindex-om.....	24
Graf 4: Korelacija med OI in SNar.....	25
Graf 5: Korelacija med OI in HNar.....	25
Graf 6: Korelacija med OI in GNar.....	25
Graf 7: Korelacija med OI in Xt.....	26
Graf 8: Korelacija med OI in Ram.....	26
Graf 9: Korelacija med OI in ZM2.....	26
Graf 10: Dobljene vrednosti R^2 med primerjavo Ohridskega indeksa z drugimi topološkimi indeksi.....	27
Graf 11: Korelacija med OI in entropijo.....	30

Graf 12: Korelacija med OI in ΔH_{vap}	31
Graf 13: Korelacija med OI in ΔH_{vap}°	31
Graf 14: Korelacija med OI in acentričnim faktorjem	32
Graf 15: Dobljene vrednosti R^2 pri primerjavi vrednosti OI s fizikalno-kemijskimi lastnostmi izomerov oktana	33
Graf 16: Korelacija med OI in ZM1	36
Graf 17: Korelacija med OI in ZMV1	36
Graf 18: Korelacija med OI in ZM2	37
Graf 19: Korelacija med OI in ZM2V	37
Graf 20: Korelacija med OI in Qindex-om	38
Graf 21: Korelacija med OI in SNar	38
Graf 22: Korelacija med OI in Xt	38
Graf 23: Korelacija med OI in Dz	39
Graf 24: Korelacija med OI in Ram	39
Graf 25: Dobljenje vrednosti R^2 pri primerjavi vrednosti OI z vrednostmi drugih topoloških deskriptorjev za poliaromatske ogljikovodike	40
Graf 26: Korelacija med OI in LogP	42
Graf 27: Korelacija med OI in vreliščem	42
Graf 28: Korelacija med OI in tališčem	43

KAZALO TABEL

Tabela 1: Seznam lastnosti potrebnih za preiskave kemikalij (Basak, 2019).....	10
Tabela 2: Ohridski indeks in vrelišča prvih osmih cikloalkanov	17
Tabela 3: 18 strukturnih izomerov oktana s pripadajočimi strukturnimi formulami in molekulskimi grafi	19
Tabela 4: Vrednosti topoloških deskriptorjev za strukturne izomere oktana	23
Tabela 5: Vrednosti izbranih fizikalno-kemijskih lastnosti strukturnih izomerov oktana	28
Tabela 6: Korelacijski koeficienti med vrednostmi Ohridskega indeksa in fizikalno-kemijskimi lastnostmi izomerov oktana.....	32
Tabela 7: Vrednosti topoloških deskriptorjev za poliaromatske ogljikovodike	34
Tabela 8: Vrednosti izbranih fizikalno-kemijskih lastnosti poliaromatskih ogljikovodikov	40

KAZALO DIAGRAMOV

Diagram 1: Načrt praktičnega dela	15
---	----

KAZALO SLIK

Slika 1: Hierarhična razvrstitev kemodeskriptorjev in biodeskriptorjev za uporabo v QSAR (Basak, 2016).....	9
Slika 2: Primer enostavnega grafa G.....	11
Slika 3: Cikel C6	12
Slika 4: Polni grafi Kn.....	13
Slika 5: Zvezde Sn	13
Slika 6: Strukturna formula propena.....	13
Slika 7: Molekulski graf 2,2-dimetilbutana	14

OKRAJŠAVE UPORABLJENE V RAZISKOVALNI NALOGI

OI- Ohridski indeks	GNar- Narumijev geometrijski topološki indeks
QSAR- Kvantitativna analiza razmerja med strukturo in aktivnostjo	Xt- Celostni strukturni indeks povezanosti
QSPR- Kvantitativna analiza razmerja med strukturo in lastnostmi	Ram- Indeks razvejanosti
ZM1- Prvi Zagrebški indeks	BP- vrelišče
ZM1V- Prvi valenčni Zagrebški indeks	MP- tališče
ZM2- Drugi Zagrebški indeks	TSA- skupna ploščina
ZM2V- Drugi valenčni Zagrebški indeks	S- Entropija
Qindex- Kvadratični indeks	MON- oktansko število
SNar- Narumijev enostavni topološki indeks	LogP- Koeficient particije med vodo in oktanolom
HNar- Narumijev harmonični topološki indeks	

1. UVOD

Uporaba strukturnih molekulskih deskriptorjev za napoved fizikalno-kemijskih lastnosti ter biološke aktivnosti novih kemijskih entitet postaja vse pogostejša smer razvoja na področju farmacevtske industrije in okoljske toksikologije (Hawkins et al., 2006). Kvantitativna analiza razmerja med strukturo in aktivnostjo (QSAR) je ena izmed najpogostejših metod za modeliranje lastnosti molekul (Cherkasov et al., 2014). Parthasarathi in Dhawan (2018) sta QSAR opisala kot: *»tehniko, ki poskuša predvideti aktivnost, reaktivnost in lastnosti neznane skupine molekul glede na analizo enačbe, ki strukturo molekul poveže z izmerjeno aktivnostjo ter lastnostmi.«* Glavni cilj QSAR analize je torej razvoj kar se da preprostega modela, ki omogoča določitev ter napoved potencialnega vpliva kemikalij, materialov na zdravje ljudi in različne ekosisteme. Poleg tega so tudi dobra referenca za morebitno dodatno modifikacijo molekul, ki se bo odražala v izboljšani aktivnosti in zelenih lastnosti (Vilar in Contanzi, 2012).

Pred sprejetjem nove kemijske entitete v splošno uporabo je potrebnih za določitev lastnosti in aktivnosti ogromno predhodnih eksperimentov. To zahteva veliko energije, materialov in delovne sile. Matematična napoved fizikalno-kemijskih lastnosti skrajša čas in pomeni prihranke v energiji in materialih (Gao et al., 2016). QSAR lahko torej razumemo kot znanstveni varčevalni načrt, kot pomoč pri iskanju inovativnih trajnostnih rešitev in priložnost za razvoj bolj zelene kemije.

QSAR se je skozi leta razvila v zelo kompleksno metodo, kjer so možni različni pristopi. Moja mentorica me je seznanila s t.i. Ohridskim indeksom (OI), ki ga je leta 1989 razvil Ivan Gutman. Ker korelacije tega indeksa s fizikalnimi in kemijskimi lastnostmi še nikdar niso bile raziskane, sem bila zelo navdušena in počaščena, da je ta izziv pripadel meni. Ohridski indeks spada med 2D topološke deskriptorje, torej opisuje 2D prikaze molekul. Direktna in običajno preprosta matematična narava sta gotovo največji prednosti 2D deskriptorjev (Cherkasov et al., 2014). V svoji nalogi sem raziskovala korelacije vrednosti Ohridskega indeksa s fizikalno-kemijskimi lastnostmi določene skupine molekul, torej moj pristop sodi pod QSPR, oziroma, kvantitativno analizo razmerja med strukturo in lastnostmi.

1.1. RAZISKOVALNO VPRAŠANJE

Kako uspešno lahko napovemo fizikalno-kemijske lastnosti izomerov oktana ter poliaromatskih ogljikovodikov s pomočjo Ohridskega indeksa?

Uspešnost je v tej nalogi definirana s pomočjo korelacijskih indeksov in Randićevih atributov.

1.2 METODA DELA

Za določitev uporabnosti topološkega indeksa lahko uporabimo seznam atributov dobrega topološkega indeksa, ki jih navaja Randić (1991). Te so: direktna interpretacija strukture, dobra korelacija z vsaj eno fizikalno ali kemijsko lastnostjo, zmožnost dobrega ločevanja med izomeri, ki temelji na strukturalni formuli molekule in ne na fizikalno-kemijskih lastnostih, postopne strukturne spremembe pa se odražajo v postopnih spremembah vrednosti fizikalno-kemijskih lastnostih.

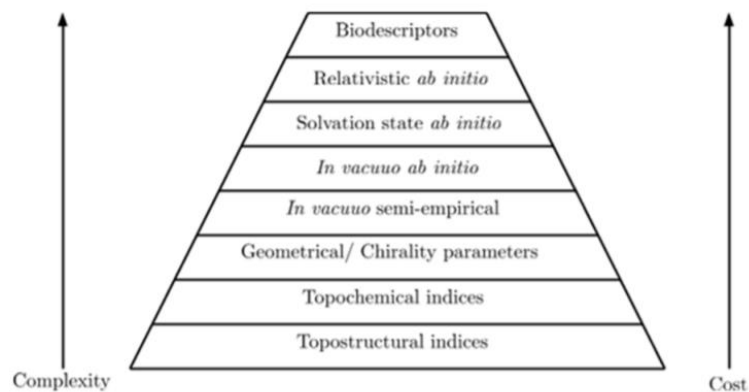
Naslednji korak je bila izbira skupine molekul za testiranje Ohridskega indeksa. Za začetek sem se odločila za strukturne izomere oktana, saj so relativno preprosta skupina s primernim številom molekul, ki bo omogočila vpogled v uspešnost Ohridskega indeksa v razlikovanju med izomeri. Nadaljevanje so predstavljali policiklični aromatski ogljikovodiki (PAH).

2 TEORETIČNO OZADJE

Sylevster je leta 1878 prvi predlagal, da je strukturna formula molekule posebna vrsta grafa, kar je bila inovativna teorija brez napovednega potenciala. Kasneje je uporaba kombinatorike predvidela obstoj natanko dveh izomerov butana (C_4H_{10}), ker obstajata natanko dva drevesna grafa s štirimi vozlišči, če upoštevamo le ogljike. To je bil teoretični model z omejenim napovednim potencialom, saj je predvidel le obstoj kemijske vrste, ni pa bil zmožen predvideti lastnosti. Basak (2016) je vsaka empirična lastnost (P) preslikal v skupino kemijskih struktur in te v množico realnih števil ter jih tako empirično uredil. Zaključil je, da je za napoved lastnosti na podlagi strukture potreben ne-empirični, strukturni razvrščevalni ključ, ki bo podoben empiričnemu razvrščanju te množice po lastnosti P. To je tudi osnova QSAR modela.

2.1 QSAR MODEL

Hierarhija deskriptorjev je zgrajena iz kombinacije topoloških, geometričnih, kvantnih in kemijskih deskriptorjev ter biodeskriptorjev. Začne se z najpreprostejšimi grafovskimi invariantami ter zaključuje z deskriptorji, ki zahtevajo ogromne količine laboratorijskih podatkov.



SLIKA 1: Hierarhična razvrstitev kemodeskriptorjev in biodeskriptorjev za uporabo v QSAR (Basak, 2016).

V tabeli 1 je naveden delni seznam lastnosti, ki morajo biti določene v procesu odkritja zdravil za zaščito zdravja ljudi in okolja. Določitev vseh lastnosti za posamezno snov je izjemno drago, ne glede na to ali so določene *in vivo* ali *in vitro*. Ena izmed alternativ je uporaba približkov lastnosti s pomočjo QSAR ali druge metode znotraj hierarhije. Štirje glavni stebri razvoja uporabnega QSAR sistema so (Basak, 2013):

- a) Razpoložljivost kvalitetnih eksperimentalnih podatkov.
- b) Podatki za zadostno količino spojin (relativno dobra velikost vzorca).
- c) Razpoložljivost ustreznih deskriptorjev, ki količinsko opredelijo molekulsko strukturo.
- d) Uporaba primernih metod za razvoj in validacijo modela.

Pred začetkom raziskave sem se morala seznaniti z osnovami teorije grafov, področjem matematike, na katerem temelji tudi QSPR.

TABELA 1: Seznam lastnosti potrebnih za preiskave kemikalij (Basak, 2019)

Physicochemical	Pharmacological / Toxicological
Molar volume	Macromolecular level
Boiling point	: Receptor binding (K_d)
Melting point	: Michaelis constant (K_m)
Vapor pressure	: Inhibitor constant (K_i)
Water solubility	: DNA alkylation
Dissociation constant (pK_a)	: Unscheduled DNA synthesis
Partition coefficient	Cell level
: Octanol-water ($\log P$)	: Salmonella mutagenicity
: Air-water	: Mammalian cell transformation
: Sediment-water	Organism level (acute)
Reactivity (electrophilicity)	LD ₅₀ (mouse, rat)
	LC ₅₀ (fathead minnow)
	Organism level (chronic)
	: Bioconcentration factor
	: Carcinogenicity
	: Reproductive toxicity
	: Delayed neurotoxicity
	: Biodegradation profile

2.2 TEORIJA GRAFOV

Vse definicije so povzete po *Introduction to Graph Theory* (Wilson, 2015).

Graf $G = (V(G), E(G))$ je definiran kot neprazna končna množica $V(G)$ elementov, imenovanih vozlišča ter končne množice $E(G)$, ki je sestavljena iz neurejenih parov elementov $V(G)$, imenovanih povezave grafa. Vozlišča običajno označimo s črkami, na primer u, v, w, z, \dots , medtem ko povezave najpogosteje ponazorimo z vozlišči. Na primer, $e = uv$, kjer sta $u, v \in V(G)$. V takem primeru sta vozlišči u in v sosednji ter incidentni povezavi e . Graf brez zank in večkratnih povezav imenujemo enostavni graf.

Stopnja (ali valenca) vozlišča v grafa G je število povezav incidentnih vozlišču v . Stopnjo zapišemo kot $\deg v$. Ta pojem je še posebej pomemben za kemijo, saj je pojem valenca analogen številu kemijskih vezi, ki določen atom povezujejo s sosednjimi. Vozlišče stopnje 0 se imenuje izolirano vozlišče, vozlišče stopnje 1 pa terminalno vozlišče. Če so vsa vozlišča grafa G iste stopnje r , potem je graf regularen.

Hitro je postalo razvidno, da je vsota stopenj vseh vozlišč v grafu sodo število, natančneje dvakratnik števila povezav. To odkritje izhaja iz Eulerjeve rešitve slavnega problema königsberških mostov, ki velja za 'prvi članek teorije grafov' (1736). Eulerjevo odkritje danes poznamo pod imenom *Lema o rokovanju* (Wilson in Watkins, 1990).

Če je lema G graf, potem velja:

$$\sum_{v \in V(G)} \deg v = 2|E(G)|$$

Dokaz je očiten. Povezava poteka skozi dve vozlišči, torej vsaka h končni vsoti prispeva natanko 2.

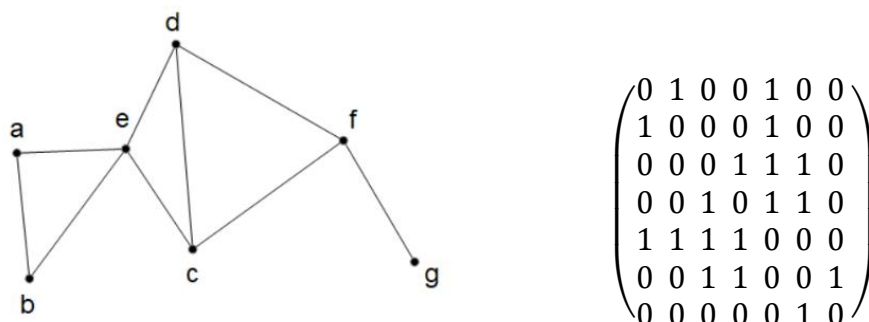
Posledice leme o rokovanju so:

- V vsakem grafu je vsota stopenj vseh vozlišč sodo število.
- V vsakem grafu je število točk lihe stopnje sodo.
- Če ima graf G n vozlišč in je regularen stopnje r , ima G natanko $\frac{1}{2} \times nr$ povezav.

Matrike so še en način predstavitve grafov. Matrika je razpredelnica števil, v povezavi s teorijo grafov največkrat govorimo o incidenčni matriki in matriki sosednosti. Sama se bom osredotočila le na slednjo. Matrika sosednosti $M(G)$ je matrika razsežnosti $n \times n$, v kateri vsak element v j -tem stolpcu i -te vrstice pove številu povezav, ki povezujejo vozlišči i in j (Wilson in Watkins, 1990).

Spodaj je podan enostavni graf G s pripadajočo matriko sosednosti. Ker sta vozlišči a in b povezani z eno povezavo, zapišemo število 1 v drugi stolpec prve vrstice ter v prvi stolpec druge vrstice. Če med dvema vozliščema ni povezave, v matriko sosednosti zapišemo 0, kar lahko v spodnjem primeru vidimo med vozliščema a in c .

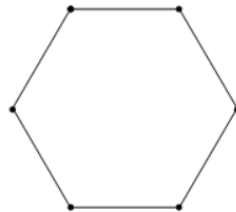
Če pobliže pogledamo, vidimo, da je vsak element glavne diagonale matrike sosednosti (iz zgornjega levega kota proti spodnjemu desnemu) enak 0, saj je G enostavni graf, torej ne vsebuje nobene zanke. V sosednostni matriki enostavnega grafa sta zmeraj prisotni le števili 0 in 1, saj glede na definicijo med dvema vozliščema zmeraj poteka največ ena povezava.



SLIKA 2: Primer enostavnega grafa G

Iskanje najkrajše poti med dvema točkama v določenem sistemu je zelo uporabna aplikacija teorije grafov. Tako moramo definirati sprehod v grafu. Sprehod dolžine n v grafu G je

zaporedje povezav grafa G oblike $u_1u_2, u_2u_3, u_3u_4, \dots, u_{n-1}u_n$. Tak sprehod poimenujemo sprehod med vozliščema u_1 in u_n . Če so vse povezave sprehoda različne, potem sprehod poimenujemo enostavni sprehod. Če so v enostavnem sprehodu vsa vozlišča različna, potem sprehod imenujemo pot. Sprehod je sklenjen, če $u_1 = u_n$. Če so v sklenjenem sprehodu vse povezave in vse točke različne, potem ga poimenujemo cikel in označimo s C_n . Graf C_n je regularen stopnje 2 in ima n povezav. Spodaj je podan cikel C_6 .

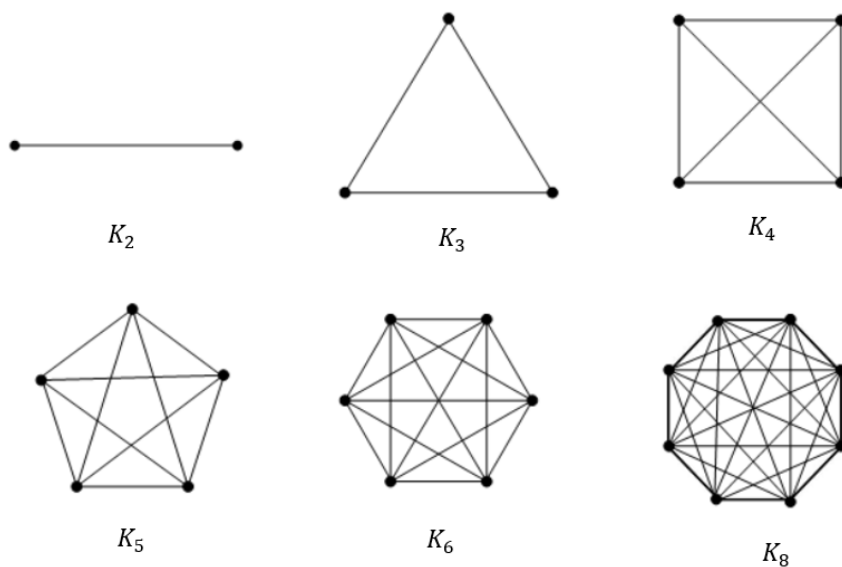


SLIKA 3: Cikel C_6

To nas vodi do še ene definicije. Graf G je povezan, če obstaja pot med poljubnim parom vozlišč, sicer je nepovezan (Wilson in Watkins, 1990).

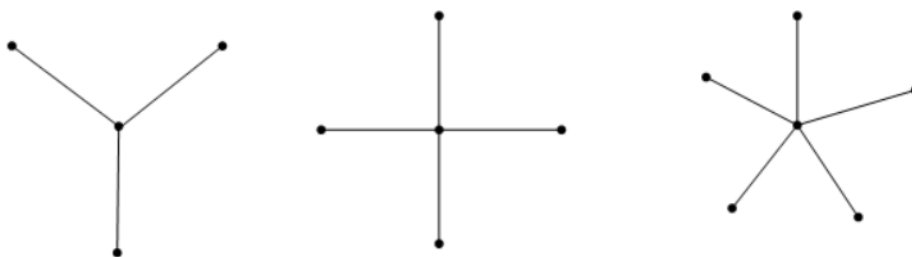
2.2.1 PRIMERI GRAFOV

Polni graf je enostavni graf, v katerem je vsak par različnih vozlišč povezan z natanko eno povezavo. Polni graf z n točkami označimo s K_n . Takšen graf je regularen. Slika 4 prikazuje primere polnih grafov.



SLIKA 4: Polni grafi K_n

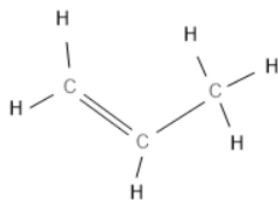
Zvezda je vrsta grafa, v katerem ima natanko $n - 1$ vozlišč stopnjo 1, eno samo vozlišče pa stopnjo $n - 1$. Označimo ga s S_n . Slika 5 prikazuje zvezde S_4 , S_5 in S_6 .



SLIKA 5: Zvezde S_n

2.3 MOLEKULSKI GRAF

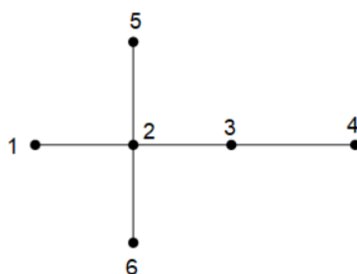
Strukturna formula je le eden izmed načinov predstavitve molekul (Rouvray, 1973). Strukturna formula propena C_3H_6 je podana na sliki 6.



SLIKA 6: Strukturna formula propena

Strukturna formula nam posreduje več informacij kot molekulska. Medtem ko nam molekulska formula pove le koliko katerih atomov je prisotnih v določeni molekuli, nam strukturna formula prikaže tudi vezi med temi atomi ter razporeditev v prostoru. Z drugimi besedami, strukturna formula nam poda informacije o topologiji molekule (Rouvray, 1973).

V procesu modeliranja zmeraj upoštevamo le določene aspekte molekulske strukture, preostale pa ignoriramo (Basak, 2016), torej vsak koncept molekulske strukture je le hipotetična skica organizacije molekul (Einstein, 1954). Teorija grafov nam tako razkriva še eno hipotetično skico, še en način predstavitve molekule, in sicer, kot graf. Če atome v molekuli predstavimo kot vozlišča, kemijske vezi pa kot povezave, dobimo kemijski ali molekulski graf (Burch, 2019). Dobljene grafe pogosto poenostavimo tako da odstranimo vodikove atome, saj bi v nasprotnem primeru geometrična struktura pri kompleksnejših molekulah postajala vse manj očitna (Wilson and Watkins, 1990). Graf z odstranjenimi vodikovimi atomi v angleščini imenujemo 'hydrogen-depleted molecular graph' (Burch, 2019). Primer takšnega grafa je prikazan na sliki 7.



SLIKA 7: Molekulski graf 2,2-dimetilbutana

Za voljo lažje primerjave različnih molekul, kodiramo kemijsko strukturo molekule v specifično število, kar nam omogočajo topološki indeksi. Gao idr. (2016) so topološki indeks opisali kot »*ne-empirično numerično vrednost, ki molekulsko strukturo preslika v število*«. Omogočajo nam numerični prikaz celotne strukture katerekoli kemijske vrste v relativno preprosto matematično enačbo (Rouvray, 1973), ter možnost iskanja morebitnih korelacij s fizikalno-kemijskimi lastnostmi (Basak idr., 1987).

Obstajajo tri glavne vrste topoloških indeksov: takšni, ki temeljijo na stopnji, razdalji in topološki indeksi, ki temeljijo na lastnih vrednostih grafa.

V svoji nalogi sem raziskovala Ohridski indeks (OI), topološki indeks temelječ na stopnji vozlišča, ki ga je razvil Ivan Gutman.

3 PRAKTIČNI DEL

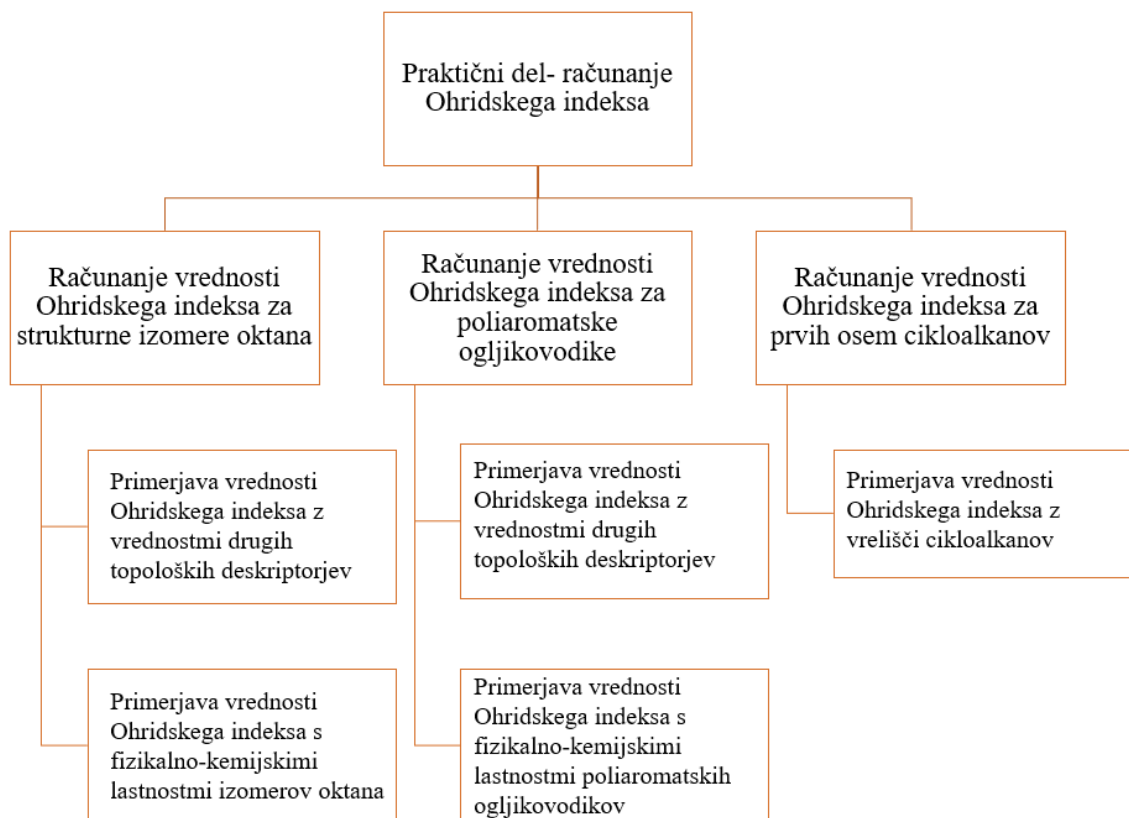


DIAGRAM 1: Načrt praktičnega dela

3.1 RAČUNANJE OHRIDSKEGA INDEKSA

Ohridski indeks se izračuna po formuli, kjer stopnja vozlišča predstavlja število C-C vezi, ki jih tvori posamezni ogljikov atom.

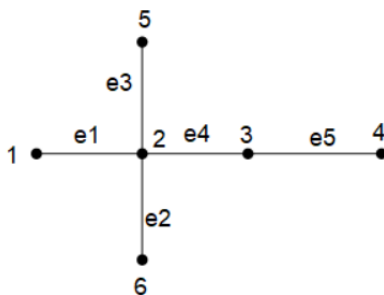
Za vsako povezavo (vez) izračunamo:

$$OI(u, v) = \frac{deg(u)}{deg(v)} + \frac{deg(v)}{deg(u)}$$

Ohridski indeks grafa G je vsota Ohridskih indeksov posameznih povezav:

$$OI(G) = \sum_{uv \in E(G)} \left(\frac{deg(u)}{deg(v)} + \frac{deg(v)}{deg(u)} \right)$$

Primer izračuna OI:



Ohridski indeks 2,2-dimetilbutana:

$$e_1 = \frac{1}{4} + 4 = \frac{17}{4} \quad e_2 = \frac{1}{4} + 4 = \frac{17}{4} \quad e_5 = \frac{1}{2} + 2 = \frac{5}{2}$$

$$e_3 = \frac{1}{4} + 4 = \frac{17}{4} \quad e_4 = \frac{1}{2} + 2 = \frac{5}{2} \quad \mathbf{OI} = 3 \times \frac{17}{4} + 2 \times \frac{5}{2} = \frac{71}{4}$$

Ta enostaven račun lahko postane pri kompleksnejših molekulah časovno zamuden (na primer poliaromatskih ogljikovodikih), zato je bila nujna uporaba ustreznega računalniškega programa. Za namene te raziskave smo tako uporabili računalniški program Matevža Črepnjaka (2019), ki je omogočil avtomatski izračun Ohridskega indeksa za posamezno molekulo na podlagi matrike sosednosti. Koda je podana v dodatku (z dovoljenjem Matevža Črepnjaka).

3.2 STATISTIČNA ANALIZA

3.2.1 OSNOVNI STATISTIČNI KONCEPTI

Pred nadaljnjo raziskavo sem se morala seznaniti še s temelji bivariatne analize. Bivariatna analiza proučuje razmerje med odvisno in neodvisno spremenljivko. Odvisna spremenljivka meri izid raziskave, medtem ko neodvisna spremenljivka pojasni spremembe v odvisni spremenljivki (Wazir, 2012). Neodvisna spremenljivka moje raziskave je tako vrednost Ohridskega indeksa, fizikalno-kemijske lastnosti pa odvisne spremenljivke.

Grafično orodje za predstavitev razmerja med dvema spremenljivkama v bivariatni statistiki je običajno razsevni grafikon. Neodvisno spremenljivko predstavimo na x-osi, odvisno pa na y-osi. Za določitev jakosti povezave med dvema spremenljivkama pogosto uporabimo korelacijski koeficient, ki meri tako smer kot moč linearnega razmerja. Zavedati se je treba, da korelacija ne pomeni vzročne zveze, torej najdena korelacija ne dokaže, da x povzroča y (Wazir, 2012). Korelacijski koeficient je zmeraj predstavljen s številom med -1 in 1. Vrednosti, ki ležijo blizu 0 nakazujejo šibko, vrednosti blizu -1 in 1 pa močno asociacijo.

Če najdemo korelacijo med dvema množicama podatkov, lahko linearno razmerje opišemo s t.i. 'best-fit line', torej premico, ki predstavlja najboljši približek trendom v podatkih. Strokovno jo imenujemo regresijska premica oblike: $y = ax + b$. S pomočjo regresijske enačbe lahko predvidimo vrednost odvisne spremenljivke glede na vrednost neodvisne. Če je vrednost x manjša ali večja od opazovanih vrednosti, potem oceno imenujemo ekstrapolacija, če pa leži nekje v razponu opazovanih vrednosti, potem jo imenujemo interpolacija (Wazir, 2012).

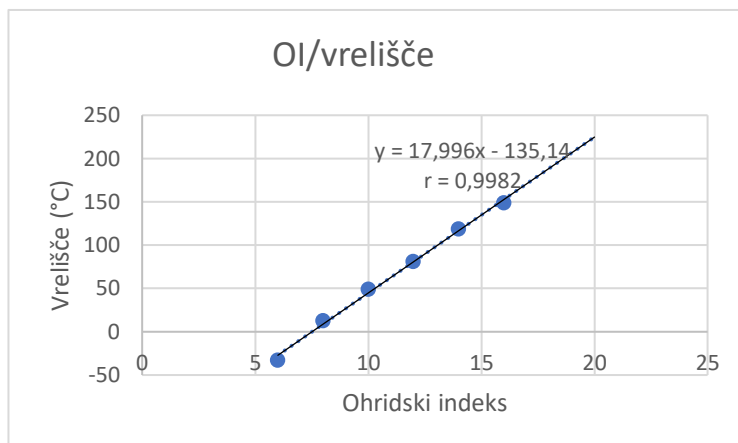
V svoji raziskavi sem kot merilo korelacije uporabila korelacijski koeficient, r .

Iz radovednosti sem poskusila izračunati Ohridski indeks za nekaj regularnih grafov in kmalu opazila trende v rezultatih. Ohridski indeks grafa regularne stopnje $r = \text{število povezav} \times 2$.

Prav tako sem izračunala Ohridski indeks grafov ciklov C_3, C_4, C_5, C_6, C_7 in C_8 . Če si te grafe zamislimo kot molekule, vidimo, da predstavljajo ciklopropan, ciklobutan, ciklopentan, cikloheksan in cikloheptan. Zato sem izračunane vrednosti Ohridskega indeksa primerjala z vrelišči prvih osmih cikloalkanov, dodala graf ter določila korelacijo s korelacijskim koeficientom. Vrednosti vrelišč so vzete iz *Handbook of Physical-chemical Properties and Environmental Fate for Organic Chemicals, Mackay (Donald, 2006)*. Rezultati so prikazani v tabeli 2.

TABELA 2: Ohridski indeks in vrelišča prvih osmih cikloalkanov

Cikloalkan	Ohridski indeks	Vrelišče (°C)
ciklopropan	6	-33.0
ciklobutan	8	12.5
ciklopentan	10	49.0
cikloheksan	12	81.0
cikloheptan	14	118.4
ciklooktan	16	149.0



GRAF 1: Korelacija med Ohridskim indeksom in vreliščem

Kot je razvidno iz grafa 1, ima korelacijski koeficient vrednost 0.9982, kar nakazuje zelo močno korelacijo.

3.3 RAČUNANJE OHRIDSKEGA INDEKSA ZA IZOMERE OKTANA

Prva skupina molekul, na kateri sem testirala Ohridski indeks so strukturni izomeri oktana. Spreminjanje vrednosti fizikalno-kemijskih lastnosti so v primeru teh molekul zelo postopne (glej tabelo 3), zato bodo dober pokazatelj, če je Ohridski indeks zmožen razlikovati med izomeri.

Najprej sem izračunala Ohridski indeks za vsakega izmed osemnajstih izomerov, nato pa sem vrednosti primerjala z:

- a) vrednostmi že uveljavljenih topoloških indeksov
- b) vrednosti izbranih fizikalno-kemijskih lastnosti

Spodaj so podani opisi topoloških indeksov, s katerimi sem primerjala Ohridski indeks.

ZM1: prvi Zagrebški indeks se izračuna kot vsota kvadratnih vrednosti stopenj vseh atomov v molekuli, z izjemo vodika (Gutman idr., 1975).

ZM2: Za izračun drugega Zagrebškega indeksa najprej za vsako vez izračunamo produkt stopenj vozlišč atomov, ki jih ta vez povezuje, kasneje pa seštejmo dobljene vrednosti za vsako vez (Gutman idr., 1975).

Qindex: Kvadratični indeks je normalizacija Prvega zagrebškega indeksa (Balaban, 1979).

SNar: Narumijev enostavni topološki indeks je topološki indeks, ki upošteva razvejanost molekul. Dobimo ga z množenjem stopenj vseh atomov (vozlišč), z izjemo vodika (Narumi, 1987).

HNar: Narumijev harmonični topološki indeks je logaritmična transformacija SNar indeksa. Število vseh nevodikovih atomov delimo z vsoto njihovih stopenj (Narumi, 1987).

GNar: Narumijev geometrijski topološki indeks je še ena logaritmična transformacija SNar indeksa. Dobimo ga z izračunom geometrične sredine stopenj vozlišč (Narumi, 1987).

Xt: Celostni strukturni indeks povezanosti je korenjena vrednost SNar indeksa (Needham, 1988).

Ram: Indeks razvejanosti dobimo s seštevkom stopenj vseh vozlišč večjih od dva, nato vsoti odštejemo 2.

V tabeli 2 so zbrani vsi strukturni izomeri oktana, s pripadajočimi strukturnimi formulami in molekulskimi grafi. Tabela 3 predstavlja neobdelane podatke o vrednostih Ohridskega indeksa ter ostalih topoloških indeksov strukturnih izomerov oktana. Vrednosti drugih indeksov so vzete iz baze podatkov: *The International Academy of Mathematical Chemistry*.

TABELA 3: 18 strukturnih izomerov oktana s pripadajočimi strukturnimi formulami in molekulskimi grafi

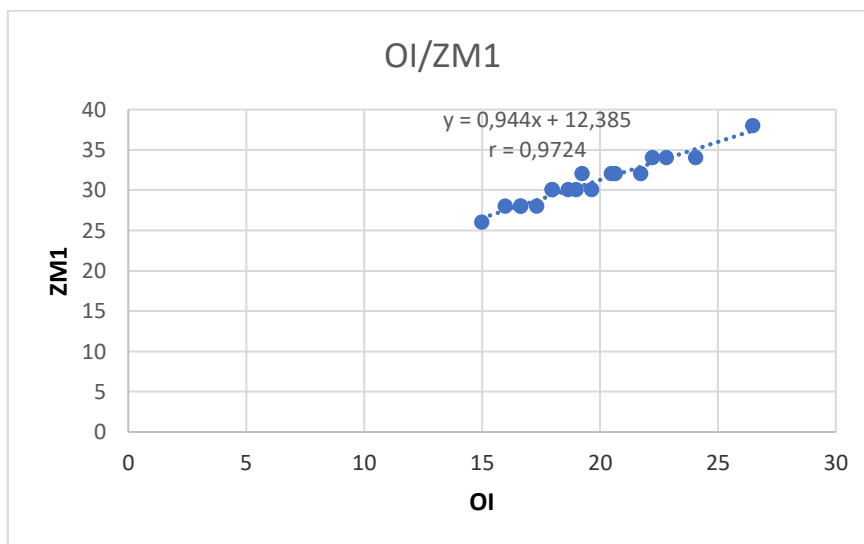
Izomer	Strukturna formula	Molekulski graf
Oktan	$ \begin{array}{cccccccc} & H & H & H & H & H & H & H \\ & & & & & & & \\ H & -C & -C & -C & -C & -C & -C & -C & -H \\ & & & & & & & \\ & H & H & H & H & H & H & H \end{array} $	
2-metilheptan	$ \begin{array}{cccccccc} & H & H & H & H & H & H & H \\ & & & & & & & \\ H & -C & -C & -C & -C & -C & -C & -C & -H \\ & & & & & & & \\ & H & & H & H & H & H & H \\ & & H-C-H & & & & & \\ & & & & & & & \\ & & H & & & & & \end{array} $	
3-metilheptan	$ \begin{array}{cccccccc} & H & H & H & H & H & H & H \\ & & & & & & & \\ H & -C & -C & -C & -C & -C & -C & -C & -H \\ & & & & & & & \\ & H & H & & H & H & H & H \\ & & & H-C-H & & & & \\ & & & & & & & \\ & & & H & & & & \end{array} $	

TABELA 4: Vrednosti topoloških deskriptorjev za strukturne izomere oktana

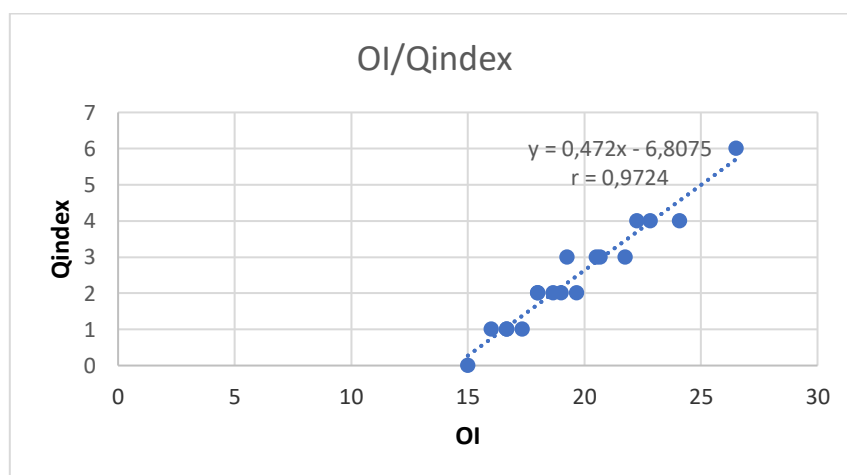
Molekula	Ohridski indeks	ZM1	ZM2	Qindex	SNar	HNar	GNar	Xt	Ram
oktan	15.00	26	24	0	4.159	1.600	1.682	0.490	0
2-metilheptan	17.33	28	26	1	3.871	1.500	1.622	0.508	1
3-metilheptan	16.67	28	27	1	3.871	1.500	1.622	0.508	1
4-metilheptan	16.67	28	27	1	3.871	1.500	1.622	0.508	1
3-etilheksan	16.00	28	28	1	3.871	1.500	1.622	0.508	1
2,2-dimetilheksan	21.75	32	30	3	3.466	1.391	1.542	0.537	2
2,3-dimetilheksan	18.67	30	30	2	3.584	1.412	1.565	0.528	2
2,4-dimetilheksan	19.00	30	29	2	3.584	1.412	1.565	0.528	2
2,5-dimetilheksan	19.67	30	28	2	3.584	1.412	1.565	0.528	2
3,3-dimetilheksan	20.50	32	32	3	3.466	1.391	1.542	0.537	2
3,4-dimetilheksan	18.00	30	31	2	3.584	1.412	1.565	0.528	2
2-metil-3-etilpentan	18.00	30	31	2	3.584	1.412	1.565	0.528	2
3-metil-3-etilpentan	19.25	32	34	3	3.466	1.391	1.542	0.537	2
2,2,3-trimetilpentan	22.83	34	35	4	3.178	1.315	1.488	0.561	3
2,2,4-trimetilpentan	24.08	34	32	4	3.178	1.315	1.488	0.561	3
2,3,3-trimetil-pentan	22.25	34	36	4	3.178	1.315	1.488	0.561	3
2,3,4-trimetilpentan	20.67	32	33	3	3.296	1.333	1.510	0.551	3
2,2,3,3-tetrametilbutan	26.50	38	40	6	2.773	1.231	1.414	0.601	4

3.3.1 PRIMERJAVA VREDNOSTI OHRIDSKEGA INDEKSA Z DRUGIMI TOPOLOŠKIMI DESKRIPTORJI ZA STRUKTURNE IZOMERE OKTANA

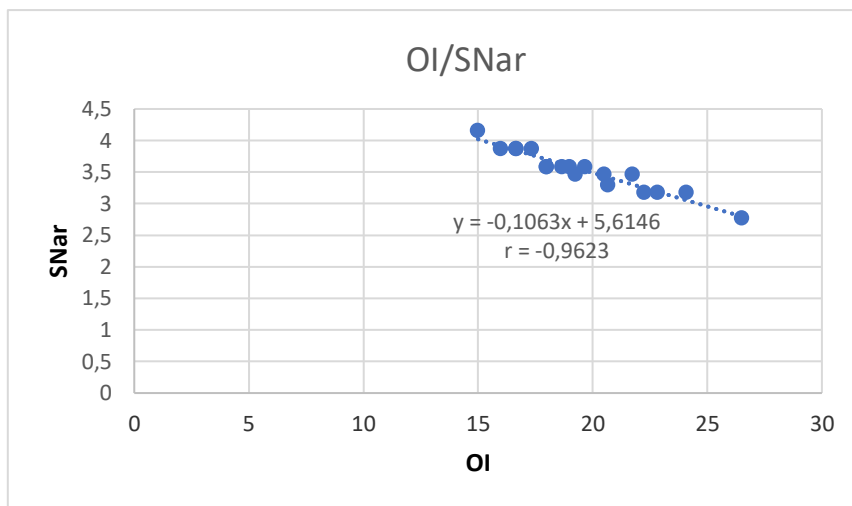
Po izračunu Ohridskega indeksa za vsakega izmed 18 izomerov, sem narisala razsevni diagram, ki primerja vrednost Ohridskega indeksa z vsakim izmed ostalih indeksov posebej. Dodala sem regresijsko premico ter korelacijski koeficient. Dobljeni grafi so prikazani spodaj.



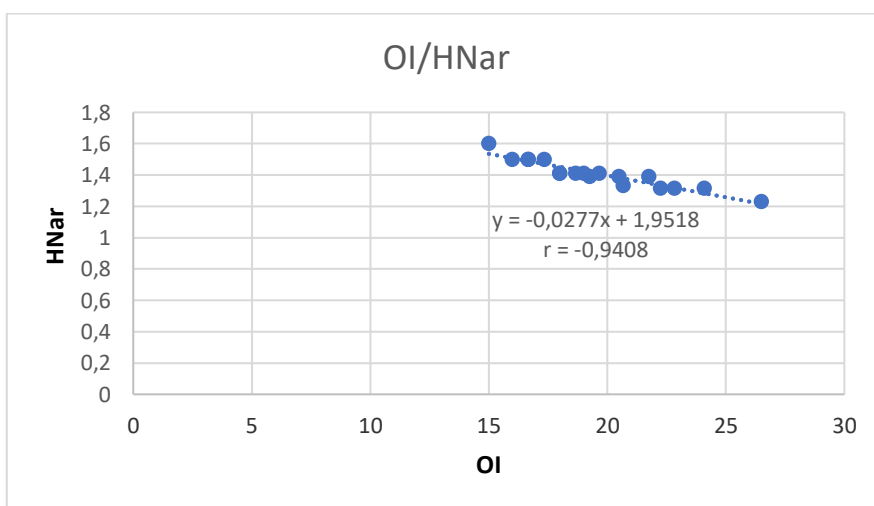
GRAF 2: Korelacija med OI in ZM1



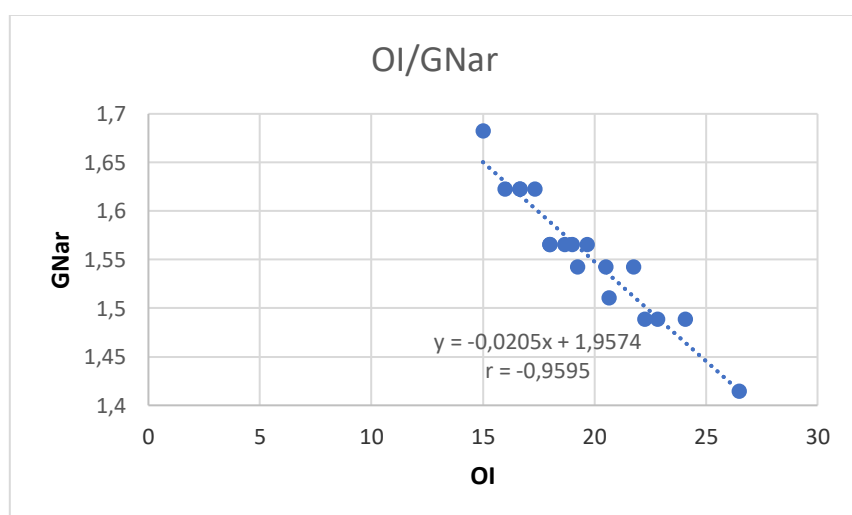
GRAF 3: Korelacija med OI in Qindex-om



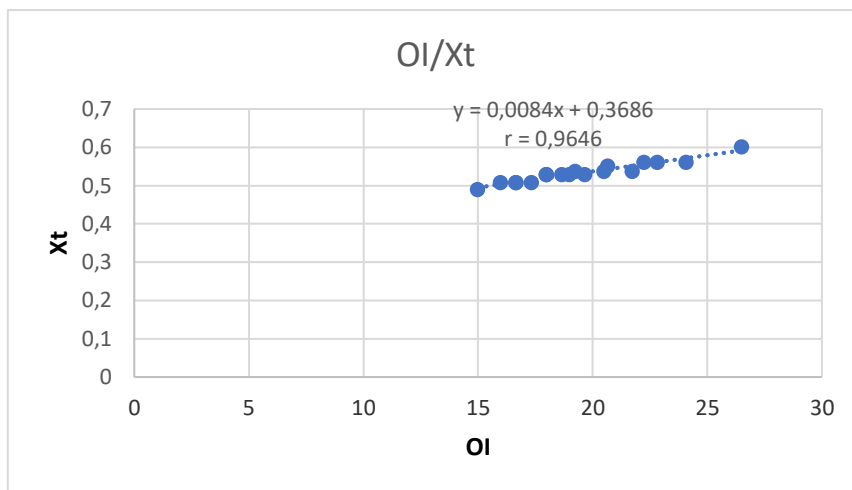
GRAF 4: Korelacija med OI in SNar



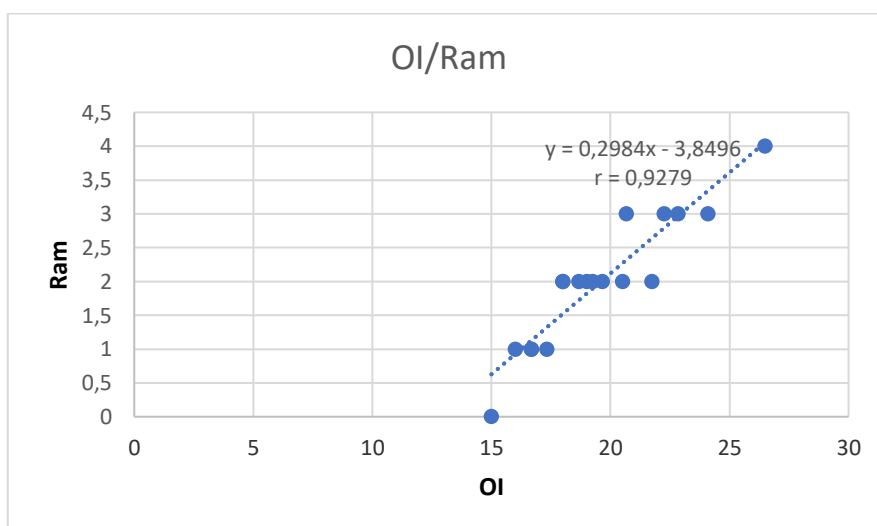
GRAF 5: Korelacija med OI in HNar



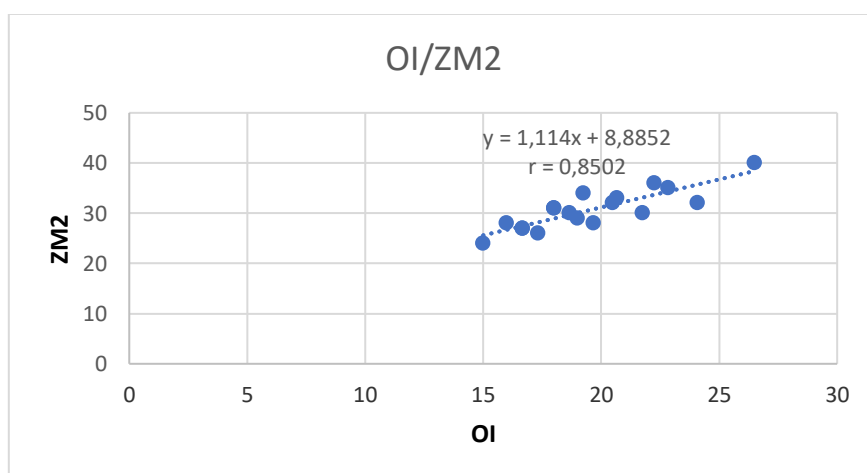
GRAF 6: Korelacija med OI in GNar



GRAF 7: Korelacija med OI in Xt



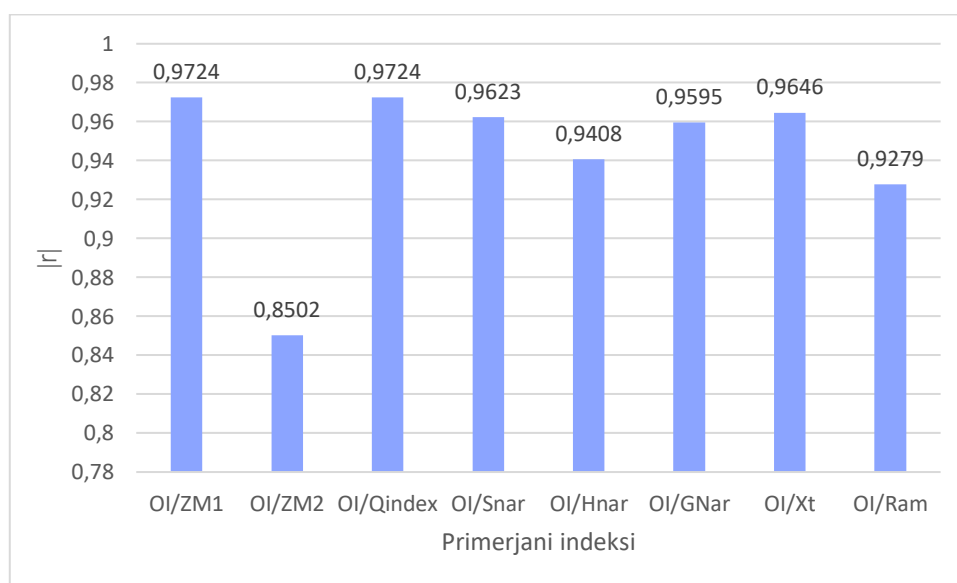
GRAF 8: Korelacija med OI in Ram



Graf 9: Korelacija med OI in ZM2

Iz zgoraj prikazanih podatkov vidimo, da Ohridski indeks korelira z vsemi primerjanimi indeksi, najboljše korelacije so bile najdene z ZM1 (0.9724), Qindex-om (0.9724), SNar (-

0.9623), GNar (-0.9595) in Xt indeksom (0.9646). Indeksa HNar in Ram sta pokazala nekoliko šibkejšo, a še zmeraj zelo močno korelacijo. Edina zmerno močna korelacija je bila najdena z ZM2 indeksom (0.8502). Odlične korelacije z že uveljavljenimi indeksi so tako dobra referenca, da bo Ohridski indeks morda koreliral tudi s fizikalno-kemijskimi lastnostmi, kar je bila naslednja stopnja moje raziskave. V tabeli 4 so zbrani podatki za nekatere fizikalno-kemijske lastnosti izomerov oktana. Vrednosti so vzete iz baze podatkov *The International Academy of Mathematical Chemistry*. Manjkajoči podatki so označeni z /.



Graf 10: Absolutne vrednosti korelacijskega koeficienta, dobljene pri primerjavi vrednosti Ohridskega indeksa z vrednostmi drugih topoloških indeksov za strukturne izomere oktana

3.3.2 KORELACIJA OHRIDSKEGA INDEKSA Z IZBRANIMI FIZIKALNO-KEMIJSKIMI LASTNOSTMI STRUKTURNIH IZOMEROV OKTANA

TABELA 5: Vrednosti izbranih fizikalno-kemijskih lastnosti strukturnih izomerov oktana

Molekula	OI	B _P (°C)	M _P (K)	C _T ($\frac{J}{K mol}$)	C _p ($\frac{J}{K mol}$)	S ($\frac{J}{K}$)	ρ ($\frac{g}{cm^3}$)	ΔH _{vap} ($\frac{kJ}{mol}$)	ΔH _{vap} ^o ($\frac{kJ}{mol}$)	ΔH _{tv} ($\frac{kJ}{mol}$)	ΔH _{tv} ^o ($\frac{kJ}{mol}$)	MON	A ($\frac{m^3}{mol}$)	ω	TSA (Å ²)	LogP	V _m (cm ³)
3,4-dimetilheksan	18.00	117.725	/	298.00	27.40	106.59	0.7200	70.2	9.316	-50.91	4.97	81.7	38.8453	0.340345	376.91	3.54	158.653
3,3-dimetilheksan	20.50	111.969	147.00	290.84	27.20	104.74	0.7100	68.5	8.973	-52.61	3.17	83.4	39.0087	0.322596	389.79	3.65	160.887
2,5-dimetilheksan	19.67	109.103	182.00	279.00	25.00	105.72	0.6935	68.6	9.051	-53.21	2.5	55.7	39.2596	0.356830	395.08	3.54	164.715
2,4-dimetilheksan	19.00	109.429	/	282.00	25.80	106.98	0.7004	68.5	9.029	-52.44	2.8	69.9	39.1300	0.344223	388.11	3.54	163.093
2,3-dimetilheksan	18.67	115.607	/	293.00	26.60	108.02	0.7121	70.2	9.272	-51.13	4.23	78.9	38.9808	0.348247	384.93	3.54	160.413
2,2-dimetilheksan	21.75	106.840	151.97	279.00	25.60	103.42	0.6953	67.7	8.915	-53.71	2.56	77.4	39.2525	0.339426	405.11	3.65	164.289
2,3,4-trimetilpentan	20.67	113.467	163.90	295.00	27.60	102.39	0.7191	68.37	9.014	-51.97	4.32	95.9	38.8681	0.317422	368.93	3.48	158.851
2,2,4-trimetilpentan	24.50	99.238	165.80	271.15	25.50	104.09	0.6919	64.87	8.402	-53.57	3.13	100.0	39.2617	0.305370	392.19	3.58	165.096
2,2,3,3-tetrametilbutan	26.50	106.470	373.80	270.80	24.50	93.06	0.8242	66.2	8.410	-53.99	4.88	/	/	0.255294	390.47	3.62	138.598
2,3,3-trimetilpentan	22.25	114.760	172.22	303.00	29.00	102.06	0.7262	68.1	8.897	-51.73	4.52	99.4	38.7617	0.293177	377.4	3.58	157.298
2,2,3-trimetilpentan	22.83	109.841	160.89	294.00	28.20	101.31	0.7161	67.3	8.826	-52.61	4.09	99.9	38.9249	0.300816	371.75	3.58	159.517
2-metil-3-etilpentan	18.00	115.450	158.20	295.00	27.40	106.06	0.7193	69.7	9.209	-50.48	5.08	88.1	38.8362	0.332433	368.10	3.54	158.807
2-metilheptan	17.33	117.647	164.16	288.00	24.80	109.84	0.6980	70.3	9.484	-51.50	3.06	23.1	39.2316	0.377916	407.85	3.61	163.653

3-etil-heksan	16.00	118.534	/	292.00	25.74	109.43	0.7136	71.70	9.476	-50.40	3.59	52.4	38.9441	0.362472	379.04	3.61	160.076
3-metil-3- etil-pentan	19.25	118.259	182.20	305.00	28.90	101.48	0.7274	69.30	9.081	-51.38	4.76	88.7	38.7171	0.306899	366.99	3.65	157.039
3-metil- heptan	16.67	118.925	152.60	292.00	25.60	111.26	0.7058	71.30	9.521	-50.82	3.29	35.0	39.1001	0.371002	397.34	3.61	161.845
4-metil- heptan	16.67	117.709	152.00	290.00	25.60	109.32	0.7046	70.91	9.483	-50.69	4.00	39.0	39.1174	0.371504	396.04	3.61	162.120
oktan	15.00	125.665	216.30	296.20	24.64	111.67	0.7025	73.19	9.915	-49.82	4.14	/	39.1922	0.397898	415.30	3.67	162.605

Legenda:

BP: vrelišče

MP: tališče

C_T : toplotna kapaciteta pri konstantni temperaturi

C_p : toplotna kapaciteta pri konstantnem tlaku

S: entropija

ρ : gostota

ΔH_{vap} : izparilna entalpija

ΔH_{vap}° : standardna izparilna entalpija

ΔH_{tv} : tvorben entalpija

ΔH_{tv}° : standardna tvorben entalpija

MON: oktansko število

LogP: koeficient particije med oktanolom in vodo

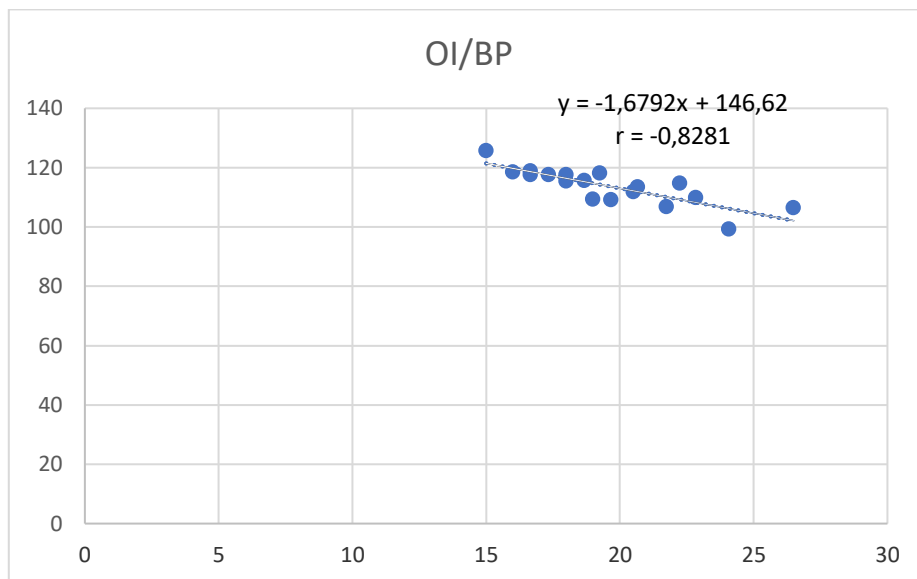
ω : acentrični faktor

TSA: skupna ploščina

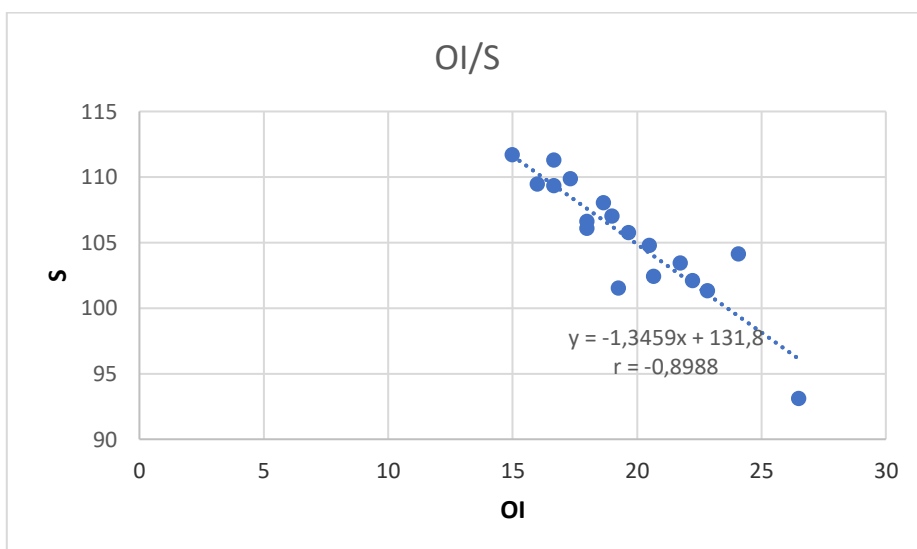
V_m : molski volumen

A: molarna refrakcija

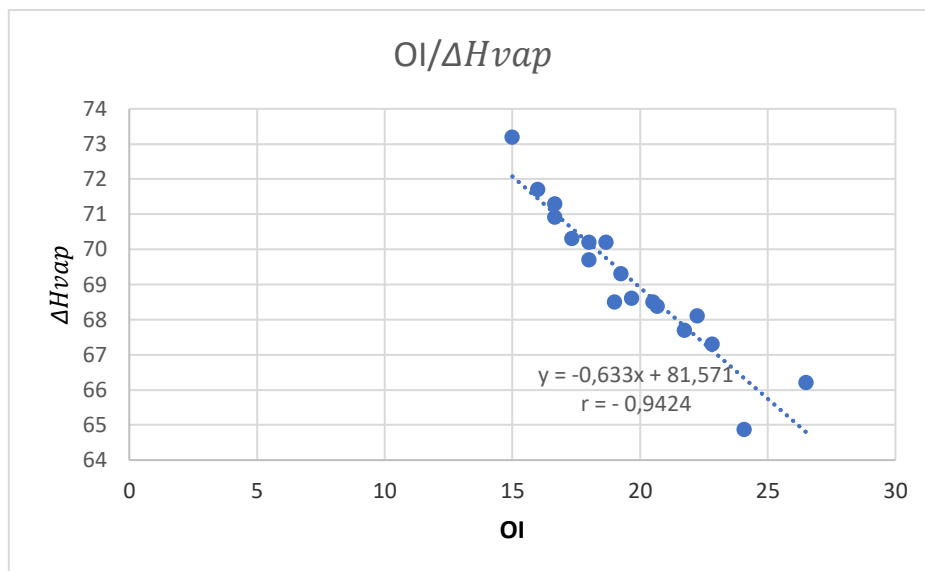
Nadaljnji postopek je bil enak. V Microsoft Excelu sem narisala grafe, ki predstavljajo asociacijo med Ohridskim indeksom in fizikalno-kemijsko lastnostjo. Dodala sem regresijsko premico ter poiskala korelacije s pomočjo korelacijskega koeficienta. Spodaj so podani le grafi, kjer je bila najdena korelacija.



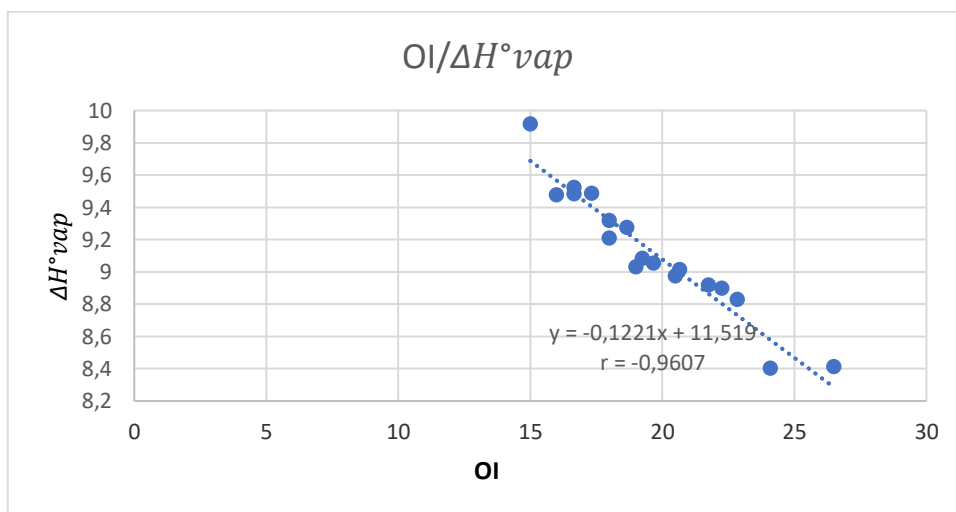
GRAF 11:Korelacija med OI in vreliščem



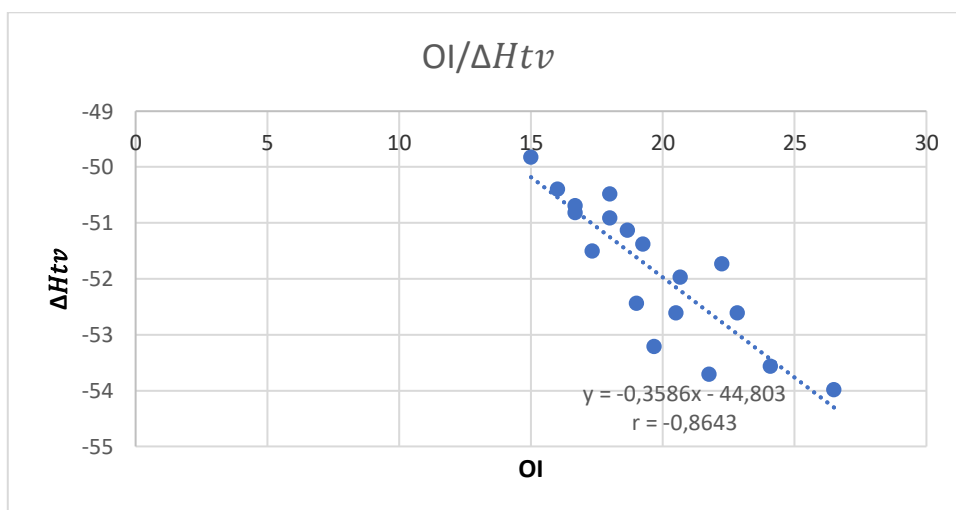
GRAF 12: Korelacija med OI in entropijo



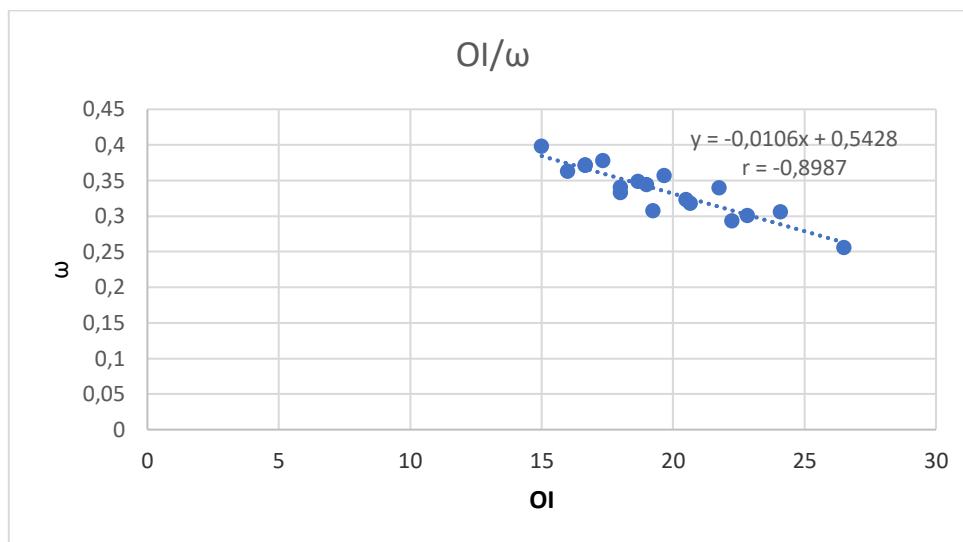
GRAF 13: Korelacija med OI in ΔH_{vap}



GRAF 14: Korelacija med OI in ΔH_{vap}°



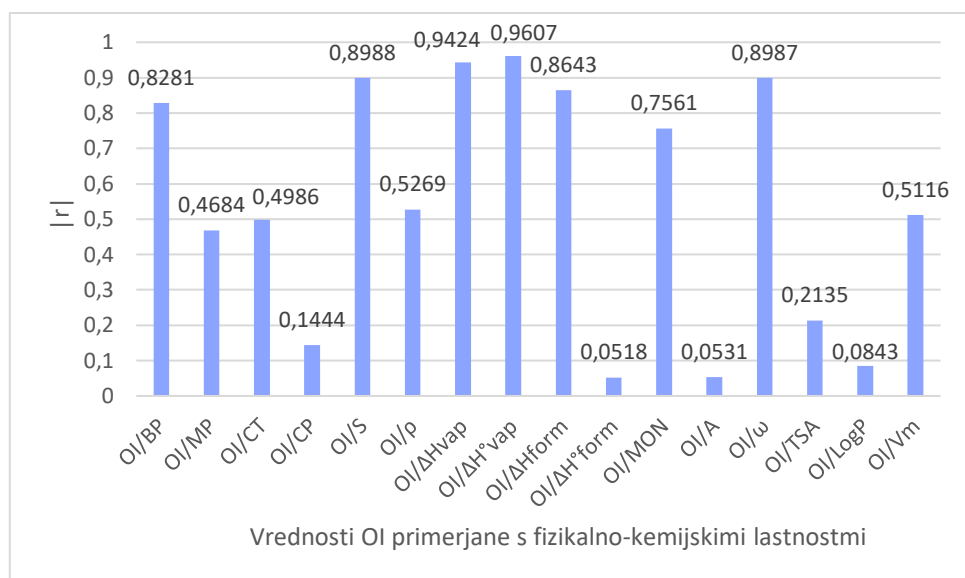
GRAF 15: Korelacija med OI in tvorbeno entalpijo



GRAF 16: Korelacija med OI in acentričnim faktorjem

TABELA 6: Korelacijski koeficienti med vrednostmi Ohridskega indeksa in fizikalno-kemijskimi lastnostmi izomerov oktana

Indeks/fizikalno-kemijska lastnost	Korelacija (r)
OI/MP	0.4684
OI/ C_T	-0.4986
OI/ C_p	0.1444
OI/ ρ	0.5269
OI/LogP	-0.0843
OI/ ΔH_{tv}°	0.0518
OI/MON	0.7561
OI/A	0.0028
OI/TSA	-0.5312
OI/ V_m	-0.5116



GRAF 17: Absolutne vrednosti korelacijskega koeficienta, dobljene pri primerjavi vrednosti OI s fizikalno-kemijskimi lastnostmi izomerov oktana

Moč korelacij se v primeru fizikalno-kemijskih lastnosti zelo razlikuje. Najvišja korelacija je bila najdena s standardno izparilno entalpijo (0.9607). Prav tako so visoke korelacije pokazale izparilna entalpija (0.9424), vrelišče (0.8281), entropija (0.8988), acentrični faktor (0.8987) in tvorben entalpija (0.8643). Ohridski indeks zmerno korelira z oktanskim številom. Z ostalimi fizikalno-kemijskimi lastnostmi ni bilo najdene korelacije. Eden izmed Randićevih atributov dobrega topološkega indeksa je dobra korelacija z vsaj eno lastnostjo; glede na to da Ohridski indeks dobro korelira s kar štirimi fizikalno-kemijskimi lastnostmi, lahko zaključimo, da Ohridski indeks izpolni ta pogoj.

3.4. RAČUNANJE VREDNOSTI OHRIDSKEGA INDEKSA ZA POLIAROMATSKE OGLJIKOVODIKE

Kot nadgradnjo smo celotni postopek ponovili še na kompleksnejši skupini molekul, poliaromatskih ogljikovodikih z 82 različnimi molekulami. Ker se struktura teh molekul postopno spreminja bodo služile, kot dober indikator, ali Ohridski indeks dobro odraža te spremembe. Za začetek sem izračunane vrednosti OI ponovno primerjala z že uveljavljenimi indeksi.

3.4.1. PRIMERJAVA VREDNOSTI OHRIDSKEGA INDEKSA Z DRUGIMI TOPOLOŠKIMI INDEKSI ZA POLIAROMATSKE OGLJIKOVODIKE

V spodnji tabeli so zbrane izračunane verzije Ohridskega indeksa za posamezni poliaromatski ogljikovodik ter vrednosti ostalih topoloških indeksov za iste molekule. Vključeni so tudi grafi,

kjer je bila najdena dobra korelacija. Vrednosti drugih indeksov so vzete iz baze podatkov *The International Academy of Mathematical Chemistry*.

Tokrat smo v primerjavo vključili še 3 dodatne topološke indekse:

Dz: Poglianijev indeks (Dz) je vsota modificiranih stopenj vozlišč, ki so definirane kot razmerje med številom valenčnih elektronov in glavnim kvantnim številom vozlišča oz. atoma (Pogliani, 1996).

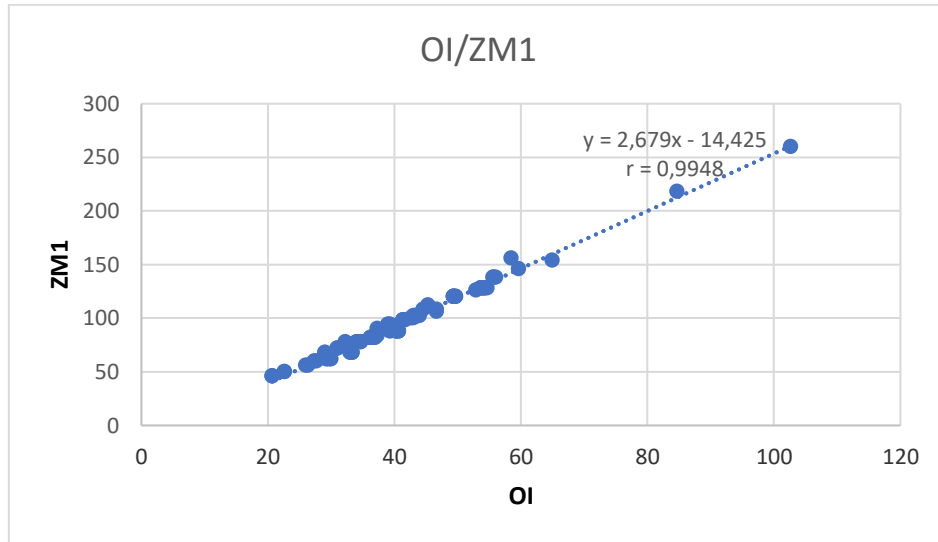
ZM1V, ZM2V: Če v definicijah prvega in drugega Zagrebškega indeksa stopnjo vozlišča zamenjamo z valenčno stopnjo vozlišča, ki je enaka razliki števila valenčnih elektronov vozlišča u in števila vodikovih atomov vezanih na u , dobimo prvi in drugi Zagrebški valenčni indeks, ki ju označimo ZM1V in ZM2V.

TABELA 7: Vrednosti topoloških deskriptorjev za poliaromatske ogljikovodike

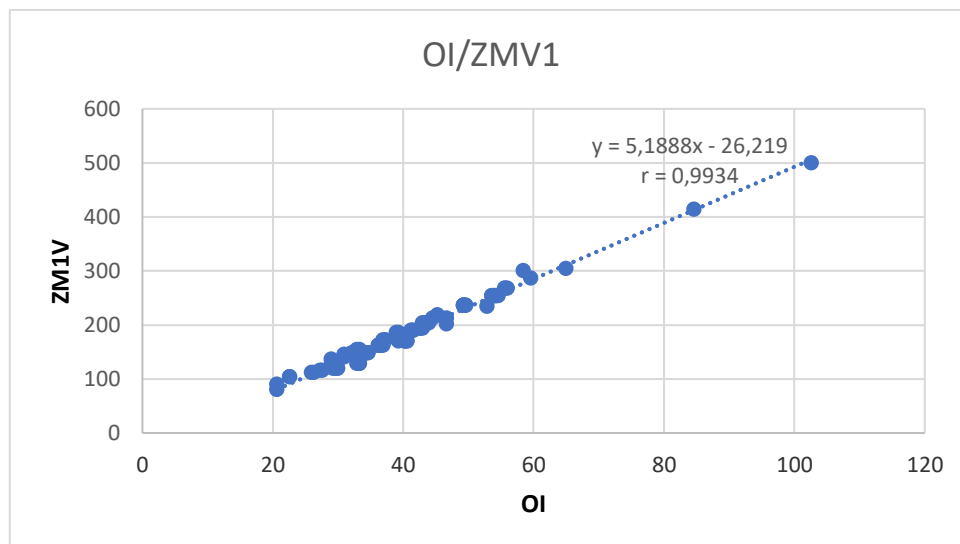
Molekula	OI	ZM 1	ZM1 V	ZM 2	ZM 2V	Qindex	SNar	HNar	GNar	Xt	Dz	Ram
naftalen	22.6667	50	104	57	118	8	7.742	2.143	2.169	0.359	20	2
1-metilnaftalen	26.0000	56	112	65	129	9	8.148	2.000	2.097	0.350	22	3
2-metilnaftalen	26.3333	56	112	64	128	9	8.148	2.000	2.097	0.350	22	3
1-etilnaftalen	27.3333	60	116	70	135	9	8.841	2.000	2.089	0.336	24	3
2-etilnaftalen	27.6667	60	116	69	134	9	8.841	2.000	2.089	0.336	24	3
2,6-dimetilnaftalen	30.0000	62	120	71	138	10	8.553	1.895	2.04	0.342	24	4
2,7-dimetilnaftalen	30.0000	62	120	71	138	10	8.553	1.895	2.04	0.342	24	4
1,7-dimetilnaftalen	29.6667	62	120	72	139	10	8.553	1.895	2.04	0.342	24	4
1,5-dimetilnaftalen	29.3333	62	120	73	140	10	8.553	1.895	2.04	0.342	24	4
1,2-dimetilnaftalen	29.3333	62	120	73	140	10	8.553	1.895	2.04	0.342	24	4
1,3,7-trimetilnaftalen	33.3333	68	128	79	149	11	8.959	1.814	1.992	0.334	26	5
2,3,5-trimetilnaftalen	33.0000	68	128	80	150	11	8.959	1.814	1.992	0.334	26	5
2,3,6-trimetilnaftalen	33.3333	68	128	79	149	11	8.959	1.814	1.992	0.334	26	5
fenalen	31.0000	72	145	87	174	13	10.633	2.229	2.266	0.307	26	4
1-fenilnaftalen	37.0000	84	172	99	201	13	12.712	2.182	2.213	0.280	32	4
2-fenilnaftalen	37.3333	84	172	98	200	13	12.712	2.182	2.213	0.280	32	4
antracen	33.3333	76	154	90	182	13	11.326	2.211	2.246	0.297	28	4
1-metilantracen	36.6667	82	162	98	193	14	11.731	2.093	2.186	0.292	30	5
2-metilantracen	37.0000	82	162	97	192	14	11.731	2.093	2.186	0.292	30	5
2,7-dimetilantracen	40.6667	88	170	104	202	15	12.137	2.000	2.135	0.287	32	6
2,6-dimetilantracen	40.6667	88	170	104	202	15	12.137	2.000	2.135	0.287	32	6
2,3-dimetilantracen	40.3333	88	170	105	203	15	12.137	2.000	2.135	0.287	32	6
9,10-dimetilantracen	39.3333	88	170	108	206	15	12.137	2.000	2.135	0.287	32	6
fenantren	33.0000	76	154	91	183	13	11.326	2.211	2.246	0.297	28	4
1-metilfenantren	36.3333	82	162	99	194	14	11.731	2.093	2.186	0.292	30	5
2-metilfenantren	36.6667	82	162	98	193	14	11.731	2.093	2.186	0.292	30	5
3-metilfenantren	36.6667	82	162	98	193	14	11.731	2.093	2.186	0.292	30	5

4-metilfenantren	36.3333	82	162	99	194	14	11.731	2.093	2.186	0.292	30	5
9-metilfenantren	36.2500	82	162	99	194	14	11.731	2.093	2.186	0.292	30	5
3,6-dimetilfenantren	40.3333	88	170	105	203	15	12.137	2.000	2.135	0.287	32	6
4,5-metilenfenantren	37.3333	90	172	113	213	18	12.83	2.308	2.352	0.279	30	6
tetracen	44.0000	102	204	123	246	18	14.909	2.250	2.289	0.259	36	6
benzo[a]antracen	43.6667	102	204	124	247	18	14.909	2.250	2.289	0.259	36	6
krizen	43.3333	102	204	125	248	18	14.909	2.250	2.289	0.259	36	6
benzo[c]fenantren	43.3333	102	204	125	248	18	14.909	2.250	2.289	0.259	36	6
trifenilen	43.0000	102	204	126	249	18	14.909	2.250	2.289	0.259	36	6
piren	39.3333	94	186	117	230	18	13.523	2.286	2.328	0.272	32	6
1-metilpiren	42.6667	100	194	125	241	19	13.929	2.170	2.269	0.268	34	7
2-metilpiren	43.0000	100	194	124	240	19	13.929	2.170	2.269	0.268	34	7
4-metilpiren	42.6667	100	194	125	241	19	13.929	2.170	2.269	0.268	34	7
2,7-dimetilpiren	46.6667	106	202	131	250	20	14.334	2.077	2.217	0.264	36	8
pentacen	54.6667	128	254	156	310	23	18.493	2.276	2.318	0.233	44	8
dibenzo[ai]antracen	54.3333	128	254	157	311	23	18.493	2.276	2.318	0.233	44	8
dibenzo[ah]antracen	54.0000	128	254	158	312	23	18.493	2.276	2.318	0.233	44	8
dibenzo[aj]antracen	54.0000	128	254	158	312	23	18.493	2.276	2.318	0.233	44	8
benzo[b]krizen	54.0000	128	254	158	312	23	18.493	2.276	2.318	0.233	44	8
dibenzo[ac]antracen	53.6667	128	254	159	313	23	18.493	2.276	2.318	0.233	44	8
picen	53.6667	128	254	159	313	23	18.493	2.276	2.318	0.233	44	8
benzo[a]piren	49.6667	120	236	151	295	23	17.107	2.308	2.352	0.242	40	8
benzo[e]piren	49.3333	120	236	152	296	23	17.107	2.308	2.352	0.242	40	8
perilen	49.3333	120	236	152	296	23	17.107	2.308	2.352	0.242	40	8
koronen	58.4583	156	300	204	390	33	21.501	2.400	2.449	0.216	48	12
antrantren	56.0000	138	268	177	342	28	19.304	2.357	2.405	0.228	44	10
benzo[ghi]perilen	55.6667	138	268	178	343	28	19.304	2.357	2.405	0.228	44	10
dibenzo[ae]piren	59.6667	146	286	186	361	28	20.69	2.323	2.368	0.220	48	10
1-metilkrizen	46.6667	108	212	133	259	19	15.315	2.151	2.239	0.256	38	7
6-metilkrizen	44.5833	108	212	133	259	19	15.315	2.151	2.239	0.256	38	7
3-metilkolantren	52.9167	126	234	159	293	24	17.512	2.211	2.302	0.239	42	9
indeno[1,2,3-cd]piren	55.5833	138	268	178	343	28	19.304	2.357	2.405	0.228	44	10
pentafen	54.3333	128	254	157	311	23	18.493	2.276	2.318	0.233	44	8
heksafen	65.0000	154	304	190	375	28	22.076	2.294	2.338	0.213	52	10
indan	20.6667	46	80	53	91	8	7.049	2.160	2.189	0.377	18	2
inden	20.6667	46	90	53	102	8	7.049	2.160	2.189	0.377	18	2
azulen	22.6667	50	104	57	118	8	7.742	2.143	2.169	0.359	20	2
acenaften	29.0000	68	126	83	152	13	9.940	2.250	2.289	0.317	24	4
acenaftilen	29.0000	68	136	83	165	13	9.940	2.250	2.289	0.317	24	4
fluoren	31.0000	72	140	87	166	13	10.633	2.229	2.266	0.307	26	4
1-metilfluoren	34.3333	78	148	95	177	14	11.038	2.100	2.200	0.301	28	5
2-metilfluoren	34.6667	78	148	94	176	14	11.038	2.100	2.200	0.301	28	5
3-metilfluoren	34.6667	78	148	94	176	14	11.038	2.100	2.200	0.301	28	5
4-metilfluoren	32.2500	78	148	95	177	14	11.038	2.100	2.200	0.301	28	5
9-metilfluoren	34.0000	78	146	96	177	14	11.038	2.100	2.200	0.301	28	5
1,2-benzofluoren	41.3333	98	190	121	231	18	14.216	2.267	2.308	0.265	34	6
fluoranthene	39.0000	94	186	118	231	18	13.523	2.286	2.328	0.272	32	6
2,3-benzofluoren	41.6667	98	190	120	230	18	14.216	2.267	2.308	0.265	34	6
3,4-benzofluoren	41.3333	98	190	121	231	18	14.216	2.267	2.308	0.265	34	6

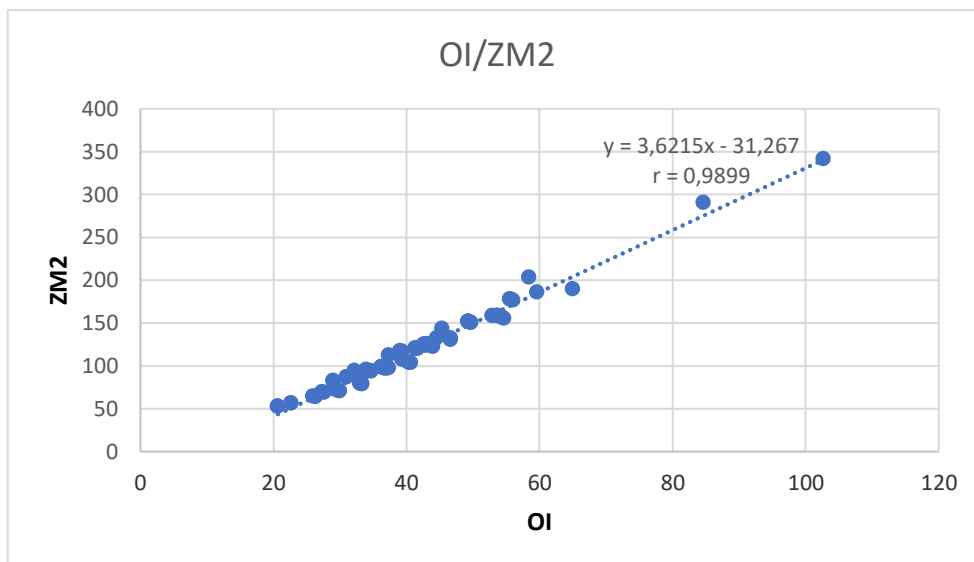
benzo[ghi]fluoranten	45.3333	112	218	144	278	23	15.720	2.348	2.395	0.252	36	8
benzo[k]fluoranten	49.6667	120	236	151	295	23	17.107	2.308	2.352	0.242	40	8
benzo[b]fluoranten	49.3333	120	236	152	296	23	17.107	2.308	2.352	0.242	40	8
benzo[j]fluoranten	49.3333	120	236	152	296	23	17.107	2.308	2.352	0.242	40	8
ovalen	84.6667	218	414	291	550	48	29.479	2.462	2.512	0.184	64	18
kvaterilen	102.6667	260	500	342	652	53	35.835	2.400	2.449	0.167	80	20



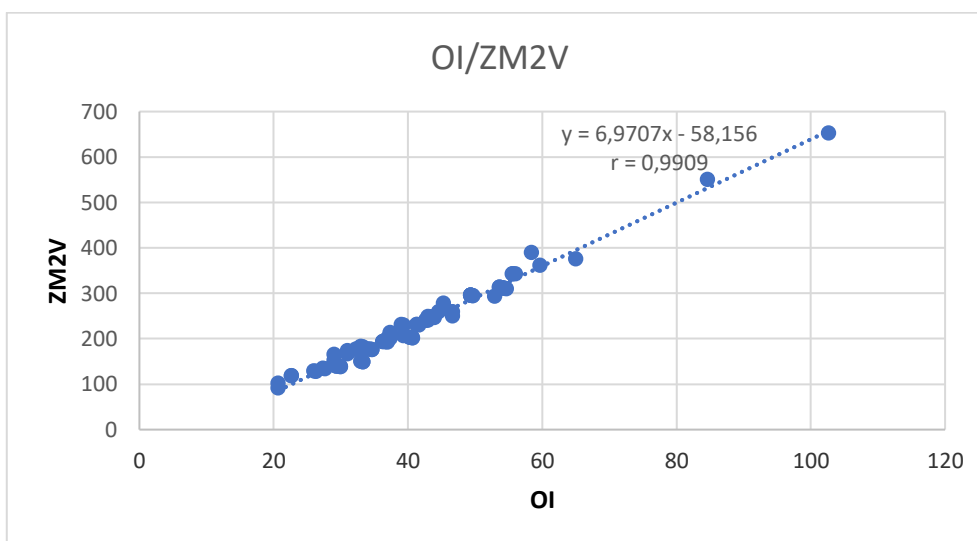
GRAF 18: Korelacija med OI in ZM1



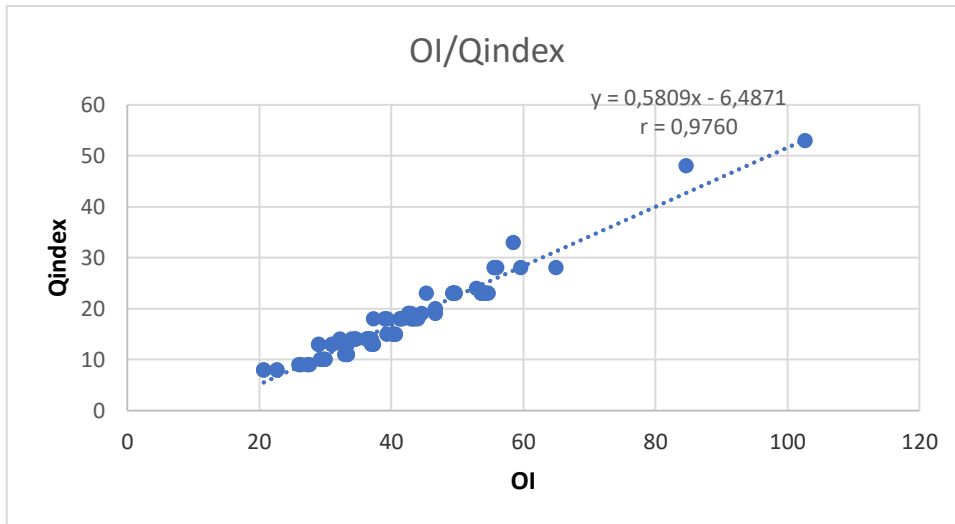
GRAF 19: Korelacija med OI in ZMV1



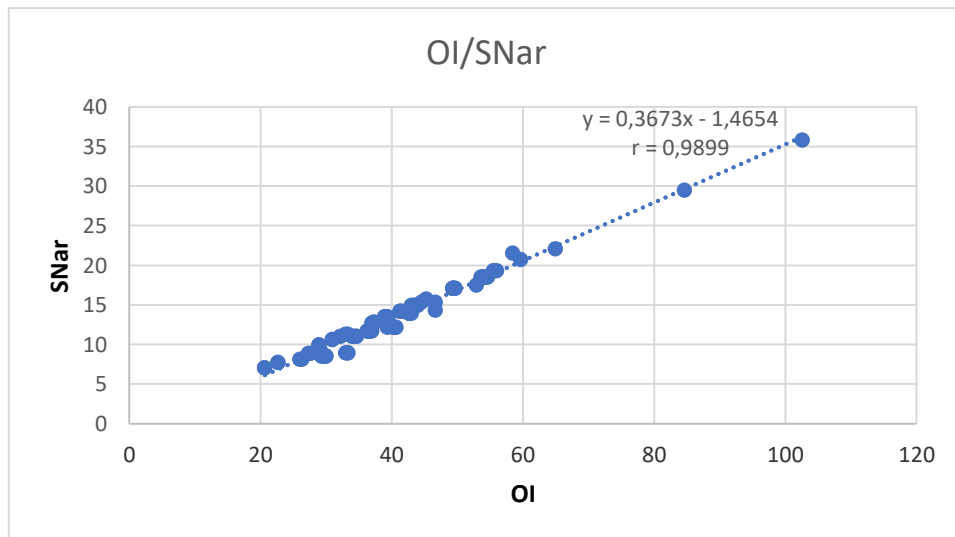
GRAF 20: Korelacija med OI in ZM2



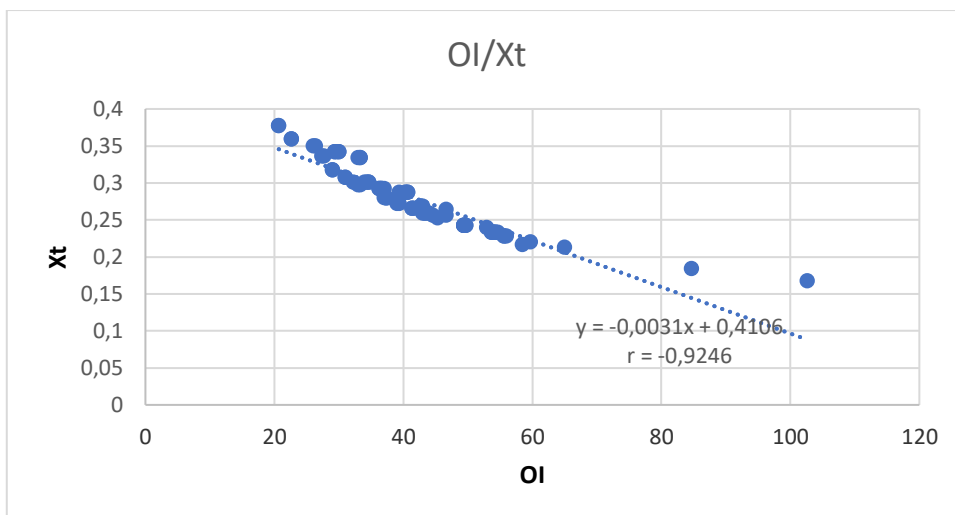
GRAF 21: Korelacija med OI in ZM2V



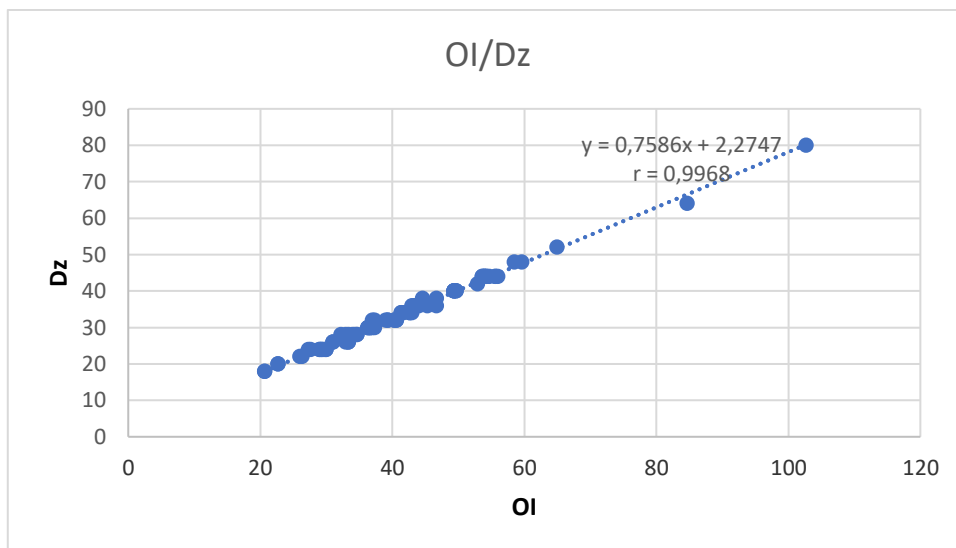
GRAF 22: Korelacija med OI in Qindex-om



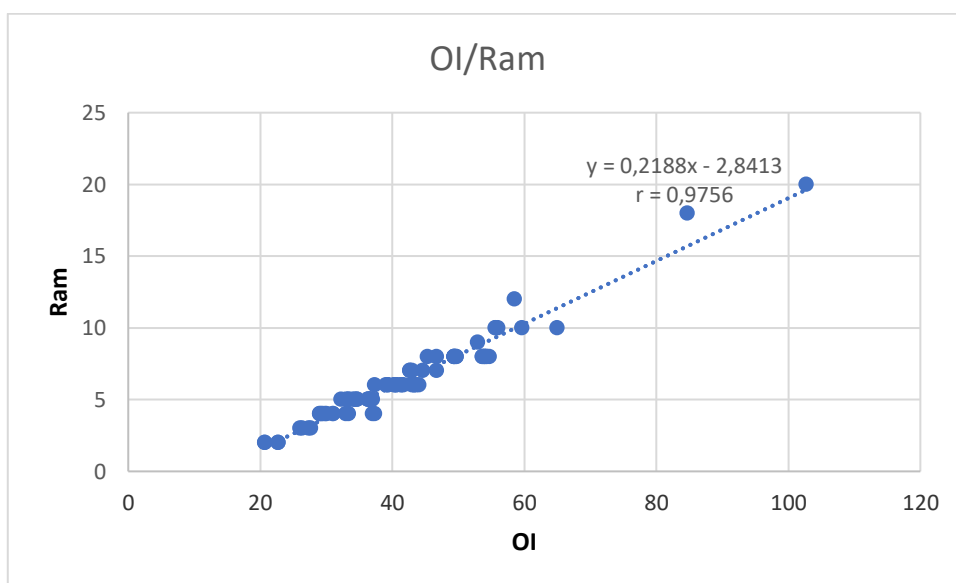
GRAF 23: Korelacija med OI in SNar



GRAF 24: Korelacija med OI in Xt

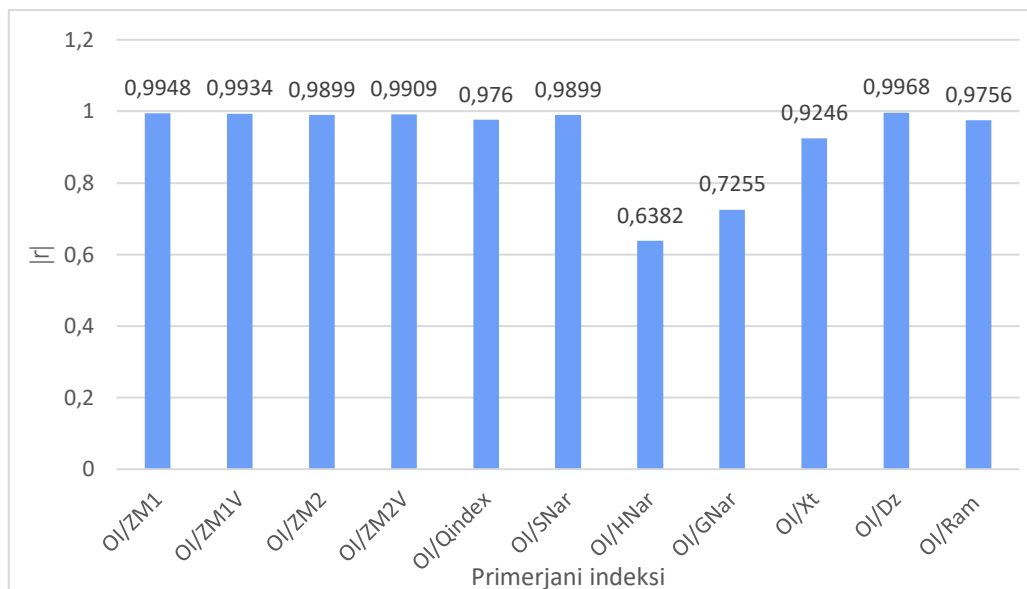


GRAF 25: Korelacija med OI in Dz



GRAF 26: Korelacija med OI in Ram

Kot je razvidno iz grafov, so korelacije med Ohridskim indeksom ter drugimi topološkimi deskriptorji ponovno izredno visoke. Najvišje, skoraj linearne korelacije so bile najdene s Prvim in Drugim Zagrebškim indeksom, kot tudi z njunima valenčnima različicama ter z Narumijevim enostavnim topološkim indeksom. Nekoliko slabše korelacije so bile najdene le s harmonično in geometrično različico Narumijevega indeksa. Za boljšo predstavo so vrednosti korelacijskega koeficienta prikazane v stolpčnem diagramu spodaj.



GRAF 27: Absolutne vrednosti korelacijskega koeficienta, dobljene pri primerjavi vrednosti Ohridskega indeksa z vrednostmi drugih topoloških deskriptorjev za poliaromatske ogljikovodike

3.4.2. KORELACIJA OHRIDSKEGA INDEKSA Z IZBRANIMI FIZIKALNO-KEMIJSKIMI LASTNOSTMI POLIAROMATSKIH OGLJIKOVODIKOV

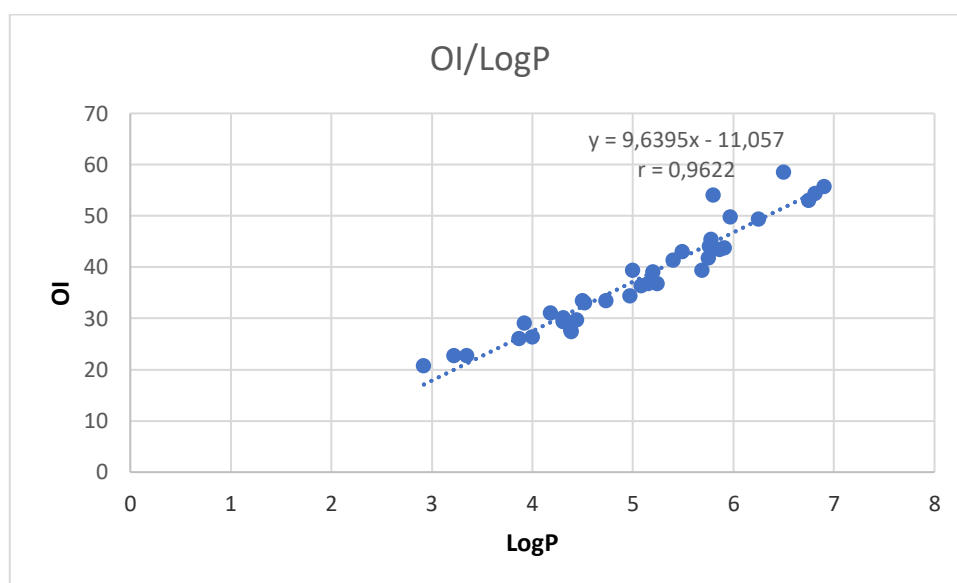
Zaradi slabe dostopnosti mi je uspelo pridobiti podatke le za 3 fizikalno-kemijske lastnosti. Podatki so vzeti iz baze podatkov *The International Academy of Mathematical Chemistry*. Manjkajoči podatki so označeni z /.

TABELA 8: Vrednosti izbranih fizikalno-kemijskih lastnosti poliaromatskih ogljikovodikov

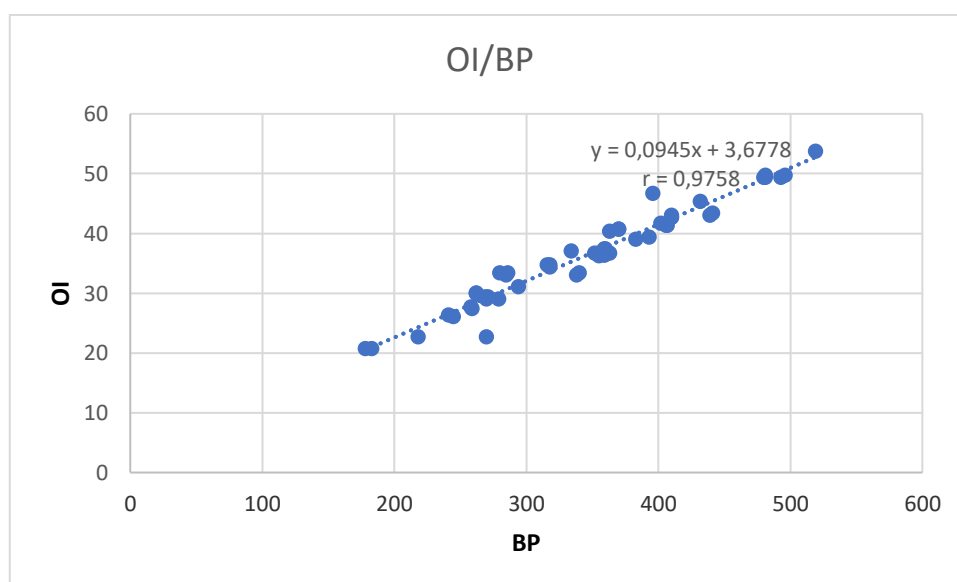
Molekula	OI	MP	BP	LogP
naftalen	22.6667	81	218	3.35
1-metilnaftalen	26.0000	-22	245	3.87
2-metilnaftalen	26.3333	35	241	4.00
1-ethylnaftalen	27.3333	-14	259	4.39
2-ethylnaftalen	27.6667	-7	258	4.38
2,6-dimetilnaftalen	30.0000	110	262	4.31
2,7-dimetilnaftalen	30.0000	97	262	/
1,7-dimetilnaftalen	29.6667	-14	263	4.44
1,5-dimetilnaftalen	29.3333	80	269	4.31
1,2-dimetilnaftalen	29.3333	-4	271	4.31
1,3,7-trimetilnaftalen	33.3333	14	280	/
2,3,5-trimetilnaftalen	33.0000	25	285	/
2,3,6-trimetilnaftalen	33.3333	101	286	4.73
fenalen	31.0000	85	/	/
1-fenilnaftalen	37.0000	45	334	/
2-fenilnaftalen	37.3333	104	360	/
antracen	33.3333	216	340	4.50
1-metilantracen	36.6667	86	363	/
2-metilantracen	37.0000	209	/	/
2,7-dimetilantracen	40.6667	241	370	/

2,6-dimetilantracen	40.6667	250	370	/
2,3-dimetilantracen	40.3333	252	/	/
9,10-dimetilantracen	39.3333	183	/	5.69
fenantren	33.0000	101	338	4.52
1-metilfenantren	36.3333	123	359	5.08
2-metilfenantren	36.6667	56	355	5.24
3-metilfenantren	36.6667	65	352	5.15
4-metilfenantren	36.3333	50	/	/
9-metilfenantren	36.2500	91	355	/
3,6-dimetilfenantren	40.3333	141	363	/
4,5-metilenefenantren	37.3333	116	359	/
tetracen	44.0000	257	/	5.76
benzo[a]antracen	43.6667	162	/	5.91
krizen	43.3333	256	441	5.86
benzo[c]fenantren	43.3333	68	/	/
trifenilen	43.0000	199	439	5.49
piren	39.3333	156	393	5.00
1-metilpiren	42.6667	70	410	/
2-metilpiren	43.0000	144	410	/
4-metilpiren	42.6667	148	410	/
2,7-dimetilpiren	46.6667	/	396	/
pentacen	54.6667	271	/	/
dibenzo[ai]antracen	54.3333	264	/	6.81
dibenzo[ah]antracen	54.0000	270	/	5.80
dibenzo[aj]antracen	54.0000	198	/	/
benzo[b]krizen	54.0000	294	/	/
dibenzo[ac]antracen	53.6667	205	/	/
picen	53.6667	/	519	/
benzo[a]piren	49.6667	177	496	5.97
benzo[e]piren	49.3333	179	493	/
perilen	49.3333	278	/	6.25
koronen	58.4583	360	/	6.50
antrantren	56.0000	261	/	/
benzo[ghi]perilen	55.6667	283	/	6.90
dibenzo[ae]piren	59.6667	234	/	/
1-metilkrizen	46.6667	161	/	/
6-metilkrizen	44.5833	257	/	/
3-metilcholanthrene	52.9167	180	/	6.75
indeno[1,2,3-cd]piren	55.5833	163	/	/
pentafen	54.3333	263	/	/
heksafen	65.0000	308	/	/
indan	20.6667	-51	178	/
inden	20.6667	-2	183	2.92
azulene	22.6667	100	270	3.22
acenaften	29.0000	96	279	3.92
acenaftilen	29.0000	93	270	/
fluoren	31.0000	117	294	4.18
1-metilfluoren	34.3333	87	318	4.97
2-metilfluoren	34.6667	104	318	/
3-metilfluoren	34.6667	88	316	/
4-metilfluoren	32.2500	71	/	/

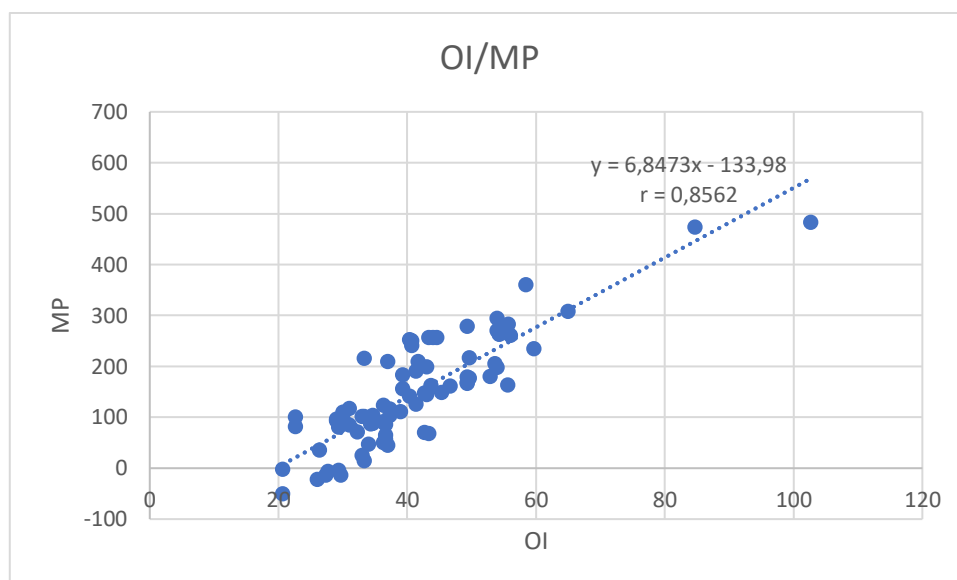
9-metilfluoren	34.0000	47	/	/
1,2-benzofluoren	41.3333	190	407	5.40
fluoranten	39.0000	111	383	5.20
2,3-benzofluoren	41.6667	209	402	5.75
3,4-benzofluoren	41.3333	125	406	/
benzo[ghi]fluoranten	45.3333	149	432	5.78
benzo[k]fluoranten	49.6667	217	481	/
benzo[b]fluoranten	49.3333	168	481	/
benzo[j]fluoranten	49.3333	166	480	/
ovalen	84.6667	473	/	/
kvaterilen	102.6667	483	/	/



GRAF 28: Korelacija med OI in LogP



GRAF 29: Korelacija med OI in vreliščem



GRAF 30: Korelacija med OI in tališčem

Tudi s fizikalno-kemijskimi lastnostmi poliaromatskih ogljikovodikov in vrednostmi Ohridskega indeksa so bile najdene odlične korelacije, še posebej z vreliščem (0.9758) ter logaritemsko vrednostjo koeficienta particije¹ (0.9622). Zmerna korelacija je bila najdena s tališčem.

4 ZAKLJUČEK

Glavni cilj moje raziskovalne naloge je bil raziskati Ohridski indeks, teoretični parameter za opis grafov, oziroma topološki indeks v QSPR analizi izomerov oktana ter poliaromatskih ogljikovodikov.

Glede na najdene korelacije z že uveljavljenimi indeksi tako v primeru izomerov oktana kot poliaromatskih ogljikovodikov lahko zaključimo, da je Ohridski indeks zares obetaven. V primeru oktanovih izomerov sem zabeležila najmočnejšo korelacijo z ZM1 in Qindex-om, ki je le normalizacija prvega Zagrebškega indeksa, zato je korelacija v obeh primerih 0.9724. Kar se tiče poliaromatskih ogljikovodikov, so bile korelacije z drugimi topološkimi indeksi generalno še višje kot pri izomerih oktana. Najmočnejša, skoraj linearna odvisnost je bila najdena s Poglianijevim indeksom, in sicer, 0.9968.

Nekaj zelo dobrih korelacij je bilo najdenih tudi s fizikalno-kemijskimi lastnostmi. Pri izomerih oktana je Ohridski indeks najboljše koreliral z entropijo (-0.8988), izparilno entalpijo (-0.9424),

¹ To je distribucijski koeficient, ki predstavlja razmerje koncentracij spojin v dveh fazah mešanice dveh topil, ki se ne mešata med seboj.

standardno izparilno entalpijo (-0.9607) ter acentričnim faktorjem (-0.8987). V primeru poliaromatskih ogljikovodikov pa z vreliščem (0.9758) ter koeficientom particije (0.9622).

Če se vrnemo k Randićevim atributom dobrega topološkega indeksa (direktna interpretacija strukture, dobra korelacija z vsaj eno fizikalno ali kemijsko lastnostjo, zmožnost dobrega ločevanja med izomeri, temeljenje na strukturi molekule in ne na fizikalno-kemijskih lastnostih, postopne strukturne spremembe se odražajo v postopnih spremembah vrednosti) vidimo, da Ohridski indeks izpolnjuje vse pogoje. Ohridski indeks ni izpeljan iz fizikalne ali kemijske lastnosti, temveč temelji na strukturi. Iz opazovanih korelacij s fizikalno-kemijskimi lastnostmi Ohridski indeks dobro korelira s kar osmimi. Glede na rezultate iz testiranja na izomerih oktana vidimo, da Ohridski indeks omogoča dobro razločevanje med izomeri, kot tudi dobro napoved nekaterih njihovih lastnosti. Ohridski indeks je relativno preprosto izračunati, še posebej če uporabimo računalniški program. Nenazadnje, tako kot struktura v primeru cikloalkanov ter poliaromatskih ogljikovodikov, se tudi vrednosti Ohridskega indeksa postopno spreminjajo, kar pomeni, da je Ohridski indeks občutljiv že za male spremembe v strukturi. Torej je Ohridski indeks zares uporaben topološki indeks. Problem nastopi pri heteroatomnih molekulah, torej molekulah, ki niso sestavljene samo iz ogljika in vodika, saj Ohridski indeks ne poda eksplicitne razlage, kako ga uporabiti na molekulah, ki vključujejo atome drugih elementov. V tej nalogi smo v izbranih molekulah stopnjo vozlišča določili le atomom ogljika, zato ni jasno, če moramo z atomi drugih elementov ravnati enako kot z ogljikom ali vodikom. Posledično bi Ohridski indeks potreboval dodelave, če bi ga želeli v prihodnje uporabiti na še kompleksnejših skupinah molekul z večjim številom elementov.

Vseeno pa topološki indeksi in QSPR predstavljajo velik potencial za prihodnost. Napredne tehnologije že sedaj generirajo ogromne količine podatkov o učinku kemikalij, QSAR in QSPR se že uporabljata v razvoju novih zdravil. Glede na to da okoljska trajnost postaja vse bolj pomembna, lahko pričakujemo zmeraj večjo potrebo po uporabi modernih načinov napovedovanja lastnosti s QSP(A)R, ki omogoča ogromne prihranke materiala, energije in delovne sile. Morda z nekaj dodelavami tudi Ohridski indeks nekoč prinese uporabno rešitev za napoved molekulskih lastnosti. Do tedaj bo pa zagotovo ostal zanimiv predmet dodatnih raziskav.

5 VIRI

Basak, S., Magnuson, V., Niemi, G., Regal, R., & Veith, G. (1987). Topological indices: their nature, mutual relatedness, and applications. *Mathematical Modelling*, 8, p. 300–305.

Basak, S. C. *Curr. Comput. Aided Drug Des.* 2013, 9, p. 449.

Basak, S. C. Use of Graph Invariants in Quantitative Structure-Activity Relationship Studies. *Croat. Chem. Acta* 2016, 89(4), p. 419–429

Burch, K. J. (2019). Chemical applications of graph theory. *Mathematical Physics in Theoretical Chemistry*, p. 261–294.

Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., ... Tropsha, A. (2014). QSAR Modeling: Where Have You Been? Where Are You Going To? *Journal of Medicinal Chemistry*, 57, p. 4977–5010.

Einstein A. in *Ideas and Opinions by Albert Einstein*, (Ed. Carl Seelig), Crown Publishers, Inc., New York, 1954, p. 18–24

Gao, W., Wang, W., & Farahani, M. R. (2016). Topological Indices Study of Molecular Structure in Anticancer Drugs. *Journal of Chemistry*, 2016, p. 1–8.

Hawkins, D. M. Basak, S. C., Kraker, J. J., Geiss, K. T., Witzmann, F. A., J. (2016). *Advances in Mathematical Chemistry and Applications*, Vol. 2, Bentham Books, Elsevier

Parthasarathi, R., & Dhawan, A. (2018). In Silico Approaches for Predictive Toxicology. *In Vitro Toxicology*, p. 91–109.

Randić, M. (1991). Generalized molecular descriptors. *Journal of Mathematical Chemistry*, 7(1), p. 155–168.

Rouvray, D. H. (1973). The Search for Useful Topological Indices in Chemistry: Topological indices promise to have far-reaching applications in fields as diverse as bonding theory, cancer research, and drug design. *American Scientist*, 61(6), p. 729–735.

Sylvester, J. J. *Am. J. Math.* 1878, 1, p. 105

Todeschini, R., & Consonni, V. (2000). *Handbook of molecular descriptors*. Weinheim: Wiley-VCH.

Vilar, S., & Costanzi, S. (2012). Predicting the Biological Activities Through QSAR Analysis and Docking-Based Scoring. *Membrane Protein Structure and Dynamics*, p. 271–284.

Wazir, I. (2012). *Standard level mathematics: developed specifically for the Ib diploma*. Harlow, Essex: Pearson.

Wilson, R. J., & Watkins, J. J. (1990). *Graphs: an introductory approach*. New York: Wiley.

Wilson, R. J. (2015). *Introduction to graph theory*. Harlow, United Kingdom: Prentice Hall.

(n.d.). Retrieved from <http://www.iamc-online.org/members/index.htm>.

6 DODATEK

Računalniška koda za izračun Ohridskega indeksa (Črepnjak, 2019):

```

#include <iostream>
#include <fstream>
#include <string>
#include <stdlib.h>
#include <stdio.h>
#include <math.h>
#include <cmath>

using namespace std;

/*
void izpisiMatriko(int mat[dolP][dolP]){
    int i;
    int j;
    for(i=0;i<dolP;i++){
        for(j=0;j<dolP;j++){
            cout << mat[i][j] << " ";
        }
        cout << endl;
    }
    return;
}
*/

int main(){
    //indeksa za sprehajanje po matriki
    int i;
    int j;
    int s;
    string vrstica;
    string imeDat = "PAH-matrike";
    string imeDat1 = imeDat + ".txt";
    ifstream datM(imeDat1.c_str());
    string imeMolekule;
    fstream rezultati;
    string rezIme = "rezultati-Ohrid-pah.xml";
    rezultati.open(rezIme.c_str(), ios::app);

    while(!datM.eof()) {
        //preberemo ime molekule
        getline(datM,imeMolekule);
        //cout << imeMolekule << endl;

        //preberemo dolžino
        getline(datM,vrstica);
        int dolP = atoi(vrstica.c_str());;

        //      cout << dolP << endl;

        //ustvarimo matriko sosednosti
        int sosedi[dolP][dolP];
        for(i=0; i<dolP; i++){
            getline(datM,vrstica);
            for(j=0; j<dolP; j++){
                sosedi[i][j] = vrstica[2*j] - 48;
            }
        }
    }
}

```

```

    }
}

//
int deg[dolP];
for(i=0; i<dolP; i++){
    deg[i] = 0;
    for(j=0; j<dolP; j++){
        deg[i] += sosedi[i][j];
    }
}

/*
for(i=0; i < dolP; i++){
    cout << deg[i] << " ";
}
cout << endl << endl;
*/

long double Oindex = 0;
long double pomozni = 0;

for(i=0; i<dolP; i++){
    for(j=0; j<dolP; j++){
        if(sosedi[i][j]==1){
            pomozni = (double)deg[i]/(double)deg[j] +
(double)deg[j]/(double)deg[i];
            //sesteva po vrsti po povezavah
            Oindex += (pomozni);
        }
    }
}

Oindex = Oindex/2;

cout << "Ohrid index za " << imeMolekule << " je " << Oindex <<
endl;
//izpis rezultatov v datoteko
rezultati << imeMolekule << ";" << Oindex << endl;

//ta getline potrebujem, ker je v naši datoteki oktani med
dvema oktanoma ena vrstica prosta
getline(datM, imeMolekule);
}

datM.close();
rezultati.close();
return 0;
}

```