

# Kvantitativna analiza tramadola i paracetamola u tabletama spektroskopskim i kemometričkim metodama

---

Glavanović, Siniša

Doctoral thesis / Disertacija

2017

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:775030>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-03**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)





Sveučilište u Zagrebu  
PRIRODOSLOVNO-MATEMATIČKI FAKULTET

Siniša Glavanović

**KVANTITATIVNA ANALIZA  
TRAMADOLA I PARACETAMOLA U  
TABLETAMA SPEKTROSKOPSKIM I  
KEMOMETRIČKIM METODAMA**

DOKTORSKI RAD

Zagreb, 2017.





University of Zagreb  
FACULTY OF SCIENCE

Siniša Glavanović

**QUANTITATIVE ANALYSIS OF  
TRAMADOL AND PARACETAMOL IN  
TABLETS BY SPECTROSCOPIC AND  
CHEMOMETRIC METHODS**

DOCTORAL THESIS

Zagreb, 2017





Sveučilište u Zagrebu  
PRIRODOSLOVNO-MATEMATIČKI FAKULTET

Siniša Glavanović

**KVANTITATIVNA ANALIZA  
TRAMADOLA I PARACETAMOLA U  
TABLETAMA SPEKTROSKOPSKIM I  
KEMOMETRIČKIM METODAMA**

DOKTORSKI RAD

Mentori: prof. dr. sc. Vladislav Tomišić  
izv. prof. dr. sc. Marina Vranić

Zagreb, 2017.





University of Zagreb  
FACULTY OF SCIENCE

Siniša Glavanović

**QUANTITATIVE ANALYSIS OF  
TRAMADOL AND PARACETAMOL IN  
TABLETS BY SPECTROSCOPIC AND  
CHEMOMETRIC METHODS**

DOCTORAL THESIS

Supervisors: Prof. Vladislav Tomišić  
Assoc. Prof. Marina Vranić

Zagreb, 2017





Ova disertacija izrađena je u Belupu d.d. u Koprivnici, u Zavodu za fizikalnu kemiju Kemijskog odsjeka Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu i u Laboratoriju pokušališta Sveučilišta u Zagrebu Agronomskog fakulteta pod vodstvom prof. dr. sc. Vladislava Tomišića i izv. prof. dr. sc. Marine Vranić.



## *Zahvala*

*Zahvaljujem se mentorima izv. prof. dr. sc. Marini Vranić i prof. dr. sc. Vladislavu Tomišiću na pomoći, prijedlozima, savjetima i podršci tijekom izrade ovog rada.*

*Veliko hvala Stanislavu Biondiću, dipl. oec. i spec. Tatjani Šešok, mr. pharm., direktorima u Belupu, koji su mi omogućili pohađanje poslijediplomskog studija.*

*Hvala Sonji Betlehem-Bebek, rukovoditelju Analitičke dokumentacije i Jasminki Mikotić-Fulir, pomoćniku direktora, na pruženom razumijevanju i podršci.*

*Veliko hvala kolegicama Nevenki Betlehem-Požgaj, Ani Kroppek i Gordani Medenjak bez čije pomoći ne bi bilo ovog rada.*

*Također, zahvaljujem se na razumijevanju i pomoći svim kolegicama i kolegama iz Analitičkih laboratorija.*

*Mojim roditeljima veliko hvala na podršci, razumijevanju, odricanju i svemu što su mi omogućili.*

*Posebno se zahvaljujem supruzi Mariji i kćerkici Marti koje su mi bile motivacija kad je bilo najteže. Hvala na svemu!*

*Siniša*



# Sadržaj

<b>SAŽETAK .....</b>	<b>XVII</b>
<b>ABSTRACT.....</b>	<b>XIX</b>
<b>§ 1. UVOD .....</b>	<b>1</b>
<b>§ 2. LITERATURNI PREGLED .....</b>	<b>4</b>
<b>2.1. Regresija metodom djelomičnih najmanjih kvadrata .....</b>	<b>15</b>
<b>2.2. Genetički algoritmi .....</b>	<b>19</b>
2.2.1. <i>Kodiranje .....</i>	23
2.2.2. <i>Inicijalizacija populacije i evaluacija rješenja .....</i>	28
2.2.3. <i>Selekcija rješenja .....</i>	29
2.2.4. <i>Križanje.....</i>	31
2.2.5. <i>Mutacija.....</i>	33
2.2.6. <i>Odabir valnih duljina GA-PLS algoritmom.....</i>	34
2.2.6.1. <i>Inicijalizacija populacije .....</i>	35
2.2.6.2. <i>Razmnožavanje i mutacija .....</i>	35
2.2.6.3. <i>Evaluacija.....</i>	36
2.2.6.4. <i>Ubacivanje novog kromosoma u populaciju .....</i>	36
2.2.6.5. <i>Kontrola ponavljanja rješenja.....</i>	38
2.2.6.6. <i>Utjecaj različitih parametara na GA-PLS algoritam .....</i>	38
2.2.6.7. <i>Provjera podskupa varijabli u kromosomima .....</i>	39
2.2.6.8. <i>Ponavljanje GA-PLS algoritma.....</i>	39
2.2.6.9. <i>Uključivanje autokorelacije između spektralnih varijabli.....</i>	40
2.2.6.10. <i>Zaustavljanje algoritma.....</i>	42
<b>2.3. Umjetne neuronske mreže.....</b>	<b>42</b>
2.3.1. <i>Struktura neuronskih mreža .....</i>	43
2.3.2. <i>Algoritam povratnog rasprostiranja pogreške.....</i>	47
2.3.3. <i>Proces učenja.....</i>	49
2.3.4. <i>Levenberg-Marquardtov algoritam .....</i>	51
2.3.5. <i>Algoritam najstrmijeg spusta .....</i>	52
2.3.6. <i>Newtonova metoda.....</i>	52
2.3.7. <i>Gauss-Newtonov i Levenberg-Marquardtov algoritam .....</i>	55
<b>§ 3. EKSPERIMENTALNI DIO .....</b>	<b>58</b>
<b>3.1. Materijali.....</b>	<b>58</b>
<b>3.2. Metode analize .....</b>	<b>59</b>
3.2.1. <i>Spektrofotometrija.....</i>	59
3.2.2. <i>NIR spektroskopija.....</i>	59
3.2.3. <i>Ramanova spektroskopija .....</i>	60

---

3.2.4.	<i>Tekućinska kromatografija ultravisoke djelotvornosti</i> .....	60
3.2.5.	<i>Obrada podataka</i> .....	61
<b>§ 4.</b>	<b>REZULTATI I RASPRAVA</b> .....	<b>62</b>
<b>4.1.</b>	<b>Spektrofotometrijska analiza</b> .....	<b>62</b>
4.1.1.	<i>Obrada podataka PLS metodom</i> .....	63
4.1.2.	<i>Obrada podataka GA-PLS metodom</i> .....	65
4.1.3.	<i>Obrada podataka ANN metodom</i> .....	67
4.1.4.	<i>Primjena razvijenih PLS, GA-PLS i ANN metoda</i> .....	69
<b>4.2.</b>	<b>Analiza NIR spektroskopijom</b> .....	<b>71</b>
4.2.1.	<i>Obrada podataka PLS metodom</i> .....	73
4.2.2.	<i>Obrada podataka GA-PLS metodom</i> .....	79
4.2.3.	<i>Obrada podataka ANN metodom</i> .....	84
<b>4.3.</b>	<b>Analiza Ramanovom spektroskopijom</b> .....	<b>87</b>
4.3.1.	<i>Obrada podataka PLS metodom</i> .....	89
4.3.2.	<i>Obrada podataka GA-PLS metodom</i> .....	93
4.3.3.	<i>Obrada podataka ANN metodom</i> .....	97
<b>§ 5.</b>	<b>ZAKLJUČAK</b> .....	<b>101</b>
<b>§ 6.</b>	<b>LITERATURNI IZVORI</b> .....	<b>103</b>
<b>§ 7.</b>	<b>ŽIVOTOPIS</b> .....	<b>XXI</b>









Sveučilište u Zagrebu  
Prirodoslovno-matematički fakultet  
**Kemijski odsjek**

Doktorska disertacija

## SAŽETAK

### KVANTITATIVNA ANALIZA TRAMADOLA I PARACETAMOLA U TABLETAMA SPEKTROSKOPSKIM I KEMOMETRIČKIM METODAMA

Siniša Glavanović

Belupo d.d., Istraživanje i razvoj, Koprivnica  
Zavod za fizikalnu kemiju, Kemijski odsjek, Prirodoslovno-matematički fakultet,  
Sveučilište u Zagrebu

Razvijeni su novi analitički postupci temeljeni na spektroskopskim metodama i kemometričkim izračunima za istovremeno kvantitativno određivanje tramadola i paracetamola u tabletama. Spektroskopski podaci dobiveni analizom tableta primjenom UV-Vis spektrofotometrije, spektroskopije u bliskom infracrvenom području (NIR) te Ramanove spektroskopije obrađeni su upotrebom metode djelomičnih najmanjih kvadrata (PLS), genetičkih algoritama spregnutih s PLS-om (GA-PLS) i umjetnih neuronskih mreža (ANN). Rezultati dobiveni kemometričkom obradom spektroskopskih podataka statistički su uspoređeni s onima određenima validiranom metodom tekućinske kromatografije ultravisoke djelotvornosti (UHPLC).

Dobiveno je dobro slaganje između količina paracetamola određenih u tabletama primjenom UV spektrofotometrije te PLS, GA-PLS i ANN algoritama i referente UHPLC metode. U slučaju tramadola GA-PLS rezultati pokazali su se pouzdaniji u odnosu na one dobivene PLS-om i ANN-om. Najjednostavniji i najpouzdaniji modeli konstruirani su korištenjem PLS metode za paracetamol i GA-PLS metode za tramadol.

Maseni udjeli paracetamola u tabletama određeni primjenom NIR spektroskopije i kalibracijskih kemometričkih modela izrađenih pomoću PLS, GA-PLS i ANN metoda vrlo se dobro slažu s onima dobivenima UHPLC-om. Rezultati dobiveni primjenom ANN-a i GA-PLS-a međusobno se ne razlikuju znatno, te su pouzdaniji od onih izračunanih upotrebom PLS metode. S druge strane, udjeli tramadola određeni NIR metodom spregnutom s GA-PLS algoritmom znatno su pouzdaniji od onih izračunanih upotrebom druge dvije metode. GA-PLS pristupom obradi podataka pojednostavio se je kalibracijski model te poboljšala točnost i preciznost dobivenih rezultata.

Ramanova spektroskopija spregnuta s PLS, GA-PLS i ANN kemometričkim metodama uspješno je upotrijebljena za kvantitativnu analizu paracetamola, pri čemu su rezultati dobiveni tim metodama usporedivi s referentnim vrijednostima. Algoritam GA-PLS još se jednom pokazao superiornijim u odnosu na PLS i ANN. U slučaju analize tramadola, nijednom uporabljenom metodom nisu dobiveni rezultati usporedivi s referentnim vrijednostima.

(112 stranica, 46 slika, 12 tablica, 107 literaturnih navoda, jezik izvornika: hrvatski)

Rad je pohranjen u Središnjoj kemijskoj knjižnici, Horvatovac 102A, Zagreb i Nacionalnoj i sveučilišnoj knjižnici, Hrvatske bratske zajednice 4, Zagreb.

Ključne riječi: kemometričke metode/ NIR spektroskopija/ paracetamol/ Ramanova spektroskopija/ tramadol/ UV spektrofotometrija

Mentori: prof. dr. sc. Vladislav Tomišić  
izv. prof. dr. sc. Marina Vranić

Ocjenitelji: prof. dr. sc. Tomica Hrenar  
prof. dr. sc. Vladislav Tomišić  
izv. prof. dr. sc. Marina Vranić  
izv. prof. dr. sc. Snežana Miljanić  
dr. sc. Mladen Litvić, v. zn. sur

Rad prihvaćen: 1. veljače 2017.





University of Zagreb  
Faculty of Science  
Department of Chemistry

Doctoral Thesis

## ABSTRACT

### QUANTITATIVE ANALYSIS OF TRAMADOL AND PARACETAMOL IN TABLETS BY SPECTROSCOPIC AND CHEMOMETRIC METHODS

Siniša Glavanović

Belupo Inc., Research and Development, Koprivnica  
Division of Physical Chemistry, Department of Chemistry,  
Faculty of Science, University of Zagreb

Several new analytical procedures based on spectroscopic methods and chemometric calculations were developed for quantitative determination of tramadol and paracetamol in tablets. Spectroscopic data obtained by analyses of tablets using UV-Vis spectroscopy, spectroscopy in near infrared region (NIR), and Raman spectroscopy were processed by means of partial least squares (PLS), genetic algorithm coupled with PLS (GA-PLS) and artificial neural networks (ANN). Results gained by chemometric processing of the spectroscopic data were statistically compared with those obtained by means of validated ultra-high performance liquid chromatographic (UHPLC) method.

A good agreement between the amounts of paracetamol determined in the tablets using UV spectrophotometry and PLS, GA-PLS, as well as ANN algorithms, and that obtained by referent UHPLC method was found. In the case of tramadol, GA-PLS results were proven to be more reliable compared to those of PLS and ANN. The simplest and the most reliable models were constructed by using PLS method for paracetamol and GA-PLS method for tramadol.

Mass fractions of paracetamol in the tablets determined by means of NIR spectroscopy and calibration chemometric models constructed using PLS, GA-PLS, and ANN methods were in accordance with those gained by UHPLC. There was no significant difference between the results obtained by using ANN and GA-PLS, and they were more reliable than that obtained by PLS method. On the other hand, fractions of tramadol determined by NIR method coupled with GA-PLS algorithm were considerably more reliable than those calculated using other two methods. GA-PLS-based approach reduced the calibration model complexity and improved the accuracy and precision of the results.

Raman spectroscopy coupled with PLS, GA-PLS, and ANN chemometric methods was successfully applied for quantitative analysis of paracetamol, whereby the results obtained by these methods were comparable with the referent values. GA-PLS algorithm was once again proven to be superior over the PLS and ANN. In the case of tramadol analysis, the results gained by applying any of the used method were not comparable with the referent values.

(112 pages, 46 figures, 12 tables, 107 references, original in Croatian)

Thesis deposited in Central Chemical Library, Horvatovac 102A, Zagreb, Croatia and National and University Library, Hrvatske bratske zajednice 4, Zagreb, Croatia.

Keywords: chemometric methods/ NIR spectroscopy/ paracetamol/ quantitative analysis/ Raman spectroscopy/ tramadol/ UV spectrophotometry

Supervisors: Professor Vladislav Tomišić  
Associate Professor Marina Vranić

Reviewers: Professor Tomica Hrenar  
Professor Vladislav Tomišić  
Associate Professor Marina Vranić  
Associate Professor Snežana Miljanić  
Dr. Mladen Litvić, Senior Research Associate

Thesis accepted: 1 February 2017



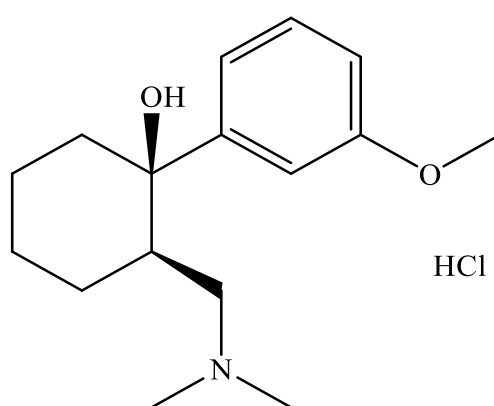
## § 1. UVOD

Svaka serija proizvedenog lijeka namijenjena tržištu treba odgovarati unaprijed definiranim zahtjevima kakvoće kako bi se osigurala njena djelotvornost, sigurnost i kvaliteta. Zahtjevi kakvoće lijeka definirani su u specifikacijama u kojima su opisani testovi i odgovarajući kriteriji prihvatljivosti kojima analizirani lijek mora odgovarati [1]. Jedan od parametara zahtjeva kakvoće koji se gotovo uvijek provjerava je količina (sadržaj) djelatne tvari u proizvodu. Najčešći analitički postupci koji se koriste u tu svrhu temelje se na tekućinskoj kromatografiji visoke ili ultravisoke djelotvornosti (engl. *High Performance Liquid Chromatography* - HPLC ili engl. *Ultra High Performance Liquid Chromatography* - UHPLC), UV spektrofotometriji, titrimetriji, plinskoj kromatografiji, a u posljednje vrijeme sve više i na spektroskopskim tehnikama kao što su infracrvena spektroskopija (engl. *Infrared* - IR), spektroskopija u bliskom infracrvenom području (engl. *Near Infrared* - NIR), Ramanova ili terahertz spektroskopija [2–4]. Naime, u farmaceutskoj industriji sve više se potiče i primjenjuje razvoj lijekova po principu kvalitete ugrađene u dizajn (engl. *Quality by Design* - QbD). QbD je sustavni pristup razvoju lijeka koji započinje definiranjem ciljeva te naglašava razumijevanje proizvoda i proizvodnog procesa, te njihovu kontrolu na temelju znanstvenog pristupa i upravljanja rizicima [5]. Kod takvog pristupa razvoju lijeka, mjerenje kritičnih parametara kvalitete i procesa provodi se u realnom vremenu primjenom prethodno navedenih spektroskopskih tehnika spregnutih s kemometričkim metodama [6].

Primjenom kemometričkih metoda za obradu spektroskopskih podataka omogućena je analiza složenih farmaceutskih pripravaka bez prethodne separacije pojedinih komponenata smjese, čak i u slučaju njihovog znatnog spektralnog preklapanja. Razvojem računala i softvera njihova je primjena postala raširena i općeprihvaćena. Glavne prednosti korištenja spektroskopskih tehnika u kombinaciji s kemometričkim metodama, u odnosu na druge tehnike koje se koriste u farmaceutskoj industriji, npr. tekućinsku kromatografiju visoke djelotvornosti, su brzina, niski troškovi analize te znatno manja štetnost za okoliš zbog manje potrošnje štetnih otapala potrebnih za analizu .

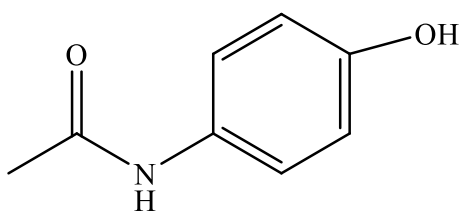
U literaturi je do danas opisana primjena samo jedna spektroskopska metoda za istovremenu kvantitativnu analizu tramadola i paracetamola u tabletama [7]. Budući da se maseni udjeli tih tvari u tabletama znatno razlikuju (udio paracetamola je 76,5 %, a tramadola 8,8 %), primjena spektroskopskih metoda za određivanje njihove količine vrlo je zahtjevna.

Tramadol, (1RS,2RS)-2-[(dimetilamino)metil]-1-(3-metoksifenil)cikloheksanol hidroklorid (slika 1), je centralno djelujući atipični opioid koji se sastoji od dva enantiomera, od kojih oba doprinose njegovoj aktivnosti, ali putem različitih mehanizama. Prvi puta je sintetiziran 1962. g., a od 1977. primjenjuje se za smanjenje boli. Tramadol ima slabi afinitet za  $\mu$  opioidne receptore, dok ga za receptore  $\delta$  ili  $\kappa$  uopće nema. Inhibira ponovnu pohranu noradrenalina (norepinefrina) i serotonina vezanjem na njihova neuronska mjesta pohrane što istovremeno smanjuje aferentnu signalizaciju boli i povećanje inhibitornog eferentnog signala. U usporedbi s drugim  $\mu$ -receptorskim agonistima nema klinički relevantne hemodinamske učinke i samo u manjoj mjeri produžuje gastrointestinalni tranzit [8,9].



Slika 1. Strukturna formula tramadola.

Paracetamol, (*N*-(4-hidroksifenil)acetamid (slika 2), jedan je od najpopularnijih i najviše korištenih analgetika i antipiretika bez protuupalnog djelovanja. Otkriven je 1893. g., a u široj medicinskoj praksi koristi se već više od pola stoljeća. Mehanizam djelovanja mu je kompleksan i do danas nije u potpunosti razjašnjen. Njegov analgetski učinak je uglavnom središnji, posredstvom nekoliko mehanizama, koji uključujuće inhibiciju sinteze prostaglandina, supraspinalnu aktivaciju silaznih serotoninskih putova, inhibiciju puta dušikovog oksida, te interakciju aktivnih metabolita s endogenim kanabinoidnim sustavom. Zbog nedostatka značajnog djelovanja na periferne prostaglandine, paracetamol se bolje podnosi od nesteroidnih protuupalnih lijekova (engl. *Non-Steroidal Anti-Inflammatory Drugs* - NSAID) i ne izaziva gastrointestinalne nuspojave [10,11].



Slika 2. Strukturna formula paracetamola.

Oralno primijenjena kombinacija fiksne doze tramadola i paracetamola u obliku tableta je indicirana za simptomatsko liječenje umjerene do jake boli. Ta kombinacija lijekova učinkovita je u smanjenju boli odraslih bolesnika u postoperativnom periodu nakon manjih operacija, smanjenu mišićno-koštane boli, bolne dijabetičke periferne neuropatije, odnosno bolnih migrena. Kako tramadol i paracetamol imaju komplementarne načine djelovanja te u organizmu djeluju na više različitih mjesta, njihova kombinacija pruža bolju analgeziju protiv različitih vrsta i izvora boli [12].

U sklopu ovog istraživanja pripravljene su Paracetamol + Tramadol 325 mg + 37,5 mg tablete s različitim masenim udjelima djelatnih tvari, odnosno tramadola i paracetamola, te su analizirane primjenom UV, NIR i Ramanove spektroskopije. Spektrometrijski podaci obrađeni su primjenom metode djelomičnih najmanjih kvadrata (engl. *Partial Least Squares* - PLS), genetičkih algoritama (engl. *Genetic Algorithms* - GA) spregnutih s PLS metodom (GA-PLS), te umjetnih neuronskih mreža (engl. *Artificial Neural Networks* - ANN) s ciljem kvantitativnog određivanja masenih udjela djelatnih tvari. Tablete su istovremeno analizirane i primjenom validirane metode tekućinske kromatografije ultravisoke djelotvornosti, a nađene količine djelatnih tvari statistički su uspoređene s onima izračunanim primjenom kemometričkih algoritama za obradu spektroskopskih podataka.



## § 2. LITERATURNI PREGLED

Upotreba kemometričkih kalibracijskih tehnika za obradu spektroskopskih podataka ima vrlo široku primjenu u kontroli kvalitete lijekova naročito u analizi smjesa i višekomponentnih farmaceutskih formulacija kod kojih postoji značajno spektralno preklapanje. Prednost upotrebe tih tehnika leži u činjenici da prilikom analize nije potrebno odvojiti pojedine komponente analiziranih smjesa.

U literaturi je opisan znatan broj primjera primjene metode djelomičnih najmanjih kvadrata, genetičkih algoritama spregnutih s metodom djelomičnih najmanjih kvadrata te metode umjetnih neuronskih mreža za kvantitativnu obradu spektrometrijskih podataka u analizi lijekova. U nastavku je dan pregled odabranih radova.

Ziémons i sur. opisali su upotrebu NIR spektroskopije i djelomičnih najmanjih kvadrata za kvantitativnu analizu paracetamola u sirupu [13]. NIR spektroskopija spregnuta s PLS-om uspješno je primijenjena za razvoj robustnog kalibracijskog modela za predviđanje količine paracetamola u formulaciji s relativno malim udjelom djelatne tvari (koncentracija lijeka u analiziranoj formulaciji bila je  $20 \text{ mg mL}^{-1}$ ). Tijekom pripreme kalibracijskih uzoraka uključene su varijabilnosti prisutne kod rutinske proizvodnje lijeka kao što su proizvodnja više serija istog lijeka u nizu (kampanjska proizvodnja), varijabilnosti u koncentraciji djelatne tvari te fizikalnih i kemijskih svojstava različitih serija pomoćnih tvari. Nadalje, NIR analizu provela su dva analitičara, a temperature uzoraka bile su  $25$  i  $30$  °C. Kao referentna metoda koristila se tekućinska kromatografija visoke djelotvornosti. Prije izrade kemometričkih modela primjenom PLS-a, izračunane su prva derivacije NIR spektara, koji su potom normalizirani primjenom *standard normal variate* (SNV) metode. Za odabir optimalnog broja latentnih varijabli u kalibracijskom modelu primijenjena je metoda *cross-validacije* slučajnog odabira podskupa kalibracijskih uzoraka. Za izradu modela upotrijebljene su 4 latentne varijable. Koeficijent determinacije izračunan za rezultate dobivene NIR i referentnom metodom u validacijskom nizu uzoraka bio je 0,993, a korijen srednjeg kvadrata pogreške predviđanja  $0,26 \text{ mg mL}^{-1}$ .

U radu Bittnera i sur. PLS kemometrička metoda korištena je za izradu kalibracijskog modela za određivanje raspodjele veličine čestica sirovine amoksicilin trihidrata [14]. Niz sirovina amoksicilin trihidrata analiziran je svjetlosnim mikroskopom pomoću kojega je izračunana brojčana raspodjela veličina čestica te NIR spektroskopijom u difuzno-

refleksijskom načinu. Podaci dobiveni svjetlosnim mikroskopom za  $D_{90}$  čestice korišteni su kao referentne vrijednosti. Primjenom PLS metode za obradu snimljenih spektara uspješno su kreirani modeli za predikciju raspodjele veličine čestica analizirane sirovine. Nadalje, potvrđeno je da su u slučaju čestica manjih od 85  $\mu\text{m}$  izmjerene apsorbancije obrnuto proporcionalne veličini čestica sirovine, te da postoji značajna linearna povezanost između tih vrijednosti.

Metoda djelomičnih najmanjih kvadrata upotrijebljena je za izradu kemometričkog modela za analizu vlage u tabletama s različitim djelatnim tvarima NIR spektroskopijom [15]. Analizirane tablete sadržavale su različite djelatne tvari te su bile različite formulacija i oblika (okrugle, izdužene i ovalne), a maseni udio vode bio je u području od 2 do 13 %. Za svaku tabletu mjerenja su provedena s obje strane u difuzno-refleksijskom načinu. Karl-Fischerova titracija korištena je kao primarna metoda za određivanje udjela vode u analiziranim uzorcima. Prva ili druga derivacija NIR spektara upotrijebljene su za izradu kalibracijskih modela primjenom PLS metode, a broj latentnih varijabli odabran je *cross-validacijom* uz izostavljanje jednog uzorka. Razvijena metoda uspješno je primijenjena za analizu validacijskog niza tableta te je potvrđeno vrlo dobro slaganje rezultata predložene NIR i referentne metode. Međutim, istaknuto je da ukoliko se novorazvijena NIR metoda želi uspješno primijeniti i za druge formulacije ili oblike, potrebno je u kalibracijski model uključiti uzorke tih proizvoda.

Sulub i sur. upotrijebili su PLS metodu za analizu ujednačenosti sadržaja djelatne tvari u tabletama NIR spektroskopijom u sklopu procesne kontrole lijeka [16]. Tablete u kojima je maseni udio djelatne tvari 29,4 % analizirane su pomoću 5 različitih NIR spektrometara smještenih na 3 različite lokacije čime je simuliran uobičajeni prijenos NIR metode iz istraživanja i razvoja u kontrolu kvalitete ili proizvodnju. Za svaki NIR instrument izrađen je kalibracijski model pomoću PLS-a. Razvijeni modeli uspoređeni su pomoću vrijednosti korijena srednjeg kvadrata pogreške previđanja (engl. *Root Mean Square Error of Prediction* - RMSEP) izračunanih na temelju količina djelatnih tvari određenih u validacijskom skupu uzoraka NIR i primarnom HPLC metodom. Pomoću svih instrumenata dobiveni su slični rezultati i vrlo dobro slaganje s referentnim vrijednostima. Dodatno, ispitana je robustnost razvijenih modela na način da je jedan od modela proglašen primarnim, a predikcija količina djelatne tvari u tabletama određena iz spektara snimljenih na preostala 4 uređaja. U ovom su se slučaju izračunane RMSEP vrijednosti za pojedine uređaje značajnije razlikovale. Isti postupak je ponovljen, ali prije upotrebe za kvantitativnu analizu NIR spektri dijela kalibracijskih uzoraka su međusobno standardizirani primjenom *piecewise direct*

*standardization* (PDS) i *wavelet hybrid direct standardization* (WHDS) metoda. Primjenom tih metoda omogućen je uspješan prijenos kalibracijskog modela između različitih NIR instrumenata čime je znatno ubrzan i olakšan prijenos razvijene metode. Nadalje, u radu je naglašena brzina predloženog NIR postupka u odnosu na kromatografsku metodu. Autori navode da je za analizu 30 tableta HPLC-om potrebno oko 5 h dok se rezultati primjenom NIR-a dobiju za svega 12 minuta.

Također, Koller i sur. upotrijebili su metodu djelomičnih najmanjih kvadrata za obradu NIR spektara s ciljem određivanja koncentracije djelatne tvari u realnom vremenu tijekom procesa miješanja sirovina te određivanja završne točke homogenizacije [17]. Mjerenja su provedena u refleksijskom načinu pomoću NIR spektrometra opremljenog s probom s optičkim vlaknima koja je bila u direktnom kontaktu sa smjesom acetilsalicilne kiseline i laktoze monohidrata čiji je proces homogenizacije proučavan. Rezultati dobiveni NIR metodom potvrđeni su primarnom spektrofotometrijskom metodom. U radu je naglašeno da je za optimalni smještaj probe u mješaču potrebno jako dobro poznavati tečenje proučavanog praška kako bi se homogenizacija uspješno pratila pomoću samo jedne probe. Naime, jedna proba može detektirati uspostavu uravnoteženog stanja, odnosno završne točke miješanja, samo na odabranoj lokaciji, međutim ukoliko u mješaču dolazi do raslojavanja, različita mjesta će u različitim vremenima postići ravnotežna stanja. Zbog toga, ukoliko nije dobro poznato tečenje proučavane smjese, za mjerenja se može upotrijebiti više proba istovremeno. Autori naglašavaju da upotreba NIR spektroskopije za praćenje miješanja ima niz prednosti u odnosu na najčešće korištene analitičke postupke, kao što su oni koji uključuju HPLC ili UV-Vis spektrometriju. Naime, navedena spektroskopska metoda je brza, nedestruktivna i omogućava mjerenje u realnom vremenu. Kod dokazivanja ujednačenosti sadržaja u nekoj smjesi primjenom neke od klasičnih metoda, glavni izvor pogreške je uzorkovanje prilikom kojeg može doći do raslojavanja već homogene smjese ili kontaminacije analizirane smjese [18,19].

Tawakkola i sur. upotrijebili su metode regresije glavnih komponenata (engl. *Principle Component Regression* - PCR) i djelomičnih najmanjih kvadrata za spektrofotometrijsku kvantitativnu analizu meoksiprila i hidroklorotiazida u tabletama [20]. Dobiveni rezultati uspoređeni su s onima određenima primjenom HPLC-a i univariatne metode kalibracije. Univariatna kalibracija provedena je primjenom *extended ratio subtraction* metode (EXRSM) spregnute s *ratio subtraction* metodom (RSM). Za izradu modela u slučaju obje multivariatne metode korištena su 4 faktora koji su određeni *cross-validacijom* izostavljanja jednog uzorka. Dodatno, razvijene metode su provjerene analizom otopina koje su sadržavale poznate koncentracije analita. Na temelju izračunanih vrijednosti *recovery*-a i usporedbe rezultata s

onima određenima HPLC-om autori zaključuju da su multivariatne metode superiornije u odnosu na upotrijebljenu univariatnu metodu.

Također, u literaturi se mogu naći primjeri upotrebe PLS metode za obradu Ramanovih i NIR spektara u svrhu kvantitativne analize polimorfa u farmaceutskim oblicima [21,22]. Promjena polimorfne forme djelatne tvari u čvrstim farmaceutskim oblicima može značajno utjecati na topljivosti, brzinu oslobađanja i bioraspoloživost lijeka. Stoga je iznimno važno kvantificirati i utvrditi takve promjene u toj vrsti farmaceutskih oblika. Kao primarna metoda koja se najčešće koristi za identifikaciju i kvantifikaciju polimorfa je rentgenska difrakcija na praškastom uzorku. Hennigan i sur. primijenili su NIR i Ramanovu spektroskopiju spregnute s PLS-om kako bi kvantificirali smjesu dva polimorfa, FII i FIII, piracetama u tabletama. Kalibracijski uzorci pripremljeni su miješanjem dvaju polimorfa s ostalim komponentama tableta u različitim masenim omjerima i to u području od 0 do 10 %. NIR spektri proizvedenih tableta snimljeni su u refleksijskom načinu, dok su Ramanovi spektri istih uzoraka snimljeni mjerenjem raspršenog zračenja koje je prošlo kroz uzorak i raspršenog s njegove površine. Spektri su predobrađeni primjenom *multiplicative scatter* korekcije (engl. *Multiplicative Scatter Correction* - MSC), a potom je izračunana njihova druga derivacija pomoću Savitzky–Golayeve metode kako bi se dodatno pojačale spektralne razlike dvaju polimorfa. Tako obrađeni spektri upotrijebljeni su za izradu kvantitativnog modela upotrebom PLS algoritma. Jedna latentna varijabla upotrijebljena je za izradu modela na temelju NIR spektara, dok su za modele izrađene iz Ramanovih spektara koji su snimljeni mjerenjem zračenja raspršenog s površine i raspršenog zračenja koje je prošlo kroz uzorak upotrijebljene 3, odnosno 1 varijabla. Najlošiji rezultati dobiveni su pomoću modela koji je konstruiran iz Ramanovih spektara snimljenih mjerenjem zračenja raspršenog s površine. Autori to objašnjavaju činjenicom da spektri snimljeni u refleksijskom načinu sadrže manje informacija u odnosu na one snimljene u transmisijskom načinu zbog manjeg volumena uzorka koji je izložen laseru kod ovog načina mjerenja. Utjecaj načina snimanja na kvalitetu spektara dobro je poznat. S druge strane, Hennigan i sur. najbolje su rezultate dobili pomoću modela konstruiranog iz Ramanovih spektara snimljenih mjerenjem raspršenog zračenja koje je prošlo kroz uzorak. Također, NIR spektroskopijom su dobivene vrlo slične vrijednosti onima najboljima, a metoda ima samo neznatno veći limit detekcije od Ramanove. Kao moguće objašnjenje autori navode da su u Ramanovom spektru zbog vrlo oštih i dobro odvojenih vrpca razlike između polimorfa uočljivije u odnosu na one u NIR spektru gdje su vrpce vrlo široke i međusobno se preklapaju.

Objavljeno je dosta radova u kojima je opisana upotreba PLS-a i NIR spektroskopije za predviđanje oslobađanja djelatne tvari iz farmaceutskih oblika. Freitas i sur. koristili su

navedene metode za predviđanje profila oslobađanja klonazepama iz tableta [23]. Ispitivanjem oslobađanja djelatne tvari iz čvrstih farmaceutskih oblika moguće je relativno jednostavno i brzo dobiti informacije o mehanizmima koji se događaju u lijeku tijekom njegovog izlaganja mediju, a dobiveni podaci mogu se povezati s bioraspoloživosti [24]. Zbog toga se testovi oslobađanja intenzivno koriste u fazi razvoja lijeka kako bi se razvila i optimizirala formulacija koja osigurava kvalitetan i djelotvoran proizvod. Autori su niz tableta različitih formulacija i tvrdoća istovremeno analizirali NIR spektroskopijom u refleksijskom načinu te primjenom aparata za ispitivanje oslobađanja djelatne tvari [25], a za kvantifikaciju upotrijebljena je HPLC metoda. NIR spektri su bez predobrade korelirani s količinom oslobođene djelatne tvari u određenim vremenskim intervalima primjenom PLS metode. Za izradu modela upotrijebljeno je 10 faktora, a njihov broj određen je primjenom metode *cross-validacije* izostavljanjem jednog uzorka. Razvijena NIR metoda provjerena je analizom validacijskog niza tableta. Vrijednosti koeficijenta determinacije između rezultata primarne i predložene metode dobiveni su u rasponu od 0,80 do 0,92. Autori su ovim radom pokazali da je PLS pogodna kemometrička metoda za izradu kalibracijskih modela, a NIR spektroskopija pogodna tehnika za predviđanje oslobađanja djelatne tvari. Razvijena NIR metoda u usporedbi s primarnom tehnikom brza je i nedestruktivna te se može upotrijebiti u razvoju novog lijeka.

U radu Darwisha i sur. uspoređeni su kvantitativni kemometrički modeli razvijeni upotrebom genetičkih algoritama spregnutih s PLS i ANN metodama (GA-ANN), umjetnih neuronskih mreža, djelomičnih najmanjih kvadrata te analize glavnih komponenata spregnute s ANN metodom (PCA-ANN) [26]. Razvijeni modeli korišteni su za analizu amlodipina, valsartana i hidroklorotiazida u tabletama UV spektrofotometrijskom metodom. Za kalibraciju pripremljene su otopine u kojima je koncentracija amlodipina bila u području od 2 do 10  $\mu\text{g mL}^{-1}$ , valsartana od 20 do 40  $\mu\text{g mL}^{-1}$ , a hidroklorotiazida od 3,125 do 15,625  $\mu\text{g mL}^{-1}$ . UV spektri snimljeni su u području od 200 do 400 nm i obrađeni primjenom PLS1, GA-PLS, ANN, GA-ANN i PCA-ANN algoritama. Broj latentnih varijabli u modelima konstruiranima primjenom PLS1 i GA-PLS metoda određeni su na temelju RMSECV vrijednosti izračunanih *cross-validacijom* izostavljanjem jednog uzorka. U slučaju PLS-a za sve tri djelatne tvari korištena su 3 faktora, dok su za model izračunan GA-PLS-om za amlodipin upotrijebljene 2 latentne varijable, a za valsartan i hidroklorotiazid 3. Neuronske mreže sastojale su se od 3 sloja, a broj neurona u skrivenom sloju je variran. Tako je u slučaju ANN-a za amlodipin i valsartan upotrijebljeno 10 čvorova, a za hidroklorotiazid 1. U slučaju GA-ANN algoritma upotrijebljeno je 7 neurona za amlodipin, 10 za valsartan i 1 za hidroklorotiazid. Za PCA-ANN

mrežu za sve djelatne tvari upotrijebljena su 3 čvora. U skrivenom i izlaznom sloju korištena je linearna prijenosna funkcija. Sve mreže trenirane su primjenom Levenberg–Marquardtovog algoritma. Razvijene metode provjerene su analizom neovisnog niza otopina amlodipina, valsartana i hidroklorotiazida, a komercijalne tablete (Exforge HCT<sup>®</sup> 5/160/12,5, 5/160/25) dodatno su analizirane primjenom HPLC metode. Na temelju podataka dobivenih analizom validacijskih otopina izračunane su vrijednosti *recovery*-a i RMSEP-a, a količine lijekova u tabletama analizirane UV-om i HPLC-om uspoređene su primjenom *F* i *t*-testa. Na temelju izračunanih statističkih vrijednosti pokazalo se da su sve predložene metode uspješno upotrijebljene za analizu tableta te da se rezultati statistički ne razlikuju od onih određenih kromatografskom analizom. Međutim, upotreba genetičkih algoritama za odabir valnih duljina smanjila je kompleksnost PLS1 modela za predviđanje koncentracije amlodipina, te su vrijednosti RMSEP-a i *recovery*-a za sva tri lijeka smanjene što ukazuje da GA-PLS metoda ima veću rezolucijsku moć od PLS1 postupka. Također, GA je omogućio upotrebu manjeg broja neurona u skrivenom sloju tijekom ANN analize. S druge strane, upotreba analize glavnih komponenata nije znatno utjecala na poboljšanje kvalitete ANN modela. Autori zaključuju da je razlog tome što GA uvodi najrelevantnije valne duljine za predviđanje koncentracije djelatnih tvari.

Također, metoda djelomičnih najmanjih kvadrata, umjetnih neuronskih mreža spregnuta s analizom glavnih komponenata te genetički algoritmi u kombinaciji s PLS-om upotrijebljeni su za obradu spektrofotometrijskih podataka tijekom kvantitativne analize kombinacije paracetamola, ibuprofena i kofeina u tabletama [27]. Pripremljen je niz otopina u kojima je koncentracija paracetamola bila od 0,60 do 11,0  $\mu\text{g mL}^{-1}$ , ibuprofena od 1,00 do 24,0  $\mu\text{g mL}^{-1}$ , te kofeina od 1,00 do 18,0  $\mu\text{g mL}^{-1}$ . UV spektri otopina snimljeni su u području od 200 do 400 nm koji su potom upotrijebljeni za izradu kalibracijskih modela primjenom navedenih kemometričkih metoda. Za izradu modela primjenom PLS1 algoritma upotrijebljeno je 5 faktora za paracetamol, 6 za ibuprofen i 4 za kofein. Broj faktora odabran je na temelju izračunanih vrijednosti sume kvadrata reziduala pogreške predviđanja, a statistička značajnost pojedine latentne varijable utvrđena je primjenom *F*-testa. Upotrebom genetičkog algoritma odabrana su spektralna područja lijekova koja sadrže najviše informacija o njihovoj koncentraciji, a potom su ta područja upotrijebljena za izradu modela PLS-om. I u ovom slučaju upotreba GA smanjila je kompleksnost modela za sve analite u odnosu na one konstruirane PLS-om kada su korišteni cijeli spektri. Tako su GA-PLS algoritmom za paracetamol u model uključene 3 latentne varijable, za ibuprofen 4, a za kofein njih 3. Broj

varijabli za ANN analizu smanjen je upotrebom PCA pa je tako za paracetamol korišteno 5 glavnih komponenata, za ibuprofen 6, a za kofein 4. Upotrebom PCA broj neovisnih varijabli je reduciran pa je smanjeno i vrijeme potrebno za trening mreže. Neuronska mreža sastojala se od ulaznog, skrivenog i izlaznog sloja. U skrivenom sloju za ibuprofen i paracetamol smještene su 4 neurona, a za kofein njih 3. I u ovom radu razvijeni modeli provjereni su analizom validacijskog niza otopina djelatnih tvari te komercijalnih uzoraka kapsula. GA-PLS algoritam se ponovno pokazao kao superiorna metoda u odnosu na druga dva upotrijebljena statistička postupka. Kao što je već istaknuto, GA je smanjio kompleksnost modela, a rezultati dobiveni ovom metodom točniji su i precizniji od onih izračunanih PCA-ANN i PLS metodama. Svi razvijeni modeli uspješno su upotrijebljeni za analizu smjese djelatnih tvari iako postoji značajno spektralno preklapanje spektara pojedinih analita.

Nadalje, Elkhoudary i sur. usporedili su kvantitativne modele kreirane PLS, GA-PLS, ANN, GA-ANN i PCA-ANN metodama za analizu smjese metronidazola, diloksanida, spiramicina i kliokinola u tabletama [28]. UV spektri otopine lijekova u kojima je koncentracija metronidazola, spiramicina i kliokinola bila u području  $1,5 - 4,5 \mu\text{g mL}^{-1}$ , a diloksanida u rasponu  $1,2 - 3,6 \mu\text{g mL}^{-1}$  snimljeni su u području od 220 do 400 nm te upotrijebljeni za izradu kalibracijskih modela. Linearni kalibracijski modeli konstruirani su primjenom PLS i GA-PLS algoritama, te PCA-ANN i GA-ANN metodama kod kojih je u skrivenom i u izlaznom sloju mreža upotrijebljena linearna prijenosna funkcija. Nelinearni modeli izrađeni su primjenom PCA-ANN i GA-ANN algoritama u kojima su kao prijenosne funkcije upotrijebljene tangens hiperbolni i linearna funkcija. Za trening svih mreža upotrijebljen je Levenberg–Marquardtov algoritam, a kako bi se dodatno skratilo vrijeme učenja broj varijabli je smanjen upotrebom GA i PCA. Primjenom genetičkog algoritma za metronidazol odabrane su 22 varijable, za spiramicin njih 30, za diloksanid 27 te 28 za kliokinol. Razvijeni modeli međusobno su uspoređeni na temelju rezultata analize otopina smjese djelatnih tvari i uzoraka različitih komercijalnih tableta. Prema određenim vrijednostima *recovery*-a i RMSEP-a zaključeno je da su najtočniji i najprecizniji rezultati izračunani primjenom nelinearnih modela konstruiranih GA-ANN-om i PCA-ANN-om. Također, linearne i nelinearne ANN metode daju pouzdanije rezultate od GA-PLS-a, dok su rezultati dobiveni PLS-om najlošiji. Elkhoudary i sur. zaključuju da izmjerene apsorbancije otopina za kalibraciju i validaciju najvećih koncentracija diloksanida i kliokinola izlaze izvan linearnog područja stoga nelinearni modeli kreirani primjenom ANN algoritmom daju najtočnije rezultate.

Primjer upotrebe umjetnih neuronskih mreža te PCA-ANN metode za nelinearnu kalibraciju u svrhu određivanja sadržaja djelatnih tvari, tj. antipirina i kofeina, u tabletama NIR spektroskopijom može se pronaći u radu Doua i sur. [29]. Za analizu upotrijebljene su 73 serije tableta, od kojih su 43 korištene za trening mreže, 20 za nadzor učenja, a 10 za validaciju razvijenih kemometričkih metoda. Maseni udio antipirina u formulaciji je bio 86,4 %, a kofeina 8,2 %. NIR spektri snimljeni su u području od 1100 do 2500 nm u refleksijskom načinu. Mreže su se sastojale od 3 sloja te su trenirane metodom povratnog rasprostiranja pogreške. Snimljeni spektri predobrađeni su primjenom SNV i MSC metoda, dok su prva i druga derivacija izračunane upotrebom Savitzky–Golayevog algoritma. Upotrebom PCA smanjen je broj ulaznih varijabli za ANN metodu pa su u ulazni sloj mreža uvedeni faktorski bodovi (engl. *principal component scores*). Konstruirani modeli uspoređeni su na temelju izračunanih vrijednosti relativne standardne pogreške i koeficijenta determinacije. ANN i PCA-ANN algoritmi uspješno su upotrijebljeni za analizu antipirina i kofeina u analiziranim uzrocima. Međutim, modeli razvijeni upotrebom PCA-ANN-a daju bolje rezultate u odnosu na one kad nije primijenjena PCA metoda. Također, autori zaključuju da je nužno da NIR spektri budu obrađeni prije upotrebe ANN-a kako bi se razvio kvalitetan nelinearni kalibraciji sustav. Tako je model s najmanjom vrijednosti relativne standardne pogreške i najvećim koeficijentom determinacije dobiven iz spektara koji su obrađeni SNV metodom.

Genetički algoritmi u sprezi s umjetnim neuronskim mrežama korišteni su za kvantitativno određivanje cefaleksina u prašku [30]. Pripremljeno je 120 uzoraka praška u kojem je koncentracija djelatne tvari mijenjana u području  $\pm 10\%$  od nominalne količine. NIR spektri snimljeni su u području od 800 do 1100 nm u difuzno-refleksijskom načinu. Spektralni podaci obrađeni su primjenom ANN algoritma i metode djelomičnih najmanjih kvadrata. Broj ulaznih varijabli za ANN smanjen je primjenom metoda *uninformative variable elimination* (UVE), *mutual information* (MI) i modificiranim genetičkim algoritmom. Za trening mreže i istovremenu provjeru učenja upotrijebljena su 72 odnosno 24 uzorka, dok je provjera razvijenih metoda provedena analizom 24 neovisne serije praška. Točnost razvijenih modela uspoređena je na temelju izračunanih vrijednosti korijena srednjeg kvadrata pogreške. U ovom radu pokazano je da svi nelinearni modeli konstruirani ANN-om daju bolje rezultate od PLS-a. Odabir valnih duljina MI, UVE i GA algoritmima znatno utječe na rezultate ANN modela te skraćuje proces učenja mreže. Modificirani GA kojim se odabire stalni broj ulaznih varijabli za ANN smanjuje strukturu mreže te razvijeni GA-ANN model u usporedbi s ostalim nelinearnim kalibracijskim modelima ima najbolje performanse.



U radu Srathaphuta i sur. opisan je primjer upotrebe genetičkih algoritama u kombinaciji s PLS metodom za kvantitativnu spektrofotometrijsku analizu vitamina B12 u tabletama [31]. Otopinama koje su sadržavale vitamin B12 u području koncentracija od 3,2 do 12,8 mg mL<sup>-1</sup> te ostale komponente tabletne formulacije snimljeni su spektri od 200 do 800 nm. Za obradu primjenom PLS i GA-PLS algoritama korišteno je spektralno područje od 354 do 600 nm u kojem je vizualnom analizom utvrđeno da postoji najmanje preklapanje analita i ostalih komponenta formulacije. Upotrebom genetičkog algoritma iz odabranog područja selektirane su 33 valne duljine koje su korištene za izradu modela PLS-om. Za odabir optimalnog broja latentnih varijabli u razvijenim modelima upotrijebljena je vrijednost sume kvadrata reziduala pogreške predviđanja (engl. *Predicted Residual Error Sum of Squares* - PRESS). Na temelju tih vrijednosti utvrđeno je da je za model konstruiran PLS-om potrebno 9 faktora, a za onaj izrađen GA-PLS-om njih 3. Osim značajnog pojednostavljenja modela primjenom GA-PLS-a poboljšana je točnost i preciznost rezultata. Tako su za PLS metodu izračunani *recovery* i RMSEP 119,3 % i 1,27 %, a za GA-PLS 101,4 % i 0,18 %, što je u slučaju vrijednosti korijena srednjeg kvadrata pogreške predikcije poboljšanje od 86 %.

Arcos i sur. koristili su genetičke algoritme spregnute s metodom djelomičnih najmanjih kvadrata za spektrofotometrijsku analizu smjese indometacina i acemetacina [32]. Te dvije molekule imaju vrlo sličnu kemijsku strukturu zbog čega su im UV spektri gotovo identični. Kako bi se dodatno istaknule spektralne razlike pripremljene su otopine lijekova različitih koncentracija u pH-području od 5 do 8. Otopinama su snimljeni spektri u području od 200 do 400 nm te su podaci obrađeni pomoću PLS i GA-PLS metoda. Na temelju dobivenih rezultata autori zaključuju da odabir valnih duljina primjenom GA pozitivno utječe na sposobnost predviđanja kreiranog modela jer su iz spektra izostavljena područja u kojima apsorbira upotrijebljeno otapalo, tj. acetonitril, dok su s druge strane odabrani dijelovi spektra u kojima se apsorbancije analita razlikuju od drugih komponenta. Nadalje, promjena u kemijskom sustavu upotrebom otapala različitih pH-vrijednosti nije povećala razlike među UV spektrima indometacina i acemetacina pa samim time nisu ni poboljšane performanse razvijenih modela. Primjena GA pomogla je razumijevanju korelacije između koncentracije analita i pojedinih dijelova snimljenih UV spektara.

Pregledom literature može se pronaći velik broj kvantitativnih analitičkih metoda za pojedinačnu analizu tramadola i paracetamola u farmaceutskim proizvodima i biološkim materijalima. Tako su opisane metode protočne analize injektiranjem uz infracrveni detektor s Fourierovom transformacijom [33], infracrvene spektroskopije s Fourierovom transformacijom [34], infracrvene spektroskopije u bliskom području [13,35–37], UV-Vis

[38,39], i florescencijske spektroskopije [40], protočne analize injektiranjem uz spektrofotometrijski detektor [41], HPLC [42], tekućinske kromatografije visoke djelotvornosti sa spektrometrom masa kao detektorom (engl. *High Performance Liquid Chromatography-Mass Spectrometry* - HPLC-MS) [43], tekućinske kromatografije visoke djelotvornosti s tandemskom spektrometrijom masa (engl. *High Performance Liquid Chromatography/Tandem Mass Spectrometry* - HPLC-MS/MS) [44], plinske kromatografije (engl. *Gas Chromatography* - GC) [45,46], plinske kromatografije s masenim spektrometrom kao detektorom (engl. *Gas Chromatography - Mass Spectrometry* - GC-MS) [47,48], kapilarne elektroforeze s elektrokemiluminiscentnom detekcijom [49,50], voltimetrije [51,52], protočne analize injektiranjem uz pravokutnovalnom voltametrijskom i amperometrijskom detekcijom [53], titrimetrije [54], i protočne analize injektiranjem s kemiluminescencijskom detekcijom [55]. Za određivanje paracetamola u višekomponentnim smjesama lijekova upotrijebljene su derivacijska spektrofotometrija [56,57] i spektrofotometrija spregnuta s multivarijantnim kalibracijskim tehnikama [27,58,59].

Međutim, unatoč velikom broju objavljenih radova u kojima su pojedinačno analizirani tramadol i paracetamol, u samo nekoliko njih opisana je istovremena analiza ove dvije djelatne tvari u farmaceutskim proizvodima i ljudskoj plazmi. Tekućinska kromatografija visoke djelotvornosti obrnutih faza [60], derivacijska spektrofotometrija drugog reda [7] i diferencijalna pulsna voltimetrija sa staklastom ugljikovom elektrodom kao senzorom [61] upotrijebljeni su za analizu tramadola i paracetamola u lijekovima, dok je metoda tekućinske kromatografije visoke djelotvornosti s masenim detektorom [62] korištena za analizu ta dva lijeka u ljudskoj plazmi. Radovi u kojima su prikazane metode za analizu proučavanih djelatnih tvari u tabletama detaljnije će biti opisane u nastavku.

U radu Kamblea i sur. opisana je stabilitetno-indikativna HPLC metoda za analizu tramadola i paracetamola u tabletama [60]. Kao nepokretna faza korištena je Intersil C18 kolona dimenzija 250 × 4,6 mm i veličine čestica 5 μm dok je kao pokretna faza upotrijebljena smjesa vodene otopine fosfatne kiseline pH 3,4 i metanola (60 : 40, volumni omjer). Protok pokretne faze postavljen je na 1 mL min<sup>-1</sup>, a detekcija oba analita je bila pri 218 nm. Djelatne tvari ekstrahirane su iz homogeniziranih uzoraka tableta u 100 mL metanola, nakon čega je otopina filtrirana, razrijeđena 4,0 mL na 100 mL pokretnom fazom te injektirana u HPLC sustav. Kako bi se dokazala selektivnost razvijene metode provedena je studija prisiljene razgradnje farmaceutskog proizvoda. Uzorci tableta izloženi su kiselim, lužnatim, oksidativnim, termičkim i fotolitičkim uvjetima nakon čega su analizirani predloženim analitičkim postupkom. Također, metoda je uspješno validirana prema zahtjevima opisanim u

odgovarajućoj ICH smjernici o validaciji analitičkih postupaka za primjenu u farmaceutskoj industriji [63]. Kamblea i sur. zaključili su da je razvijena metoda brza i prikladna za praćenje stabilnosti proizvoda.

Primjena derivacijske spektrofotometrijske metode drugog reda za istovremenu analizu tramadola i paracetamola u tabletama može se pronaći u radu Ines Toral i sur. [7]. Primjenom Savitzky–Golayevog algoritma iz snimljenih UV spektara otopina standarda dviju djelatnih tvari izračunane su druge derivacije. Spektri su snimljeni u području od 200 do 400 nm, a za kvantifikaciju tramadola odabrana je vrijednost druge derivacije spektra pri 285,7 nm, dok je za paracetamol odabrana ona pri 308 nm. Valna duljina za analizu tramadola odabrana je na temelju činjenice da u toj točki druga derivacija spektra paracetamola siječe x os pa određena vrijednost deriviranog spektra odgovara samo tramadolu. S druge strane, pri 308 nm tramadol ne apsorbira pa izmjerene vrijednosti pri toj valnoj duljini odgovaraju samo paracetamolu. Linearnom regresijom iz određenih vrijednosti pri navedenim valnim duljinama i koncentracije otopina standarda određeni su kalibracijski pravci za analizirane lijekove. Točnost razvijene metode provjerena je analizom otopine smjese djelatnih tvari pri čemu je za paracetamol dobiven *recovery* 97,4 %, a za tramadol 100,6 %. Također, predložena metoda primijenjena je i za analizu komercijalnih tableta, međutim budući da nije upotrijebljen referentni postupak dobiveni rezultati nisu uspoređeni s referentnim vrijednostima. Iako je u radu zaključeno da se razvijeni analitički postupak može upotrijebiti za rutinsku kontrolu lijeka, prema vrijednosti *recovery*-a za paracetamol, metoda ne zadovoljava zahtjev za točnost koji za većinu kvantitativnih analiza za određivanje količine djelatne tvari u lijekovima iznosi 98,0 – 102,0 % [64].

Ghorbani-Bidkorbeh i sur. upotrijebili su diferencijalnu pulsnu voltametriju sa staklastom ugljikovom elektrodom kojoj je površina modificirana nanošenjem suspenzije ugljikovih nanočestica za analizu tramadola i paracetamola u tabletama [61]. Nanošenjem suspenzije nanočestica na staklastu elektrodu povećana je njena elektroaktivna površina što je znatno povećalo vrijednosti strujnih vrhova. Tanki sloj modifikatora pokazao je katalitički učinak i povećao je brzinu procesa prijenosa elektrona. Sve navedeno povećalo je osjetljivost razvijenog senzora koji je pokazao dobru selektivnost i osjetljivost za određivanje paracetamola i tramadola u širokom području koncentracija. Razvijena metoda uspješno je primijenjena za analizu komercijalnih tableta koje su sadržavale 325 mg paracetamola i 37,5 mg tramadola te je validirana prema odgovarajućoj ICH smjernici [63].

Za kvantitativnu obradu dobivenih spektroskopskih podataka u ovom radu korištene su slijedeće kemometričke metode: regresija metodom djelomičnih najmanjih kvadrata, genetički

algoritmi povezani s PLS metodom i umjetne neuronske mreže. Svaka od tih metoda bit će opisana u nastavku.

## 2.1. Regresija metodom djelomičnih najmanjih kvadrata

Metodu djelomičnih najmanjih kvadrata razvio je 1966. g. Herman Wold s primjenom u ekonomiji. Međutim, vrlo brzo metoda se počela primjenjivati u kemometrici i kemijskom inženjerstvu [31,65].

PLS metoda kombinira značajke analize glavnih komponenata (engl. *Principal Component Analysis* – PCA) i multivarijatne linearne regresije (engl. *Multivariate Linear Regression* – MLR), a svrha joj je predikcija ili analiza skupa ovisnih varijabli ( $Y$ ) iz skupa neovisnih varijabli, prediktora ( $X$ ). Predikcija se postiže ekstrakcijom ortogonalnih faktora (latentnih varijabli) s najboljim potencijalom predikcije iz skupa prediktora [66,67].

Općenito, za veliki broj varijabli moguće je koristiti MLR, međutim ukoliko je broj prediktorskih varijabli velik u usporedbi s brojem ovisnih varijabli vjerojatno je da se MLR neće moći provesti zbog višestruke kolinearnosti neovisnih varijabli. U takvom slučaju matrica neovisnih varijabli  $X$  je singularna te ne postoji njena inverzna matrica,  $X^{-1}$ . Da bi se riješio takav problem moguće je smanjiti broj neovisnih varijabli, primijeniti regresiju glavnih komponenata (engl. *Principal Component Regression* - PCR) ili PLS regresiju.

Kod PCR metode matrica neovisnih varijabli  $X$  se metodom dekompozicije po singularnim vrijednostima (engl. *Singular Value Decomposition* - SVD) rastavlja kao:

$$X = U S V^T \quad (1)$$

s time da vrijedi uvjet ortogonalnosti (ortonormiranosti)

$$U^T U = V^T V = I \quad (2)$$

gdje je

$U$  – matrica lijevih singularnih vektora

$V$  – matrica desnih singularnih vektora

$S$  - dijagonalna matrica singularnih vrijednosti

$I$  – matrica identiteta

Singularni vektori su u matricama poredani prema odgovarajućim singularnim vrijednostima koje su jednake pozitivnom korijenu vlastitih vrijednosti disperzijske matrice  $X^T X$ . Lijevi singularni vektori (tj. stupci matrice  $U$ ) koriste se za predikciju ovisnih varijabli  $Y$  koristeći standardnu regresiju. Ortogonalnost (ortonormiranost) singularnih vektora rješava problem višestruke kolinearnosti neovisnih varijabli  $X$ . Međutim, kod PCR analize glavne komponente se biraju tako da najbolje opisuju neovisne varijable  $X$  te nema nikakve garancije da su te glavne komponente ujedno relevantne i za ovisne varijable  $Y$ .

Za razliku od PCR metode PLS metoda pronalazi komponente (latentne vektore) od  $X$  koji su relevantni i za  $Y$  varijable. Preciznije, PLS metoda rastavlja  $X$  i  $Y$  varijable na latentne vektore, ali tako da oni opisuju što je više moguće kovarijance između  $X$  i  $Y$  varijabli. Najčešće korišteni algoritam za izračun singularnih vrijednosti i singularnih vektora je PLS-NIPALS algoritam (engl. *Non-linear Iterative Partial Least Squares* - NIPALS) [66,68].

Matrica neovisnih varijabli rastavlja se kao  $X = T P^T$  pri čemu vrijedi  $T T^T = I$ . Analogno kao i kod PCA  $T$  je matrica faktorskih bodova (engl. *score*), a  $P$  je matrica opterećenja (engl. *loading*). Ovisne varijable  $Y$  se procjenjuju kao  $\hat{Y} = T B C^T$  gdje je  $B$  dijagonalna matrica regresijskih koeficijenata (engl. *regression weights*), a  $C$  matrica koeficijenta („*weight matrix*“) ovisnih varijabli. Svaki stupac matrice  $T$  naziva se latentnim vektorom.

Na početku PLS-NIPALS algoritma kreiraju se dvije matrice  $E = X$  i  $F = Y$ , čiji se stupci zatim centriraju i normiraju, a potom se slučajnim odabirom generira vektor  $u$ . Nakon toga slijedi iterativni dio algoritma koji se sastoji od četiri koraka:

$$1. \quad w \propto E^T u \quad (3)$$

$$2. \quad t = E w \quad (4)$$

$$3. \quad c \propto F^T t \quad (5)$$

$$4. \quad u = F c \quad (6)$$

simbol  $\propto$  u prethodnim izrazima označava da se vektori nakon izračuna normiraju, tj. da vrijedi  $w^T w = c^T c = I$ . Gore navedeni koraci ponavljaju se sve dok se ne postigne konvergencija vektora  $w$ . Nakon toga se izračunavaju vrijednost  $b$  kao  $b = t^T u$  i vektor  $p$  kao

$\mathbf{p} = \mathbf{E}^T \mathbf{t}$ . U sljedećem koraku od matrica  $\mathbf{E}$  i  $\mathbf{F}$  oduzima se doprinos prvog latentnog vektora  $\mathbf{t}$ ,  $\mathbf{E}_1 = \mathbf{E} - \mathbf{t} \mathbf{p}^T$  i  $\mathbf{F}_1 = \mathbf{F} - \mathbf{b} \mathbf{t} \mathbf{c}^T$ . Čitav postupak se ponavlja počevši od koraka 1 tako dugo dok matrica  $\mathbf{E}$  točnije, matrica  $\mathbf{X}$  ne postane nul-matrica. Vektori  $\mathbf{t}$ ,  $\mathbf{u}$ ,  $\mathbf{w}$ ,  $\mathbf{c}$  i  $\mathbf{p}$  spremaju se u odgovarajuće matrice, a vrijednosti  $b$  u dijagonalnu matricu  $\mathbf{B}$ .

Kad se postigne konvergencija vektora  $\mathbf{w}$  korištenjem izraza od (3) do (6) moguće je dobiti sljedeći niz jednažbi:

$$\mathbf{w} \propto \mathbf{E}^T \mathbf{u} \propto \mathbf{E}^T \mathbf{F} \mathbf{c} \propto \mathbf{E}^T \mathbf{F} \mathbf{F}^T \mathbf{t} \propto \mathbf{E}^T \mathbf{F} \mathbf{F}^T \mathbf{E} \mathbf{w} \quad (7)$$

Iz ovoga slijedi da je vektor  $\mathbf{w}$  vlastiti vektor  $\mathbf{X}^T \mathbf{Y} \mathbf{Y}^T \mathbf{X}$ , a vektor  $\mathbf{c}$  vlastiti vektor  $\mathbf{Y}^T \mathbf{X} \mathbf{X}^T \mathbf{Y}$ . Prethodni umnošci matrica su dva simetrična produkta matrica  $(\mathbf{X}^T \mathbf{Y}) (\mathbf{X}^T \mathbf{Y})^T$  i  $(\mathbf{X}^T \mathbf{Y})^T (\mathbf{X}^T \mathbf{Y})$  umnoška matrica  $(\mathbf{X}^T \mathbf{Y})$ . Zbog toga je druga interpretacija vektora  $\mathbf{w}$  i  $\mathbf{c}$  kao parova singularnih vektora matrice kovarijance  $\mathbf{X}^T \mathbf{Y}$ . Prvi par singularnih vektora  $\mathbf{w}$  i  $\mathbf{c}$  čine jedinstven par normiranih *vaganih* vektora koji maksimiziraju izraz  $\mathbf{w}^T (\mathbf{X}^T \mathbf{Y}) \mathbf{c} = (\mathbf{X} \mathbf{w})^T (\mathbf{Y} \mathbf{c}) = \mathbf{t}^T \mathbf{u}$ . Iz ovoga pak slijedi da se PLS faktori  $\mathbf{t}$  i  $\mathbf{u}$  biraju tako da maksimiziraju kovarijancu  $\mathbf{X}$  i  $\mathbf{Y}$  varijabli.

Kovarijanca faktora  $\mathbf{t}$  i  $\mathbf{u}$  može se prikazati kao umnožak triju komponenti:

$$\text{cov}(\mathbf{t}, \mathbf{u}) = s(\mathbf{t}) s(\mathbf{u}) r_{t,u} \quad (8)$$

čijim se kvadriranjem dobiva sljedeća jednažba:

$$\text{cov}^2(\mathbf{t}, \mathbf{u}) = \text{var}(\mathbf{t}) \text{var}(\mathbf{u}) r_{t,u}^2 \quad (9)$$

gdje je:

$s(\mathbf{t})$  – standardna devijacija faktora  $\mathbf{t}$

$s(\mathbf{u})$  – standardna devijacija faktora  $\mathbf{u}$

$r_{t,u}$  – koeficijent korelacije faktora  $\mathbf{t}$  i  $\mathbf{u}$

$\text{var}(\mathbf{t})$  – varijanca faktora  $\mathbf{t}$

$\text{var}(\mathbf{u})$  – varijanca faktora  $\mathbf{u}$

$r_{t,u}^2$  – koeficijent determinacije.

Iz jednažbi (8) i (9) vidljivo je kako kovarijanca kod PLS regresije povezuje dva skupa podataka preko njihovih latentnih varijabli. Faktori  $\mathbf{t}$  i  $\mathbf{u}$  trebaju imati veliku varijancu ( $\text{var}(\mathbf{t})$  i

$\text{var}(\mathbf{u})$ ), ali isto tako moraju biti i međusobno jako povezani ( $r^2_{t,u}$ ).

Kao što je već rečeno predikcija ovisnih varijabli dobiva se iz izraza  $\hat{\mathbf{Y}} = \mathbf{T} \mathbf{B} \mathbf{C}^T$ . Glavno pitanje koje se ovdje postavlja je koliko je latentnih varijabli potrebno zadržati u modelu da se dobije najbolja predikcija novih ovisnih varijabli. Općenito, broj latentnih varijabli se odabire na temelju *cross-validacije* odnosno što je niže moguće vrijednosti korijena srednjeg kvadrata pogreške *cross-validacije* (engl. *Root Mean Square Error of Cross-Validation* - RMSECV) [69].

Vrijednost RMSECV se izračunava iz sljedećeg izraza:

$$\text{RMSECV}_k = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_{-i,k})^2}{n}} \quad (10)$$

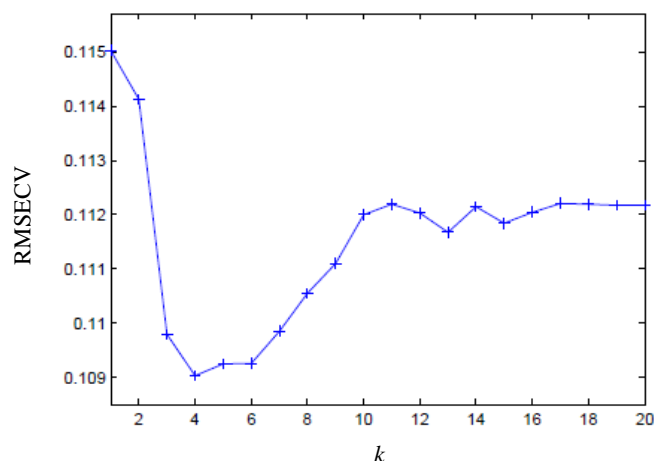
pri čemu je:

$y_i$  – vrijednost  $i$ -te varijable

$\hat{y}_{-i,k}$  - vrijednost  $i$ -te varijable dobivena predikcijom iz modela u koji je uključeno  $k$  latentnih varijabli te je iz modela izostavljena  $i$ -ta varijabla

$n$ - ukupan broj varijabli.

Iz grafičkog prikaza RMSECV o broju latentnih varijabli ( $k = 1, 2, \dots, k_{\max}$ ) moguće je relativno jednostavno odabrati broj faktora za model koji će imati najmanju pogrešku predikcije (slika 3).



Slika 3. Utjecaj broja latentnih varijabli na vrijednost RMSECV.

## 2.2. Genetički algoritmi

Kasnih 80-ih godina prošlog stoljeća u kemometriku je uvedena nova skupina računskih metoda, takozvanih prirodnih računskih metoda. Kao što im i samo ime govori razvoj tih metoda inspiriran je prirodnim procesima. Najznačajnije metode iz te skupine koje se najčešće koriste u kemometrici su: genetički algoritmi, umjetne neuronske mreže i simulirano kaljenje (SA) [65,70].

Genetički algoritmi imaju vrlo široku primjenu, pa se tako osim u kemometrici koriste u ekonomiji za predviđanje kretanja cijena dionica [71], optimiziranju telekomunikacijskih veza [72], izradi i optimiziranju prometnih ruta [73], enkripciji i dekodiranju podataka [74], marketingu i prodaji [75], izradi kompjuterskih igara [76], dizajnu automobila [77], robotici [78], ...

Kao što je pokazano u Literaturnom pregledu, u kemometrici GA se najčešće koristiti za selekciju valnih duljina iz spektralnih podataka, ali i za računalno modeliranje molekula [79], predikciju strukture proteina [80], optimizaciju kemijskih procesa [81], optimizaciju kromatografskih metoda [82],...

U nastavku bit će opisana primjena genetičkih algoritama spregnutih s metodom djelomičnih najmanjih kvadrata za odabir valnih duljina iz spektralnih podataka.

Genetički algoritmi su numeričke optimizacijske metode koje oponašaju biološku evoluciju. Ova vrsta algoritama koristi Darwinovu strategiju „opstanka najjačih“ i reprodukcijске



operatore da bi se poboljšala kvaliteta rješenja određenog problema [70]. Iako se koriste za rješavanje kemometričkih problema, valja naglasiti da GA izravno ne rješavaju problem, već predstavljaju strategiju pretraživanja multivarijatnog prostora. Npr., određeni genetički algoritmi koriste se za selekciju određenih svojstava i/ili optimalnu prilagodbu određenih parametara. Odabrana svojstva ili parametri pomoću GA zatim se prosljeđuju funkciji kojom se vrši evaluacija odabranih parametara kao rješenja određenog problema. Dakle, GA odabiru samo potencijalna rješenja nekog problema, a njihova evaluacija provodi se nekom drugom funkcijom. Kod odabira valnih duljina iz spektralnih podataka evaluacija se najčešće provodi metodom djelomičnih najmanjih kvadrata, multivarijantnom linearnom regresijom ili umjetnim neuronskim mrežama.

Prostor pretraživanja, čak i kod najjednostavnijih problema je velik. Npr., ako neki objekt sadrži  $N$  svojstava te je cilj pronaći skup svojstava koji ga najbolje opisuju ili klasificiraju, jednostavna predodžba mogućeg rješenja problema bio bi niz 0 i 1 dužine  $N$ , pri čemu 1 na nekoj poziciji (*locusu*) označava da je određeno svojstvo odabrano, a 0 da nije (slika 4). Ukupan broj mogućih rješenja za neki broj svojstava  $N$  jednak je  $2^N$  iz čega je vidljivo da čak ako je i broj  $N$  relativno mali (npr.  $N = 50$ ) ukupan broj mogućih rješenja je velik ( $1,126 \cdot 10^{15}$ ). Cilj genetičkih algoritama je pronaći optimalan ili gotovo optimalan skup svojstava procjenjujući samo manji dio mogućih rješenja iz cijelog prostora pretraživanja.

0	1	0	1	0	0	1	0	1	1
---	---	---	---	---	---	---	---	---	---

Slika 4. Prikaz odabranih svojstava za objekt s 10 svojstava ( $N = 10$ ). Svojstva na pozicijama 2, 4, 7, 9 i 10 su odabrana.

U jednostavnom biološkom sustavu sposobnost jednostavnog organizma da opstane u nekom okolišu direktno ovisi o njegovoj sposobnosti da iskoristi hranjive tvari i odupre se štetnim učincima iz okoliša. Tako su npr. prvi organizmi na Zemlji trebali preživjeti u znatno drugačijim uvjetima nego što vladaju danas. Dakle, ti su organizmi trebali imati drugačije skupove enzima od jednostavnih organizama danas. Genetički algoritmi oponašaju upravo tu evoluciju jednostavnih organizama da se prilagode okolišu u kojem žive da bi opstali u njemu.

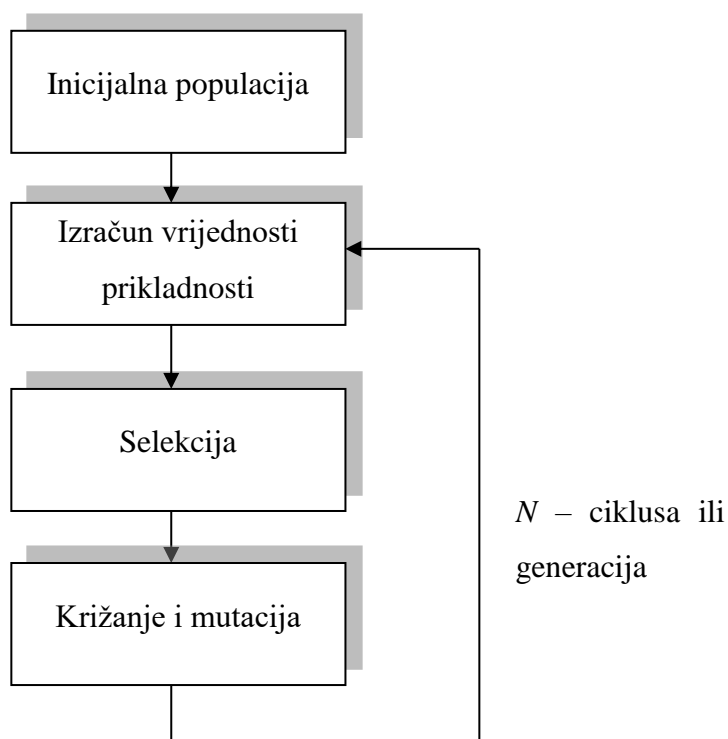
Sposobnost jednostavnog organizma da proizvede određene enzime pohranjena je u kodu koji je zapisan u njegovim kromosomima. Dijelovi kromosoma koji sadrže informacije o nekom enzimu nazivaju se geni, a skup svih gena genotipom. Skup svih enzima koje neki jednostavan organizam ima, a samim time i sposobnost da preživi u nekom okolišu naziva se fenotipom.

Kod genetičkih algoritama kromosom se naziva genetičkim vektorom. Na slici 4 je prikazan jedan takav genetički vektor koji se sastoji od niza 0 i 1. Niz 0 i 1 koji je potencijalno rješenje problema predstavlja genotip organizma, a njegova sposobnost da riješi problem je fenotip. Mjera koliko neko rješenje dobro rješava problem naziva se vrijednost prikladnosti (engl. *fitness*).

Kod razmnožavanja, kromosomi potomaka nastaju kombinacijom dijelova kromosoma roditelja. Prilikom toga postoji mala vjerojatnost da u kromosomu dođe do mutacije prije uspostave cijelog genotipa potomka. Uspostavljeni genotip proizvest će fenotip potomka, a samim time i sposobnost potomka da preživi u nekom okolišu. Darwinov „opstanak najjačih“ općenito određuje koji će se roditelji pariti i hoće li potomak biti održiv da postane roditelj.

Koristeći gore navedeni model moguće je definirati korake jednostavnog genetičkog algoritma.

1. Kreiranje početne populacije genetičkih vektora i izračun njihovih vrijednosti prikladnosti.
2. Odabir dva člana iz početne populacije za razmnožavanje na temelju izračunate vrijednosti prikladnosti.
3. Križanje roditeljskih vektora primjenom operatora križanja pri čemu nastaju novi genetički vektori.
4. Mutacija novonastalog vektora primjenom operatora mutacije.
5. Izračun vrijednosti prikladnosti genetičkog vektora potomka te zamjena najlošijeg vektora iz populacije novim.
6. Ponavljanje čitavog postupka počevši od koraka 2.



Slika 5. Koraci jednostavnog genetičkog algoritma.

Prethodno opisan jednostavni genetički algoritam (slika 5) uspijeva relativno brzo pronaći rješenje problema pretražujući samo mali broj od ukupnog broja svih mogućih rješenja. Taj proces se naziva fokusiranje. Tijekom parenja dio kromosoma se kopira s roditelja na potomka tako dugo dok parenje ili mutacija ne promijeni taj dio. Budući da se roditelji biraju na temelju vrijednosti prikladnosti postoji velika mogućnost da će potomci biti pogodni, tj. da neće biti prerano zamijenjeni budućim potomcima. Kako algoritam napreduje određena struktura će se pojaviti u svim potomcima, tj. na određenim mjestima genetičkih vektora pojaviti će se iste vrijednosti u čitavoj populaciji. Npr., ako vektor ima 10 mjesta, u populaciji se može pojaviti slijedeća shema:

$$(-, -, *, -, *, *, -, *, -, *)$$

Slika 6. Struktura (shema) genetičkog vektora.

Na slici 6 linije predstavljaju mjesta u vektoru koja se mogu mijenjati (pozicije 1, 2, 4, 7 i 9), a mjesta označena zvjezdicama imaju stalne vrijednosti (pozicije 3, 5, 6, 8 i 10). Ukoliko

operator mutacije ne promijeni bilo koju od tih vrijednosti svi će budući potomci imati istu shemu. Utjecaj toga je da se pretraga fokusira na 5 - dimenzionalni prostor rješenja koji predstavljaju mjesta u genetičkom vektoru umjesto na cjelokupni 10 – dimenzionalni prostor. Kako algoritam napreduje loša rješenja iz populacije se zamjenjuju boljim pa zbog toga GA relativno brzo pronalazi optimalna rješenja.

Proces kopiranja dijela genetičkog vektora pogodnih roditelja u vektor potomka naziva se eksploatacija (engl. *exploitation*), a stvaranje potomaka čiji se genetički vektor značajno razlikuje od vektora roditelja naziva se istraživanje (engl. *exploration*). Eksploatacija potiče nastanak određene strukture (sheme) u populaciji i na taj način se pretraga prostora fokusira samo na manji dio ukupnog prostora rješenja. Ukoliko pak operator parenja unese promjenu u genetski vektor roditelja, potomci će biti nakon nekog vremena drugačiji od svojih roditelja te će se na taj način vršiti istraživanje i drugih dijelova prostora. Slično, operator mutacije potiče istraživanje prostora i sprječava stvaranje određenih struktura u nekoj populaciji i fokusiranje pretrage.

GA s najvećom brzinom konvergencije prema optimalnom rješenju imat će operator križanja s jednom ili dvije točke prekida i neće sadržavati operator mutacije. Međutim, kod takvog tipa genetičkih algoritama postoji velika mogućnost da se ne pronađe globalno optimalno rješenje, već samo lokalno. Iz gore navedenog slijedi da će genetički algoritam s dobro uravnoteženom eksploatacijom i istraživanjem trajati nešto duže, jer će biti potrebno neko vrijeme da se uspostavi određena shema u populaciji, ali će postojati veća vjerojatnost da se pronađe bolje globalno rješenje.

### 2.2.1. Kodiranje

Genetički vektori se prije daljnjih koraka algoritma moraju kodirati u računalu razumljiv jezik, tj. moraju se prevesti u binarni kod, odnosno nizove 0 i 1. Ovisno o problemu koji se pokušava riješiti algoritmom upotrebljavaju se različite sheme kodiranja. Najčešće upotrebljavane su: *gene-based*, *node-based* i *delta encoding*.

Kod *gene-based* kodiranja postoje veze jedan naprama jedan između broja gena i određene značajke problema koja se rješava. Upravo kod odabira valnih duljina iz skupa spektralnih podataka koristi se ovaj tip kodiranja. Npr. genetički vektor oblika (0, 0, 1, 1, 0, 1...)

značio bi da valne duljine 1, 2 i 5 nisu, a valne duljine 3, 4 i 6 jesu odabrane iz nekog skupa podataka.

*Node-based* kodiranje predstavlja rutu ili raspored. Npr., ukoliko se u nekom laboratoriju treba napraviti nekoliko eksperimenata, genetički vektor (2, 7, 5...) značio bi da se će se najprije provesti eksperiment 2, zatim eksperiment 7 i na kraju eksperiment broj 5.

*Delta kodiranje* koristi se nešto manje od prethodna dva načina kodiranja. Radi se zapravo o *gene-based* načinu kodiranja s iznimkom da je vrijednost svakog gena izračunana kao relativna vrijednost u odnosu na genetički vektor koji služi kao predložak. Npr., ako je cilj optimizirati uvjete neke kemijske reakcije, genetički vektor kao predložak sadržavao bi neke realne vrijednosti reakcijskih uvjeta (koncentracije, temperature, pH ...). Svako potencijalno rješenje problema bio bi vektor koji bi sadržavao pozitivne i negativne promjene u odnosu na vrijednosti definirane genetičkim vektorom koji služi kao predložak.

Kod odabira valnih duljina *gene-based* kodiranje je vrlo jednostavno, međutim ukoliko je potrebno optimizirati niz realnih parametara potrebno je ograničiti realne parametre na neki konačni broj vrijednosti (slika 7). Ako parametra ima vrijednost cijelog broja npr. 5, genetičkim vektorom duljine 3 bita moguće je ukupno prikazati  $3^2 = 8$  različitih vrijednosti (0-7). Binarni zapis broja 5 tada bi bio 101 ( $V = 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$ ). Međutim, ukoliko se radi o realnom broju  $R$ , koji je potrebno kodirati u binarni broj ( $BN$ ) duljine  $B$  bitova, prije svega potrebno je definirati raspon  $[l, u]$  koji taj broj može imati. Nakon toga je tek moguće izračunati cijeli broj  $V'$  prema izrazu:

$$V' = \frac{R-l}{u-l} (2^B - 1) \quad (11)$$

Broju  $V'$  zatim se pridružuje najbliži cijeli broj  $V$  koji se potom kodira u binarni zapis.

Faktor  $(2^B - 1)$  određuje koliko je maksimalno brojeva moguće prikazati u nekom rasponu pomoću binarnog broja duljine  $B$ . Iz navedenog slijedi da što je veći broj  $B$  to je moguće više brojeva prikazati, tj. može se postići veća rezolucija. Binarni parametar transformira se natrag u realni pomoću izraza:

$$R = l + (u-l) \frac{1}{2^B - 1} V \quad (12)$$

pri čemu je  $V$  cijeli broj koji odgovara binarnom zapisu. Npr., ukoliko realni broj  $R = 3,3$  može imati vrijednosti u rasponu od  $l=0,0$  do  $u = 10,0$  i želi se prikazati u binarnom obliku s duljinom

od  $B = 4$  tada je

$$V' = \frac{3,3 - 0,0}{10,0 - 0,0} (2^4 - 1) = 4,95 \quad (13)$$

Najbliži cijeli broj koji je se zatim pridružuje realnom broju  $V'$  je broj  $V = 5$  čiji je binarni zapis  $BN = 0101$ . Binarni niz 1011 ( $V = 11$ ) predstavlja realni broj 7,3.

$$R = 0,0 + (10,0 - 0,0) \frac{1}{2^4 - 1} 11 = 7,3 \quad (14)$$

Niz 1100 ( $V = 12$ ) predstavlja realni broj 8,0. Iz ovoga je vidljivo da za gore prikazani primjer 4 bitna prezentacija odgovara rezoluciji od 0,7. Da se bi se postigla rezolucija od 0,1 potrebno je omogućiti da faktor  $(2^B - 1)$  bude veći od 101, odnosno broj bitova mora biti veći od 7.

Cijeli broj	<table border="1"><tr><td>74</td><td>10</td></tr></table>	74	10
74	10		
Realni broj	<table border="1"><tr><td>07,4</td><td>10,0</td></tr></table>	07,4	10,0
07,4	10,0		
Binarni broj	<table border="1"><tr><td>10111</td><td>01010</td></tr></table>	10111	01010
10111	01010		

Slika 7. Primjer kodiranja potencijalnih rješenja problema (npr. pH i koncentracija)

Jedan od glavnih problema binarnog način kodiranja je utjecaj operatora mutacije na generirani binarni genetički vektor. Npr., ukoliko je binarni prikaz realnog broja 88,20 dan 8 bitnim vektorom 00101101 i na njega djeluje operator mutacije tako da promijeni vrijednost samo jednog bita iz 0 u 1 tako da dobije npr. vektor 10101101, realni broj koji predstavlja novonastali vektor je 339,08. Iz ovog je primjera vidljivo da je pomoću ovog načina kodiranja vrlo teško kontrolirati razmjer utjecaja čak i samo jedne mutacije.

Da bi se smanjio utjecaj operatora mutacije na potencijalna rješenja koristi se Grayovo kodiranje (tablica 1). Grayov kod predstavlja cijele brojeve  $(0, 1, 2, \dots, 2^N - 1)$  kao binarni niz duljine  $N$  tako da se Grayove reprezentacije susjednih cijelih brojeva razlikuju samo u jednom bitu.

Tablica 1. Primjer Grayovog kodiranja za cijele brojeve od 0 do 15.

Broj	Standardni binarni broj	Grayov broj
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

Za svaki  $N$  postoji nekoliko načina Grayovog kodiranja. Jedan od primjera Grayovog kodiranja koristi sljedeću proceduru:

$$g(1) = b(1) \quad \text{za } i = 2 \text{ do } N \quad (15)$$

$$g(i) = b(i) \quad \text{ako je } b(i-1) = 0 \quad (16)$$

$$g(i) = \text{KOMPLEMENT}(b(i)) \quad \text{ako je } b(i-1) = 1 \quad (17)$$

gdje je  $b(i)$   $i$ -ti element normalnog binarnog niza,  $g(i)$  odgovarajuća vrijednost iz Grayovog binarnog niza,  $b(1)$  i  $g(1)$  su prvi (krajnje lijevi) elementi binarnog odnosno Grayovog niza, a  $N$  je duljina nizova. U tablici 1 prikazane su reprezentacije prvih 15 cijelih brojeva u standardnom binarnom i Grayovom kodu.

Grayov kod iz prethodnog primjera vrlo se lako može pretvoriti u standardni binarni kod izrazima kako slijedi:

$$b(1) = g(1) \quad \text{za } i = 2 \text{ do } N \quad (18)$$

$$b(i) = g(i) \qquad \text{ako je } b(i - 1) = 0 \qquad (19)$$

$$b(i) = \text{KOMPLEMENT}(g(i)) \qquad \text{ako je } b(i - 1) = 1 \qquad (20)$$

Kao što je već rečeno primjenom Grayovog kodiranja smanjuje se utjecaj operatora mutacije na rezultat. Npr., ako se za kodiranje koristi 8 bitni niz tada postoji 25% mogućih promjena koje mijenjaju rezultat samo za 1, a ostalih 75% uzrokuje veće promjene. Npr., ako je cijeli broj 92 reprezentacija neke realne varijable njegov Grayov kod za taj broj bio bi (01110010). Ako se promijene šesti ili osmi bit prikazanog broja, promijenit će se i odgovarajući cijeli brojevi, ali samo za 1. Ukoliko se promijeni sedmi bit cijeli broj će se promijeniti za 3. Međutim, ukoliko se promijene prvi ili drugi bit odgovarajući cijeli brojevi su 163 i 35.

Također je važno naglasiti da način kodiranja varijabli ne samo da utječe na operator mutacije već i na promjenu izgleda površine prostora pretraživanja, što je najbolje vidljivo iz sljedećeg primjera. Ukoliko pretpostavimo da je cijeli broj 92 (predstavlja realnu varijablu 180,32) dobro rješenje nekog problema, a cijeli broj 163 (realna vrijednost 319,48) najbolje rješenje problema, te da se koristi standardno binarno kodiranje nakon nekog vremena moguće je da se uspostavi struktura genetičkog vektora (-101-10-) koja se nalazi u blizini cijelog broja 92. Iz ovoga je vidljivo da najbolje rješenje neće biti nikada pronađeno budući da se standardna binarna reprezentacija broja 163 (10100011) razlikuje od standardne binarne reprezentacije broja 92 (01110010) u svim bitovima. Međutim, ukoliko se na gore navedene vrijednosti primjeni Grayovo kodiranje shema koja se uspostavlja oko broja 92 je (-111-01) čime oba rješenja, 92 (01110010) i 163 (11110010), postaju jednako vjerojatna jer se Grayovi kodovi za te brojeve razlikuju samo u vrijednosti prvog bita. Zamjena standardnog binarnog kodiranja Grayovim spojila je dva optimuma, koji su u standardnom binarnom kodiranju bili maksimalno odvojeni, u jedno novo (dobro) područje prostora pretraživanja. Iz ovoga slijedi da promjena načina kodiranja utječe na promjenu udaljenosti između optimuma što direktno utječe na pronalaženje najboljeg globalnog rješenja.



### 2.2.2. Inicijalizacija populacije i evaluacija rješenja

Nakon što se odabere način kodiranja varijabli potrebno je odabrati početnu populaciju iz koje će se zatim odabrati roditelji za parenje. Glavno pitanje koje se postavlja kod odabira inicijalne populacije je koliko je potencijalnih rješenja potrebno uključiti u početnu populaciju. Ako je  $n_p$  broj parametara koji se kodira u binarne nizove duljine  $n_b$  bitova intuitivan red veličine populacije je:

$$n_{\text{pop}} \approx n_p 2^{n_b} \quad (21)$$

Npr., ako je potrebno naći optimum vrijednosti 20 parametara, i svaki od parametara je kodiran u binarni niz duljine 8 bitova, broj mogućih rješenja bio bi reda 5000. Ovo je prilično velika populacija kojom je osigurano da je gotovo čitavi prostor pretraživanja uključen u optimizaciju. Međutim za većinu praktičnih problema najčešće se u inicijalnu populaciju uključuje nasumičnim odabirom od 50 do 100 potencijalnih rješenja. Smanjenjem veličine inicijalne populacije smanjuje se i prostor pretraživanja i znatno ubrzava proces optimizacije odnosno pronalaska optimalnog rješenja. Iz inicijalne populacije se zatim generiraju sljedeće generacije koje sadrže sve bolja i bolja rješenja.

Nakon odabira inicijalne populacije vrši se evaluacija kvalitete odabranih rješenja. Kvaliteta potencijalnih rješenja evaluira se pomoću vrijednosti prikladnosti. Vrijednosti prikladnosti dobivaju se izračunom iz evaluacijske funkcije  $y = f(x)$ , gdje je  $x$  binarna reprezentacija jednog potencijalnog rješenja (genetički vektor), a  $y$  je vrijednost prikladnosti baš za to rješenje. Upravo je izračun vrijednosti prikladnosti najčešće računski najzahtjevniji korak kod genetičkih algoritama. Evaluacijski korak također sadrži i provjeru da li potencijalno rješenje zadovoljava sva ograničenja prostora pretraživanja. Ukoliko rješenje ne odgovara postavljenim zahtjevima mogući su sljedeći koraci:

- Odbacivanje - vektori se odbacuju ukoliko ne zadovoljavaju ograničenje
- Kažnjavanje (engl. *penalization*) - vektorima koji ne zadovoljavaju sve kriterije smanjuje se vrijednost prikladnosti
- Popravljanje - vektori se popravljaju prema određenoj strategiji tako da nakon toga postaju prihvatljivi. Npr., ako je potencijalno rješenje izvan postavljenih granica,

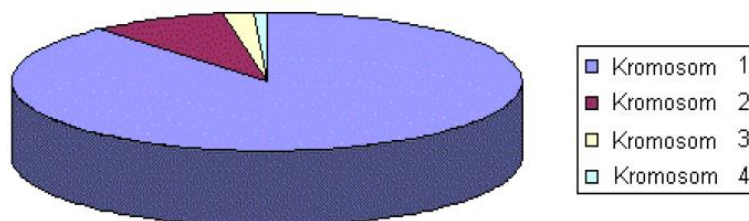
vrijednost parametra se mijenja tako da rješenje nakon toga bude unutar postavljenih granica.

Sljedeći korak je provjera da li se algoritam može zaustaviti, odnosno da li rješenje zadovoljava postavljenu vrijednost prikladnosti. U inicijalnoj populaciji mala je vjerojatnost da će se naći takvo rješenje, budući da se inicijalna populacija formira nasumično. U ovaj korak uključuje se i kriterij konvergencije. Ukoliko se rješenja nekoliko uzastopnih generacija znatno ne poboljšavaju algoritam se može zaustaviti. Vrlo jednostavan kriterij zaustavljanja algoritma, koji se također vrlo često koristi, je broj maksimalnih generacija kojim se zaustavlja algoritam ukoliko se ne postigne konvergencija.

### 2.2.3. Selekcija rješenja

Selekcija se koristi za odabir roditelja koji će se međusobno križati, odnosno koji će se od potomaka uključiti u sljedeću generaciju. Svrha ovog postupaka je očuvanje dobrih svojstava, tj. gena i njihovo prenošenje na iduću generaciju jedinki. Upravo ovaj korak je inspiriran Darwinovom teorijom opstanka najjačih.

Najčešće korištena metoda selekcije poznata je pod nazivom metoda kotača za roulette (engl. *roulette wheel*). Osnovna ideja ovog načina odabira je da se sva rješenja stave na kotač za roulette i to tako da je površina koju pojedino rješenje zauzima na kotaču proporcionalna vrijednosti prikladnosti (slika 8). Drugim riječima rješenja s najvećom vrijednosti prikladnosti imaju najveću vjerojatnost da budu odabrana za križanje.



Slika 8. Primjer selekcije rješenja metodom kotača za roulette.

Najjednostavnija procedura je generirati nasumični broj između nule i sume svih vrijednosti prikladnosti i zatim sumirati pojedinačne vrijednosti prikladnosti tako dugo dok ta suma ne

postane veća ili jednaka nasumično generiranom broju. Posljednje rješenje prilikom sumiranja postaje jedan od roditelja. Matematički se nasumični broj  $R$  određuje prema izrazu:

$$R = r \sum_{i=1}^{\mu} f_i \quad (22)$$

pri čemu je  $r$  slučajni broj u području  $[0, 1]$ ,  $\mu$  veličina populacije, a  $f_i$  vrijednost prikladnosti  $i$ -tog člana populacije.

$K$ -ti član populacije se odabire tako da je:

$$\sum_{i=1}^k f_i \geq R \quad (23)$$

a vjerojatnost da neko rješenje  $i$  bude izabrano kao roditelj dana je izrazom:

$$p_i = \frac{f_i}{\sum_{i=1}^{\mu} f_i} \quad (24)$$

Međutim, kod gore opisane metode može postati dominantno samo jedno ili dva rješenja (rješenja s najvećim vrijednostima prikladnosti) što može uzrokovati preranu konvergenciju rješenja. Da bi se to spriječilo vrijednosti prikladnosti se prije stavljanja na kotač skaliraju. Skaliranje se provodi linearno i to tako da je:

$$f'_i = a f_i + b \quad (25)$$

Koeficijenti  $a$  i  $b$  mijenjaju se s veličinom populacije.

Druga metoda koja se također vrlo često koristi naziva se  $K$ -turnirska (engl. *tournament*) selekcija. Iz populacije se odabire nasumično  $K$  članova od kojih se odabire član s najvećom vrijednosti prikladnosti. Broj  $K$  je cijeli broj između 2 i  $(N_p - 1)$  ( $N_p$  – veličina populacije).

Osim, već spomenutih metoda selekcije postoje još i druge metoda od kojih su najznačajnije:

- Linearna selekcija – vjerojatnost odabira je ovisna o rangui prikladnosti, a ne o vrijednosti prikladnosti
- *Threshold* selekcija – ovim se postupkom iz populacije uklanja  $N_w$  najgorih rješenja, a od preostalih rješenja ( $N_p - N_w$ ) se nasumično odabiru rješenja za novu generaciju. Ovim

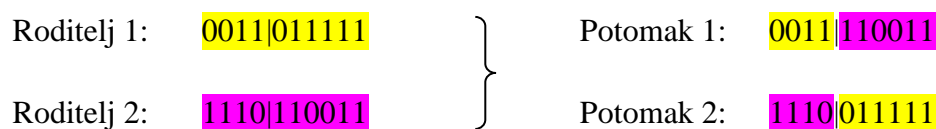
se postupkom sprječava prelazak najgorih rješenja u novu generaciju. Broj  $N_w$  je cijeli broj s vrijednošću između 0 i  $(N_p-2)$

- *Elitist* selekcija – ovim se postupkom osigurava da se iz populacije bezuvjetno (deterministički) odabere  $N_e$  najboljih rješenja, a ostala se rješenja zatim odabiru jednom od prethodno opisanih metoda. Za razliku od *threshold* selekcije ovom metodom se osigurava prijenos najboljih rješenja u sljedeću generaciju.

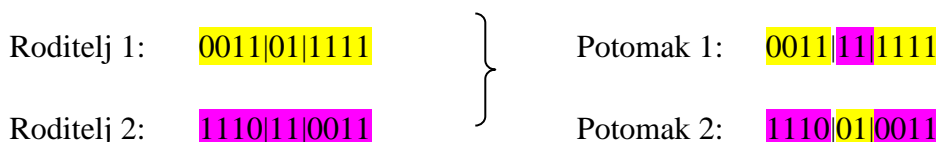
Nenasumična selekcija u populaciji uzrokuje da je svaka sljedeća generacija bolja od one prethodne. Ovdje je važno naglasiti da se ovim postupcima (osim *threshold* selekcije) najlošija rješenja ne isključuju iz postupka, ali postoji znatno manja vjerojatnost da se prenesu u sljedeću generaciju rješenja čime se sprječava prerana konvergencija.

#### 2.2.4. Križanje

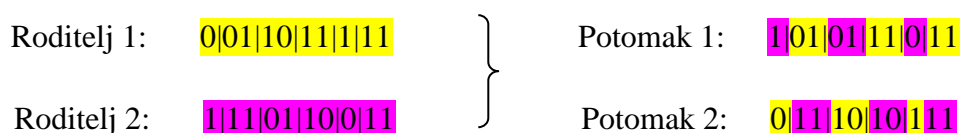
U prethodno opisanom procesu selekcije iz populacije se odabiru najbolja rješenja iz populacije čime se osigurava da buduće generacije sadržavaju dobra rješenja. Međutim, da bi se riješio određeni problem očuvanje najboljih rješenja nije dovoljno. Da bi se pronašlo optimalno rješenje i da bi se obuhvatio čitav prostor pretraživanja u populaciju je potrebno unijeti određenu varijaciju, ali na način da najvažnija informacija, koja je već prisutna u populaciji, ipak bude očuvana. Uvođenje varijacije u populaciju postiže se upotrebom genetičkih operatora križanja (rekombinacije) i mutacije. Križanje je operator koji imitira prirodni proces križanja (engl. *crossover*). Tim procesom rekombiniraju se dijelovi dvaju potencijalnih rješenja (roditelji) u novo jedno ili dva rješenja (potomci). Najjednostavniji tip križanja je križanje s jednom točkom prekida i prikazan je na slici 9a.



a) Križanje s jednom točkom prekida



b) Križanje s dvije točke prekida



c) Uniformno križanje

Slika 9. Križanje genetičkih vektora.

Dva genetička vektora podijele se u dva dijela u jednoj nasumično odabranoj točki i zatim se njihovi dijelovi međusobno rekombiniraju. Ideja čitavog ovog postupka je da će nastali potomci biti dobra, odnosno još i bolja rješenja problema od roditelja budući da su roditelji (odabrani procesom selekcije) dobra rješenja. Rekombinacija parova potencijalnih rješenja provodi se s vjerojatnošću  $P_r$  koja se najčešće odabire u rasponu vrijednosti od 0,5 do 1.

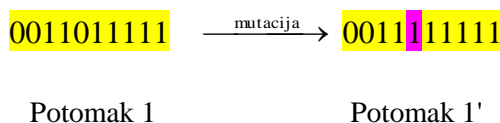
Iz gore opisanog primjera vidljivo je da se ovim operatorom u genetički algoritam uvode nova potencijalna rješenja koja se ne moraju ni nalaziti u inicijalnoj populaciji. Smatra se da se upravo zbog ovog koraka genetički algoritmi najviše razlikuju od ostalih optimizacijskih metoda. Također, budući da iz dva različita rješenja može proizaći novo (bolje) rješenje genetički algoritmi su vrlo moćan alat za pretraživanje prostora.

Osim križanja s jednom točkom prekida koristi se još i križanje s dvije točke prekida. Kod ovog tipa križanja kod svakog od roditeljskih vektora nasumično se odaberu dvije točke prekida, a zatim se razmijene dijelovi vektora koji se nalaze između točaka prekida (slika 9b).

Krajnji slučaj ovog tipa križanja je uniformno križanje kod kojeg se između dva roditelja zamjenjuje određeni broj bitova na nasumično odabranim mjestima (slika 9c).

### 2.2.5. Mutacija

Mutacija je genetički operator kojim se nasumično lokalno mijenja vrijednost bita u samo jednom kromosomu (slika 10). Ovim operatorom može se unijeti novo svojstvo u populaciju ili obnoviti izgubljeno. Npr., u slučaju kad je u čitavoj populaciji vrijednost posljednjeg bita u genetičkom vektoru jednaka 1, križanjem se ne može unijeti potrebna varijacija u populaciju. To znači da su u takvoj populaciji prisutna samo neparna rješenja, te u takvoj populaciji nije moguće dobiti paran broj primjenom samo operatora križanja. Genetički algoritam u tom slučaju može vrlo lako postati „zarobljen“ u lokalnom minimumu te neće pronaći optimalno rješenje problema.



Slika 10. Mutacija gena u kromosomu.

Vjerojatnost da se mutacija pojavi u bilo kojem od bitova genetičkog vektora jednaka je  $P_m$  i u većini slučajeva se odabire da bude mala, ispod 0,05. Nasumično se generira broj  $R$  u rasponu  $[0, 1]$  i ukoliko je on manji od postavljene vrijednosti vjerojatnosti mutacije  $P_m$  promijeni se vrijednost nasumično odabrane pozicije u genetičkom vektoru (slika 10). Veća vjerojatnost mutacije može vrlo lako uništiti nastalu strukturu koja se formira u populaciji tijekom napredovanja algoritma. Ukoliko je vjerojatnost mutacije jednaka jedan, tada je algoritam zapravo algoritam traženja slučajnom pretragom prostora rješenja.

Nakon primjene operatora križanja i mutacije trenutna se populacija može zamijeniti privremenom novom populacijom. U slučaju kad je nova populacija  $N'$  jednako velika kao i prethodna  $N_p$  zamjena je neposredna. Cijela trenutna generacija zamjenjuje se privremenom novom. Ukoliko je pak nova populacija manja od prethodne dio jedinki iz prethodne generacije rješenja mijenja se novima. Ova vrsta strategije naziva se algoritam ustaljenog stanja s

generacijskom prazninom,  $G = N/N_p$ . Međutim, kod ove vrste zamjene populacija potrebno je pripaziti da se dobra rješenja iz trenutne populacije ne zamjene lošijim rješenjima iz privremene. Da bi se to spriječilo moguće je primijeniti jedan od prethodno opisanih postupaka selekcije za odabir jedinki iz trenutne populacije za zamjenu. Ovim korakom završava jedan ciklus ili generacija genetičkog algoritma. Čitav ovaj postupak ponavlja se mnogo puta, a ovisno o složenosti postupka ponekad je potrebno zamijeniti i nekoliko tisuća generacija da bi se postigla zadovoljavajuća konvergencija.

#### 2.2.6. Odabir valnih duljina GA-PLS algoritmom

Spektralni skupovi podataka mogu se sastojati od nekoliko stotina pa do nekoliko tisuća podataka koji se danas mogu prikupiti za svega nekoliko sekundi. Metode kao što su PLS i PCA omogućavaju da se u račun uzme cijeli spektar bez da se prethodno odabere samo jedan njegov dio ili više dijelova. Zbog mogućnosti da navedene metode „izvuku“ samo važne informacije iz spektra smatralo se da su te metode neovisne o šumu i da prethodno nije potrebno odabrati samo „važne“ dijelove spektra. Međutim, kasnije je utvrđeno da je odabirom samo važnih varijabli moguće znatno poboljšati korištene matematičke modele i smanjiti njihovu složenost. Danas postoji veliki broj empirijskih dokaza da je odabir varijabli za PLS i PCA metode vrlo važan i koristan [32,70,83–85].

Iz spektra nekog jednostavnog sustava (npr. spektar jednodimenzionalnog materijala) moguće je relativno jednostavno odabrati valne duljine pri kojima je materijal spektralno aktivan. Međutim, kod spektara višekomponentnih materijala problem je znatno složeniji. Moguće je npr., da neke od komponenata nisu poznate, moguće je njihovo spektralno preklapanje ili njihova kemijska i fizikalna interakcija te više nije moguće jednostavno odabrati valne duljine pri kojima je svaka pojedina komponenta spektralno aktivna.

Zbog toga je razvijeno nekoliko različitih računskih metoda pomoću kojih je moguće iz čitavog spektra odabrati samo „značajne“ valne duljine. Neke najčešće korištene metode su: *Interactive Variable Selection*, *Uninformative Variable Elimination*, *Iterative Predictor Weighting PLS*, testovi značajnosti parametara modela i genetički algoritmi [86]. Glavna prednost odabira valnih duljina za izradu modela je znatno poboljšanje točnosti i preciznosti modela, povećanje njegove robustnosti i smanjenje kompleksnosti.

Dok je većina genetičkih algoritama razvijena da bi se koristila u kontinuiranom

prostoru, GA-PLS algoritam pretražuje rješenja u diskontinuiranom prostoru, točnije pretraga se svodi na pretraživanje samo bridova  $v$ -dimenzijske hiperkocke. Upravo su zbog toga napravljene određene promjene u GA-PLS algoritmu u odnosu na jednostavni genetički algoritam. Promjene kao i ostali koraci GA-PLS algoritma opisani su u nastavku ovog rada.

#### 2.2.6.1. Inicijalizacija populacije

U jednostavnom genetičkom algoritmu odabir varijabli za početnu populaciju vršio bi se nasumičnim odabirom te bi prema tome od ukupnog broja varijabli iz čitavog spektra bilo odabrano 50 % valnih duljina. Međutim, takav način odabira inicijalne populacije imao bi dva velika nedostatka:

- izračun bi trajao vrlo dugo za skupove podataka s mnogo varijabli
- budući da PLS ima vrlo malu osjetljivost na šum, prisutnost malog broja „dobrih“ valnih duljina bila bi dovoljna da se dobije dobar odziv, bez obzira na prisutnost loših varijabli. Drugim riječima odzivi različitih kromosoma bili bi vrlo slični pa bi se zbog toga bilo vrlo teško riješiti „loših“ valnih duljina iz kromosoma.

Zbog toga se selekcija određenog broja varijabli za inicijalnu populaciju vrši postavljanjem vjerojatnosti,  $n/v$ , gdje je  $n$  broj selektiranih varijabli, a  $v$  ukupan broj varijabli. Odabir manjeg broja varijabli smanjuje vrijeme izračuna i olakšava detekciju samo relevantnih varijabli. Osim navedenog moguće je postaviti i maksimalni broj odabranih valnih duljina da se spriječi pojava kromosoma čijom se evaluacijom dobivaju kompleksni modeli s previše latentnih varijabli.

#### 2.2.6.2. Razmnožavanje i mutacija

Kako bi se izbjegao bilo kakav utjecaj redoslijeda varijabli kao operator razmnožavanja koristi se operator uniformnog križanja (slika 9c) Budući da se općenito operatorima križanja naglašava istraživanje, a kako je cilj kod odabira valnih duljina naglasiti eksploataciju, za rješavanje ove vrste problema koristi se ne-generacijski algoritam koji se još naziva i algoritam s nultom generacijskom prazninom. Kod ove vrste algoritama potomak koji se zadržava u populaciji odmah zamjenjuje jednog člana populacije i odmah se može koristiti za



razmnožavanje.

Nakon generiranja potomaka na njih djeluje operator mutacije nakon čega se vrši evaluacija kvalitete rješenja. Najlošije jedinke iz trenutne populacije zamjenjuju se novim boljim potomcima.

### 2.2.6.3. Evaluacija

Evaluacija rješenja vrši se najčešće izračunom vrijednosti korijena srednjeg kvadrata pogreške previđanja ili vrijednosti srednjeg kvadrata pogreške *cross-validacije* (j. 10). Vrijednost RMSEP računa se prema sljedećem izrazu:

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (26)$$

pri čemu je:

$y_i$  – vrijednost  $i$ -te varijable

$\hat{y}_i$  - vrijednost  $i$ -te varijable dobivena predikcijom iz modela

$n$ - ukupan broj varijabli.

Vrijednosti varijabli  $\hat{y}_i$  izračunavaju se iz PLS modela koji se izračunava za svaki kromosom (rješenje) dobiven genetičkim algoritmom. Iz ovoga je vidljivo da je cilj GA-PLS algoritma pronaći kombinaciju varijabli koja ima najveću predikcijsku moć, odnosno ima najmanju vrijednost kvadrata pogreške predikcije ili srednjeg kvadrata pogreške *cross-validacije*.

### 2.2.6.4. Ubacivanje novog kromosoma u populaciju

Kvaliteta skupa odabranih varijabli određena je njihovom vrijednosti prikladnosti i njihovim brojem. Najbolja rješenja za određeni broj varijabli sadrže vrlo važne informacije i potrebno ih je „zaštiti“ da se ne uklone iz populacije. Kod ove vrste algoritma zaštićeni kromosom se zamjenjuje novim samo u slučaju kad novi kromosom ima jednak broj odabranih varijabli kao

i zaštićeni i daje bolju vrijednost prikladnosti od njega.

Nakon evaluacije kvalitete svakog potomka odlučuje se hoće li se on ubaciti u populaciju te koji će se od kromosoma iz populacije odbaciti. Ukoliko je novi kromosom zaštićen, on će postati članom populacije, a iz nje će se izbaciti nezaštićeni kromosom s najlošijom prikladnosti. Novi nezaštićeni kromosom postat će članom populacije samo ukoliko ima bolju prikladnost od najgoreg nezaštićenog kromosoma, koji će u tom slučaju biti uklonjen iz nje.

Tablica 2. Simulirana populacija 10 kromosoma.

Kromosom	Prikladnost	Broj odabranih varijabli
1	80.35	7*
2	78.32	7
3	75.47	8
4	70.32	5*
5	68.43	6
6	65.13	5
7	60.65	4*
8	50.43	3*
9	40.71	2*
10	30.09	1*

\* Zaštićeni kromosomi

U tablici 2 prikazana je simulirana populacija od 10 kromosoma. Vidljivo je da su slijedeći kromosomi zaštićeni: 1 (najveća vrijednost prikladnosti za  $v \leq 7$ ), 4 (za  $v \leq 5$ ), 7 (za  $v \leq 4$ ), 8 (za  $v \leq 3$ ), 9 (za  $v \leq 2$ ) i 10 (za  $v \leq 1$ ). Kromosom broj 5 nije zaštićen iako je najbolje rješenje s 6 varijabli jer kromosom broj 4 ima veću vrijednost prikladnosti s manjim brojem varijabli.

Način odlučivanja hoće li novi kromosom ući u prikazanu populaciju ili ne prikazan je u sljedeća četiri primjera:

- a) novi kromosom ima vrijednost prikladnosti 55,13 s  $v = 4$ ; kromosom je nezaštićen i budući da ima nižu vrijednost odziva od najlošijeg nezaštićenog kromosoma broj 6 (65,13) odbacuje se.
- b) novi kromosom ima vrijednost prikladnosti 66,19 s  $v = 7$ ; kromosom je nezaštićen i ima veću vrijednost prikladnosti od najlošijeg nezaštićenog kromosoma broj 6

- (65,13); novi kromosom ulazi u populaciju, a kromosom broj 6 se odbacuje.
- c) novi kromosom ima vrijednost prikladnosti 72,14 s  $v = 5$ ; kromosom je zaštićen jer ima veću vrijednost prikladnosti od zaštićenog kromosoma broj 4 s  $v = 5$ , (70,32); novi kromosom ulazi u populaciju, kromosom 4 postaje nezaštićen, a kromosom broj 6 se odbacuje.
- d) novi kromosom ima vrijednost prikladnosti 42,11 s  $v = 2$ ; kromosom je zaštićen jer ima veću vrijednost odziva od zaštićenog kromosoma broj 2 (40,71); novi kromosom ulazi u populaciju, a kromosom broj 9 se odbacuje.

#### 2.2.6.5. Kontrola ponavljanja rješenja

Tijekom jednostavnog genetičkog algoritma ne vrši se provjera ponavljanja istog rješenja u populaciji, tako da isti kromosom može biti prisutan u populaciji više puta. Kod algoritma za odabir valnih duljina iz spektralnih skupova podataka to nije dozvoljeno. Za svaki novi kromosom vrši se provjera da li u populaciji već postoji jedan takav isti. Ukoliko se utvrdi postojanje kromosoma „blizanca“ novo rješenje se odbacuje. Uvođenjem ove modifikacije u algoritam svaka formirana populacija sastoji se od novih, informativnih kombinacija varijabli.

#### 2.2.6.6. Utjecaj različitih parametara na GA-PLS algoritam

Na rezultate genetičkog algoritma za odabir valnih duljina utječu četiri parametra: veličina populacije, vjerojatnost početne selekcije,  $n/v$  (gdje je  $n$  broj selektiranih varijabli, a  $v$  ukupan broj varijabli), maksimalni broj varijabli i vjerojatnost mutacije  $P_m$ .

Proučavanjem velikog broja spektralnih skupova podataka utvrđeno je da postoji određena kombinacija navedenih parametara koja daje dobre rezultate te za svaki novi skup podataka nije potrebno vršiti njihovu optimizaciju.

Utvrđeno je da veličina populacije od 30 kromosoma s ne-generacijskim algoritmom daje dobru uravnoteženost istraživanja i eksploatacije.

Mala vrijednost vjerojatnosti inicijalne selekcije omogućava da se u kraćem vremenu istraži više kombinacija odabranih varijabli te da je optimalna vrijednost za ovaj parametar  $5/v$  odnosno prosječno odabranih 5 varijabli po kromosomu.

Kao što je već rečeno što je veći broj odabranih varijabli to je i model složeniji, a samim

time je i duže vrijeme izračunavanja vrijednosti prikladnosti. Maksimalan broj od 30 varijabli omogućava da se algoritmom dobiju dobri modeli koji nisu računski prezahtjevni.

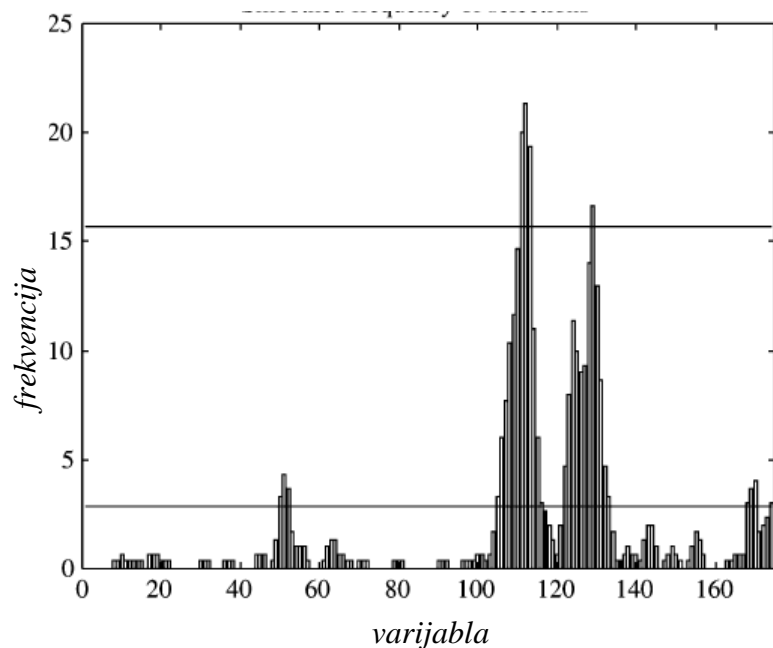
Vjerojatnost mutacije po genu od 0,01 (1%) omogućava da se u populaciju uvedu nove kombinacije odabranih valnih duljina, ali da se pri tome bitno ne promijene one već prisutne.

#### 2.2.6.7. Provjera podskupa varijabli u kromosomima

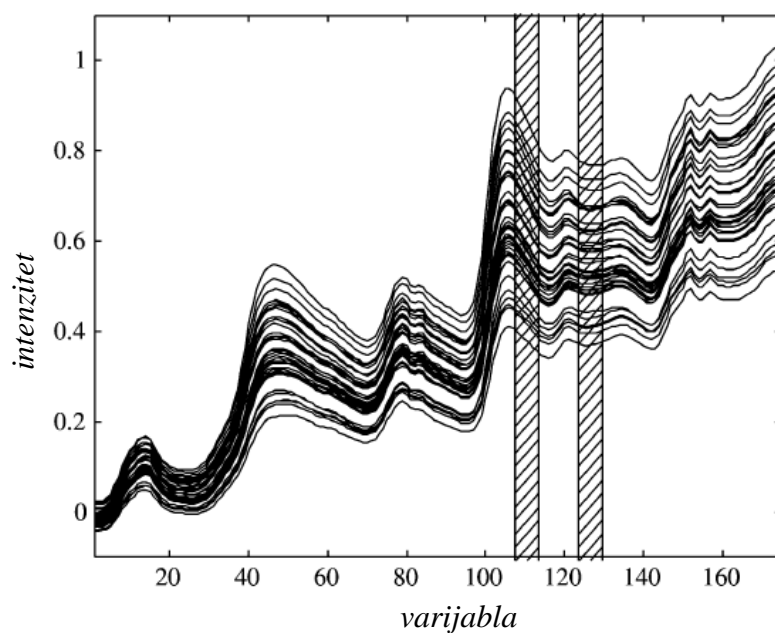
Kao što je već rečeno prilikom provođenja GA-PLS algoritma vrši se provjera ponavljanja kromosoma u populaciji. Međutim, potrebno je provesti i provjeru podskupa odabranih varijabli u kromosomima. Budući da kromosom sadrži više varijabli postoji mogućnost da jedan kromosom sadrži podskup varijabli drugog, odnosno dio odabranih varijabli koje sadrži neki kromosom  $c_1$  može biti odabrano i u drugom kromosomu  $c_2$ . Ukoliko kromosom  $c_2$  ima veću prikladnost od kromosoma  $c_1$  tada dodatne varijable prisutne u kromosomu  $c_1$  ne sadrže dodatne informacije, već predstavljaju samo šum. Uvrštavanjem kromosoma  $c_1$  u populaciju izgubio bi se dio njene raznovrsnosti, a da se pri tome ne unese dio novih informacija. Zbog toga se u takvim slučajevima kromosom  $c_1$  odbacuje.

#### 2.2.6.8. Ponavljanje GA-PLS algoritma

Krajnji rezultat jednog genetičkog algoritma je model u koji je uključeno samo nekoliko varijabli od njihova ukupnog broja. Moguće je da je u model uključen i dio ne-relevantnih varijabli, a da s druge strane neke relevantne iz cijele populacije nisu ni uključene. Uzrok tome su parametri algoritma koji su postavljeni tako da je naglašena eksploatacija, odnosno brzo pronalaženje rješenja da bi se izbjegla precjena (engl. *overfit*) podataka. Zbog toga dio prostora rješenja tijekom izvođenja algoritma uopće neće biti ni istražen, a krajnje rješenje će znatno ovisiti i o slučajnom odabiru početne populacije. Zbog svega navedenog potrebno je provesti veliki broj različitih ponavljanja genetičkog algoritma i pri tome izračunavati frekvenciju selekcije određene varijable tijekom napredovanja algoritma. Krajnji model se dobiva tako da se u njega uključe samo varijable koje su najčešće bile selektirane tijekom većeg broja ponavljanja algoritma (slika 11). Na slici 12 su prikazana spektralna područja koja su bila odbrana najviše puta tijekom izvođenja GA-PLS algoritma.



Slika 11. Frekvencija odabira valnih duljina tijekom provođenja GA-PLS algoritma.



Slika 12. Spektralna područja odabrana iz NIR spektara na temelju frekvencije selekcije (slika 11).

#### 2.2.6.9. Uključivanje autokorelacije između spektralnih varijabli

Spektralni podaci imaju jednu posebnost, a to je da su spektralne varijable vrlo jako autokorelirane (varijabla  $n$  vrlo jako je korelirana s varijablama  $n - 1$  i  $n + 1$ ). Što znači da ako je varijabla  $n$  odabrana tijekom GA kao relevantna, tada i varijable  $n - 1$  i  $n + 1$  također imaju

veliku vjerojatnost da budu odabrane. Međutim, kao što je već prije rečeno zbog naglašene eksploatacije i brzine GA nije vjerojatno da će algoritam odabrati čitavo spektralno područje oko varijable  $n$ . Promjenom vektora vjerojatnosti odabira inicijalne populacije,  $p = n/v$ , uključuje se autokorelacija među varijablama u algoritam. U prvom ciklusu GA sve varijable za inicijalnu populaciju biraju se s jednakom vjerojatnošću  $p$ , međutim u sljedećim ciklusima GA vjerojatnost  $p$  više nije jednaka za sve varijable već se ona mijenja ovisno o frekvenciji odabira pojedine varijable u prethodnom ciklusu algoritma. Tako varijable koje su najviše puta bile odabrane tijekom ciklusa algoritma imaju i najveću vjerojatnost da budu izabrane u novu populaciju. Time se usmjerava odabir varijabli za novu populaciju sljedećeg ciklusa GA na područje spektra od najvećeg interesa.

Novi elementi vektora vjerojatnosti (vagani vektor vjerojatnosti) računaju se prema sljedećim izrazima:

$$pf_i = \frac{\frac{n}{v}(R-r) + ps_i r}{R} \quad (27)$$

$$ps_i = \frac{(p_i + p_{i-1} + \dots + p_{i-k+1})}{k} \quad p_i = n \frac{sel_i}{\sum_{j=1}^v sel_j} \quad (28)$$

gdje je

$pf_i$  – konačna vjerojatnost da se  $i$ -ta varijabla pojavi u kromosomu u inicijalnoj populaciji

$R$  – ukupni broj ciklusa genetičkog algoritma

$r$  – broj provedenih ciklusa GA

$ps_i$  – vjerojatnost  $i$ -te varijable nakon izgladivanja

$k$  – veličina prozora

$n$  – prosječan broj odabranih varijabli po kromosomu za inicijalnu populaciju prvog ciklusa GA

$v$  – ukupan broj varijabli

$sel_i$  – broj (frekvencija) odabira  $i$ -te varijable nakon prvog ciklusa GA

$sel_j$  – broj (frekvencija) odabira  $j$ -te varijable nakon prvog ciklusa GA

Zahvaljujući prethodno opisanoj proceduri moguće je pomoću odabranih varijabli detektirati „značajne“ spektralne regije kao i smanjiti raznolikost dobivenih modela za isti skup podataka.

#### 2.2.6.10. Zaustavljanje algoritma

Vrlo važan korak prilikom provođenja GA je zaustavljanje algoritma. Općenito, u početku GA sa svakom novom generacijom vrijednost prikladnosti se vrlo brzo mijenja, a zatim je ta promjena sve manja i manja. Može se reći da se u početku algoritma poboljšava matematički model, ali nakon toga se počinje u njega uključivati sve više i više šuma. Zbog toga je potrebno zaustaviti algoritam vrlo brzo da se izbjegne uključivanje previše šuma u model. Primijećeno je da je najbolje zaustaviti algoritam između 50 i 200 ciklusa da se na taj način izbjegne precjena podataka.

### 2.3. Umjetne neuronske mreže

Ljudi i životinje obrađuju informacije pomoću neuronskih mreža koje se sastoje od milijardi povezanih neurona (živčanih stanica) koji međusobno razmjenjuju kratke električne impulse, akcijske potencijale. Računalni algoritmi koji oponašaju te biološke strukture nazivaju se umjetne neuronske mreže [87]. Od samog početka razvoj ANN-a odvijao se u dva smjera. Prvi, ANN-om se nastojalo, što je bolje moguće, oponašati živčani sustav, a cilj tih istraživanja bio je usmjeren prema razumijevanju funkcioniranja živčanog sustava. Prva istraživanja započeli su prije 70-ak godina McCulloch i Pitts [88], te Hebb [89]. S razvojem računala primjena umjetnih neuronskih mreža počinje se usmjeravati na njihovu prilagodbu za rješavanje praktičnih računalnih problema. Cilj tih istraživanja bio je usmjeren k primjeni novog koncepta za rješavanje računskih problema, a ne samo razumijevanju neurofiziološkog funkcioniranja živčanog sustava. Prve jednostavne mreže razvili su Rosenblatt i Widrow [90]. Te mreže jednostavnih struktura bile su sposobne učiti i izvršavati samo jednostavne zadatke. Njihov uspjeh stimulirao je dodatna istraživanja na tom području pa se počelo vjerovati da su mogućnosti umjetnih neuronskih mreža ogromne. Sve to dovelo je do pojave nerealno visokih očekivanja u mogućnosti ANN-a. Međutim, rad Paperta i Minskyog [91] u kojem su pokazali da perceptroni imaju ozbiljna ograničenja u rješavanju nelinearnih problema, odvratio je mnoge znanstvenike od daljnjeg rada na tom području pa sva istraživanja vezana uz umjetne

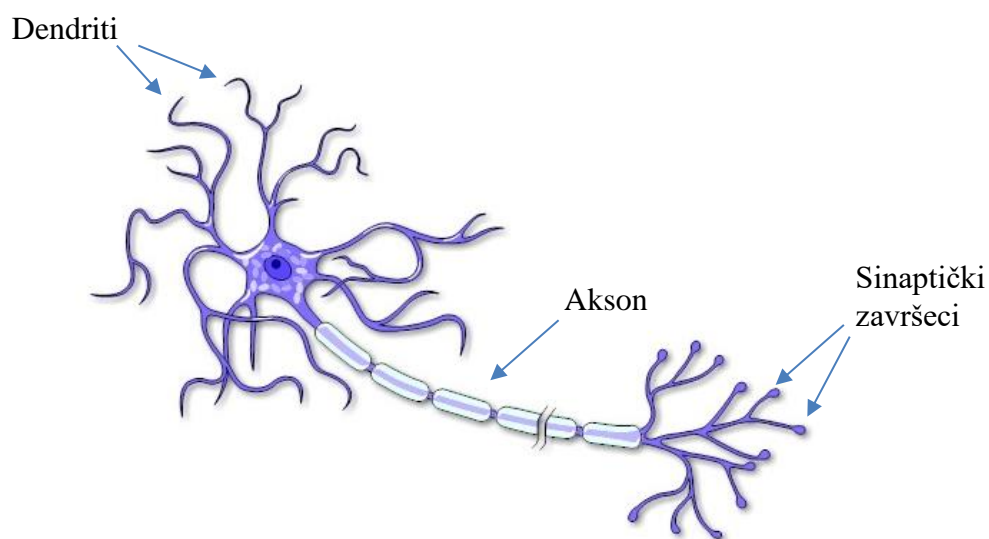
neuronske mreže gotovo da staju. Tek 1986. g. objavom knjige Rumelharta [92] o konceptu učenja povratnim rasprostiranjem pogreške (engl. *backpropagation error*), koji je ranije razvio Werbose [70], započet je ponovni interes za razvoj ANN-a. Novi koncept učenja omogućio je fleksibilnost staroj arhitekturi ANN-a i samim time mogućnost rješavanja kompleksnijih, nelinearnih problema.

Budući da se umjetne neuronske mreže sastoje od niza metoda, mogu se primjenjivati za rješavanje različitih problema. Ta vrsta algoritama danas se primjenjuje u računarstvu, strojarstvu, kemijskoj tehnologiji, znanosti o okolišu, medicini, nanotehnologiji, ekonomiji, kemiji, itd [70].

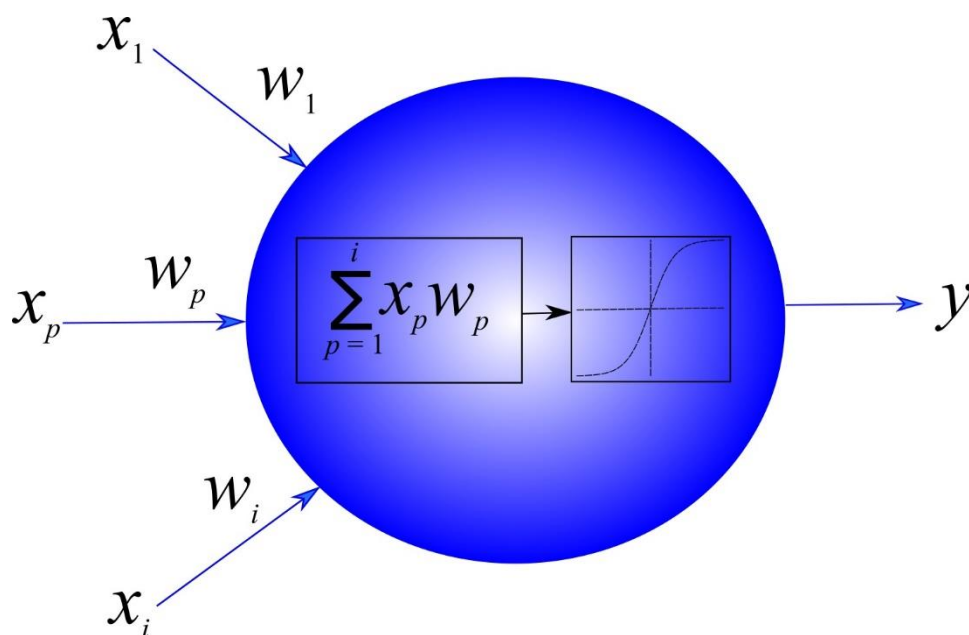
### 2.3.1. *Struktura neuronskih mreža*

Strukturalna jedinica neuronske mreže je neuron. Biološki neuroni (slika 13) primaju signal preko sinapsi koje se nalaze na dendritima ili membranama neurona. Kada primljeni signal ima dovoljno jaki intenzitet (prelazi određeni prag), neuron se aktivira i emitira signal kroz akson. Taj signal se potom može prenijeti do druge sinapse i aktivirati susjedne neurone. Rasprostiranje (propagacija) signala iz jedne živčane stanice u drugu ovisi o jačini sinapsi. U slučaju umjetnih neuronskih mreža, ulazni signali su realne ili binarne varijable,  $\mathbf{X} = (x_1, x_2, x_3, \dots, x_i)$ , koje se množe uteznim faktorima,  $\mathbf{W} = (w_1, w_2, w_3, \dots, w_i)$ , sumiraju i nakon toga prenose kroz aktivacijsku (prijenosnu) funkciju u izlazne varijable (signal). Propagacija signala i kod ANN-a ovisi o jačini sinapsi između neurona, odnosno o vrijednostima uteznih faktora (slika 14).

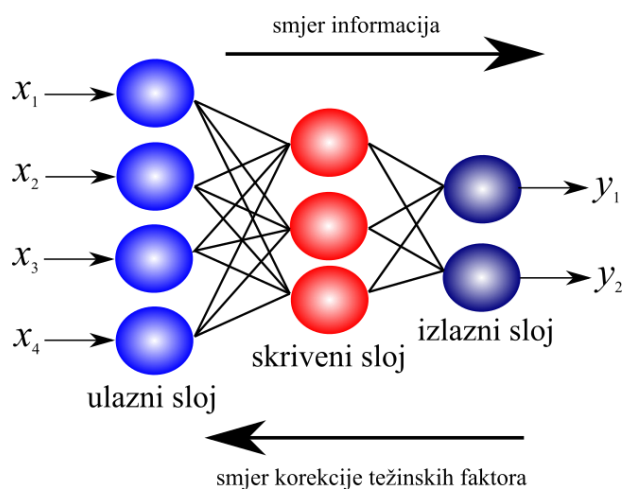




Slika 13. Biološki neuron.

Slika 14. Umjetni neuron. Vrijednosti  $x_1, x_2, \dots, x_i$  su ulazne varijable,  $w_1, w_2, \dots, w_i$ , utezni faktori, a  $y$  je izlazni signal.

Do danas je kreirano mnogo različitih struktura umjetnih mreža, ovisno o načinu kako su neuroni međusobno povezani i na koji način se signal prenosi među njima. Međutim, najčešće upotrebljavana struktura prikazana je na slici 15.



Slika 15. Struktura najčešće korištene umjetne neuronske mreže. Krugovi prikazuju neurone.

Kod te strukture, neuroni su poredani u tri sloja, ulazni, skriveni i izlazni. Svaki sloj se sastoji od barem jednog neurona (čvora) koji su na slici 15 prikazani kružićima. Svi neuroni u ulaznom sloju povezani su sa svim neuronima u skrivenom sloju. Isto tako svi neuroni iz skrivenog sloja povezani su sa svim neuronima iz izlaznog sloja. Kod ove strukture ANN-a, signal se prenosi samo s lijeva na desno, odnosno od ulaznog prema izlaznom sloju. Ulazni sloj je dio mreže koji nikada ne mijenja svoju vrijednost, te sadrži podatke (ulazne varijable) koji se uvode u mrežu. U skrivenom dijelu mreža može imati veći broj skrivenih slojeva, a u svakom od tih slojeva može biti bilo koji broj neurona. Odabir skrivenog broja slojeva, kao i broj neurona u njima, ovisi o problemu koji se pokušava riješiti primjenom algoritma. S većim brojem skrivenih slojeva i neurona raste broj izračuna koji je potrebno provesti što na kraju može biti računski vrlo zahtjevno i dugotrajno. Ukoliko se pak tijekom izgradnje strukture mreže odabere premali broj slojeva i neurona, algoritam će imati smanjenu mogućnost učenja, što naposljetku može imati za posljedicu pronalaženje rješenja koje nije optimalno.

Najčešće korištene prijenosne funkcije imaju jedan od sljedeća dva oblika (slika 16 i 17):

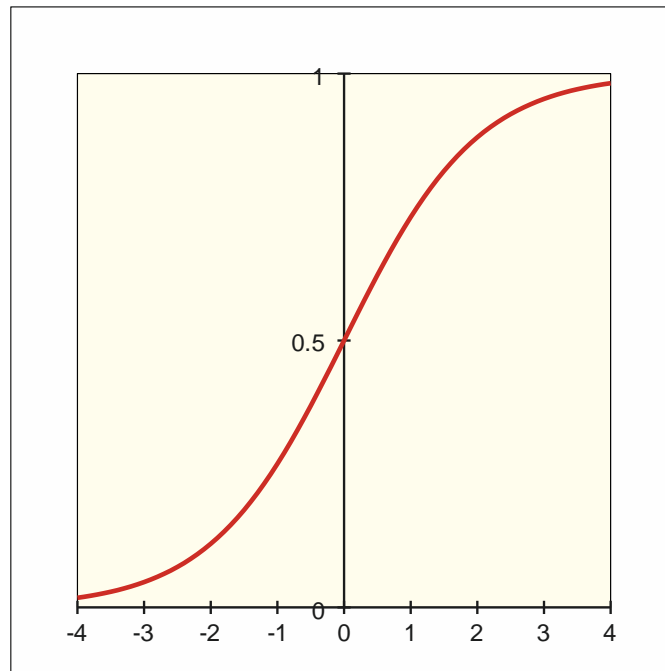
$$y_j = \frac{1}{1 + e^{-\alpha_j(\text{Net}_j - \theta_j)}} \quad (29)$$

ili

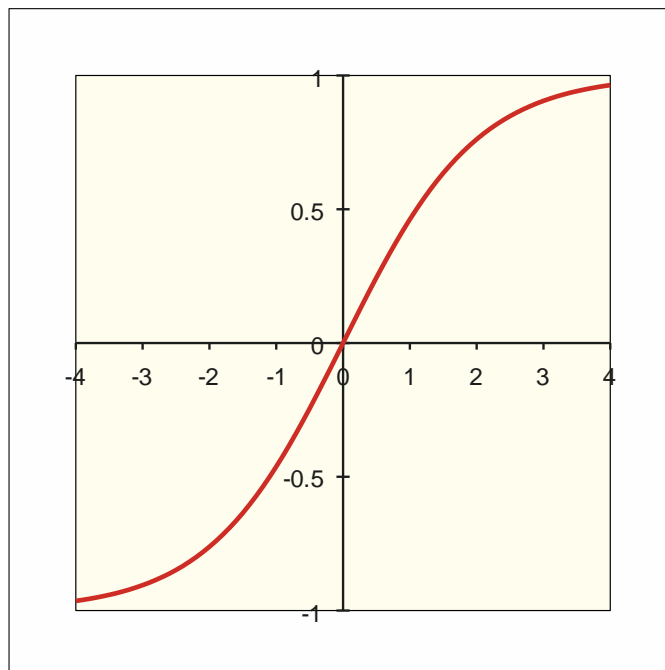
$$y_j = \frac{1 - e^{-\alpha_j(\text{Net}_j - \theta_j)}}{1 + e^{-\alpha_j(\text{Net}_j - \theta_j)}} \quad (30)$$

pri čemu je  $\text{Net}_j$  linearna kombinacija ulaznih varijabli za  $j$ -ti neuron:

$$\text{Net}_j = \sum_{p=1}^i w_{jp} x_p \quad (31)$$



Slika 16. Primjer sigmoidalne aktivacijske funkcije koja se koristi u skrivenom sloju umjetnih neuronskih mreža.



Slika 17. Hiperbolni tangens jedna od najčešće korištenih aktivacijskih funkcija u skrivenom sloju umjetnih neuronskih mreža.

Oblik odabrane prijenosne funkcije u skrivenom sloju mreže identičan je za sve neurone. Utežni faktori  $w_i$  određuju se iterativno, a taj postupak se naziva treniranjem mreže i predstavlja proces učenja. Parametri funkcije, vrijednost praga (engl. *threshold*)  $\theta_j$ , i nagib  $\alpha_j$ , mogu se također odrediti procesom učenja. Ako se kombiniraju jednadžba (31) i parametri funkcije dobije se sljedeći izraz:

$$\alpha_j(\text{Net}_j + \theta_j) = \sum_{p=1}^i \alpha_j w_{jp} x_p + \alpha_j \theta_j = \alpha_j w_{j1} x_1 + \alpha_j w_{j2} x_2 + \dots + \alpha_j w_{ji} x_i + \alpha_j \theta_j \quad (32)$$

u kojem se novi težinski faktori mogu prikazati kao umnožak dviju vrijednosti  $\alpha_j w_{jp}$ , a produkt  $\alpha_j \theta_j$  kao jedna varijabla. Ukoliko se umnožak  $\alpha_j \theta_j$  uključi u novi izraz  $\text{Net}_j$  koji sadrži ujedinjene sve parametre (težinske faktore, nagib i vrijednost praga) izraz (29) se može pojednostaviti:

$$y_j = \frac{1}{1 + e^{-\text{Net}_j}} \quad (33)$$

pri čemu je

$$\text{Net}_j = \sum_{p=1}^{i+1} w_{jp} x_p \quad (34)$$

Ovdje valja naglasiti da se uključivanjem parametara prijenosne funkcije u neuron broj utežnih faktora povećava za jedan. Povećanje broja parametara u modelu zahtjeva i veći broj objekata za trening.

### 2.3.2. Algoritam povratnog rasprostiranja pogreške

Jedna od najčešće korištenih metoda za učenje umjetnih neuronskih mreža je algoritam povratnog rasprostiranja pogreške, a naziv je dobila prema smjeru u kojem se korigiraju težinski faktori neurona tijekom procesa učenja (slika 15). U ovom slučaju težinski faktori se korigiraju od izlaznog prema ulaznom sloju, tj. u povratnom smjeru od rasprostiranja signala kada su objekti uvode u mrežu. Algoritam povratnog rasprostiranja pogreške koristi proces nadziranog načina učenja što znači da se aktualne izlazne vrijednosti,  $\mathbf{Y}$ , izračunane algoritmom uspoređuju s ciljanim vrijednostima  $\mathbf{R}$ . Zbog toga svaki  $i$ -dimenzionalni objekt  $\mathbf{X}$  mora biti povezan s  $n$ -dimenzijskim ciljanim vrijednostima  $\mathbf{R}$ . U slučaju primjene ANN-a za kvantitativnu analizu spektralnih podataka ulazne variable,  $\mathbf{X}$ , su spektri, a ciljane vrijednosti,

$R$ , su koncentracije analita. Promjena  $i$ -tog težinskog faktora  $j$ -tog neurona u  $l$ -tom sloju se vrši prema izrazu:

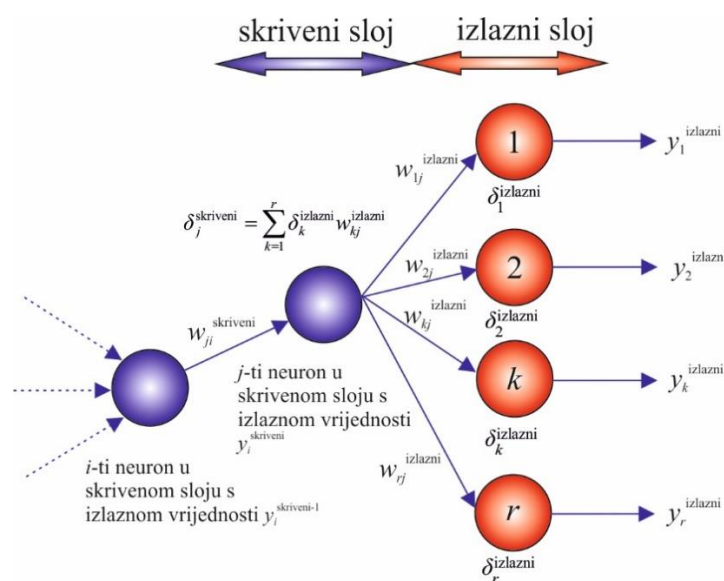
$$\Delta w_{ji}^l = \eta \delta_j^l \text{out}_j^{l-1} + \mu \Delta w_{ji}^{l, \text{prethodni}} \quad (35)$$

Prvi dio gornjeg izraza definira brzinu učenja pomoću vrijednosti konstante brzine  $\eta$  koja ima vrijednost između 0,01 i 1,0, a drugi dio izraza definira *momentum*,  $\mu$ . Vrijednost  $\Delta w_{ji}^{l, \text{prethodni}}$  je prethodna promjena težinskog koeficijenta. Član  $\delta_j^l$  je povezan s pogreškom  $j$ -tog neurona u  $l$ -tom sloju. Ulazni signal  $\text{out}_j^{l-1}$  koji uzrokuje pogrešku, dolazi do težinskog faktora  $w_{ji}$ , kao izlazna vrijednost neurona  $i$  prethodnog sloja ( $l-1$ ). Vrijednosti  $\delta_j^l$  se različito izračunavaju za neurone u izlaznom sloju u odnosu na one u skrivenom sloju kao što je prikazano sljedećim jednadžbama.

$$\delta_j^{\text{izlazni}} = (t_j - y_j^{\text{izlazni}}) y_j^{\text{izlazni}} (1 - y_j^{\text{izlazni}}) \quad \text{izlazni sloj} \quad (36)$$

$$\delta_j^{\text{skriveni}} = \left( \sum_{k=1}^{n_r} \delta_k^{\text{izlazni}} w_{kj}^{\text{izlazni}} \right) y_j^{\text{skriveni}} (1 - y_j^{\text{skriveni}}) \quad \text{skriveni sloj} \quad (37)$$

Suma u jednadžbi (37) uključuje sve neurone koji se nalaze lijevo od sloja za koji se izračunava vrijednost  $\delta_j$  (slika 18).



Slika 18. Proces učenja mreže algoritmom povratnog rasprostiranja pogreške.

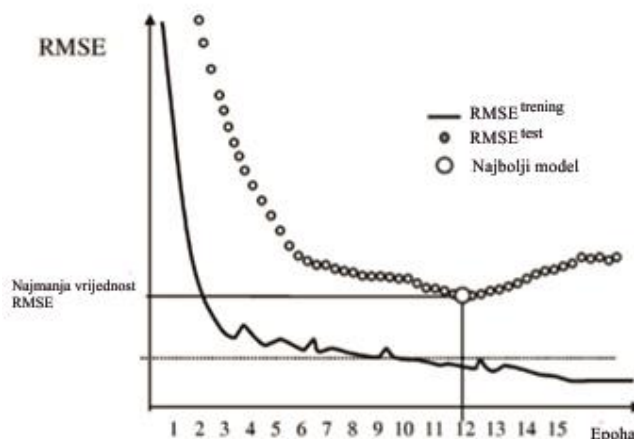
*Momentum* u jednadžbi (35) predstavlja memoriju mreže. Ovaj izraz osigurava da se vrijednosti težinskih faktora mijenjaju u istom smjeru tijekom svakog ciklusa optimizacije. Nadalje, vrijednost *momentuma* je važna kako bi se izbjeglo zaustavljanje algoritma u trenutku pronalaženja lokalnog minimuma površine odziva. Zbog ovog člana nastavlja se trend promjene vrijednosti težinskih faktora još neko vrijeme kako bi model „izašao“ iz pronađenog lokalnog minimuma. *Momentum* je uobičajeno konstanta vrijednost između 0,1 i 0,9, a u rijetkim slučajevima se mijenja tijekom procesa učenja.

### 2.3.3. Proces učenja

Učenje u slučaju umjetnih neuronskih mreža je proces optimizacije težinskih faktora mreže za specifični set podataka. U osnovi postoje dva procesa učenja neuronskih mreža, nadzirani (engl. *supervised*) i nenadzirani (engl. *unsupervised*). U ovom dijelu bit će opisan samo nadzirani način učenja budući da se ovaj tip učenja koristi kod kvantitativne obrade spektroskopskih podataka. U slučaju povratnog rasprostiranja pogreške proces učenja se kontrolira promjenom razlika ciljanih vrijednosti  $\mathbf{R}_s = (r_{s1}, r_{s2}, \dots, r_{sj}, \dots, r_{sn})$  i trenutnih izlaznih vrijednosti  $\mathbf{Y}_s = (y_{s1}, y_{s2}, \dots, y_{sj}, \dots, y_{sn})$  mreže. Kao mjera prikladnosti modela koristi se vrijednost korijena srednjeg kvadrata pogreške (engl. *Root Mean Square Error* - RMSE) koja se računa prema sljedećem izrazu:

$$\text{RMSE} = \sqrt{\sum_{s=1}^{n_t} \sum_{j=1}^n \frac{(r_{sj} - y_{sj})^2}{n_t n}} \quad (38)$$

Cilj učenja, odnosno treninga mreže je izračunati model koji ima najmanju moguću RMSE vrijednost. Ovdje valja naglasiti da najmanja RMSE vrijednost nije ona dobivena na setu podataka za kalibraciju, već ona dobivena na neovisnom setu koji nije korišten za trening. Osim eksperimentalne pogreške izmjerenih podataka na kvalitetu treninga mreže utječu i drugi parametri, kao što su dizajn mreže (broj skrivenih slojeva mreže, broj neurona u sloju, vrsta prijenosne funkcije, ...), odabir inicijalnih parametara, statistički reprezentativan odabir seta podataka za trening iz ukupne populacije, itd. Provjera modela neovisnim setom podataka može se raditi i tijekom samog procesa učenja, pa se npr. nakon određenog vremena mogu usporediti izračunane vrijednosti  $\text{RMSE}^{\text{trening}}$  i  $\text{RMSE}^{\text{test}}$ .

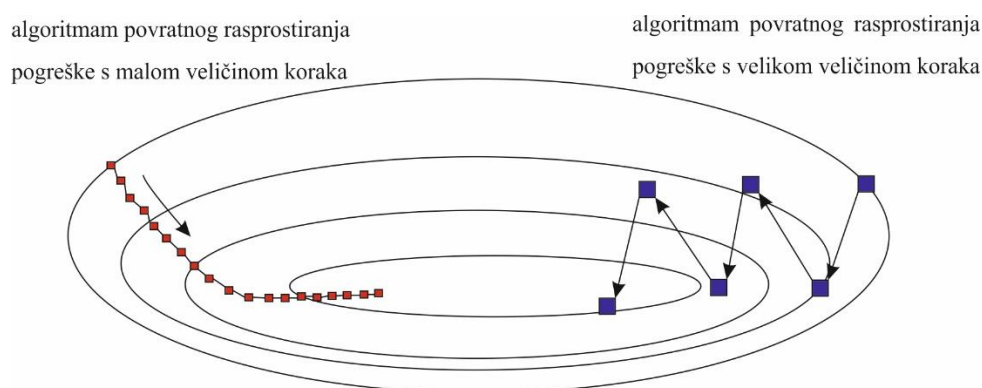


Slika 19. Promjena RMSE vrijednosti tijekom procesa učenja ANN-a.

U početku učenja određeni vremenski period obje vrijednosti RMSE-a se smanjuju, a potom kako se proces nastavlja,  $RMSE^{test}$  vrijednost može doseći svoju minimalnu vrijednost (slika 19). Nakon tog trenutka, svako smanjenje vrijednosti  $RMSE^{trening}$  zapravo ukazuje na pretreniranost (engl. *over-training*) mreže. Pretreniranost je posljedica činjenice da se model sastoji od prevelikog broja težinskih faktora koji se nakon određenog vremena počinju prilagođavati šumu. Zbog toga se težinski faktori počinju prilagođavati svim malim i nevažnim promjenama vrijednosti odziva  $R_s$  seta podataka za trening. U takvim slučajevima preporučljivo je pronaći takvu mrežu koja neće imati minimalnu  $RMSE^{test}$  vrijednost, već kod koje će se ta vrijednost stalno približavati određenom minimumu. Razlika između  $RMSE^{test}$  i  $RMSE^{trening}$  treba biti što je manja moguća, ali ipak vrijednost  $RMSE^{test}$  treba biti veća od  $RMSE^{trening}$  vrijednosti. Set podataka (test) koji se koristi za provjeru modela tijekom samog procesa treninga mreže, ne može se smatrati neovisnim jer se koristi za donošenje odluke o promjeni strukture mreže i/ili odabira novih parametara kao što su konstante brzine učenje i *momentuma*, odabir inicijalnih vrijednosti težinskih faktora, ... Zbog toga je kvalitetu odabranog modela potrebno provjeriti primjenom validacijskog seta podataka koji nije korišten u postupku treninga mreže. Tek ako se analizom tog seta dobije prihvatljiva vrijednost  $RMSE^{validacija}$  može se smatrati da je uspješno razvijen model neuronske mreže. U slučaju da ne postoji dovoljan broj podataka koji bi se podijelio u 3 grupe, grupu za trening, test i validaciju mreže, provjera modela može se provesti primjenom *leave-one-out* validacije. Kod te provjere iz cijelog seta podataka izostavlja se jedan objekt koji se ne koristi za trening mreže, već se upotrebljava za validaciju.

### 2.3.4. Levenberg-Marquardtov algoritam

Za trening malih i srednjih mreža upotrebljava se Levenberg-Marquardtov algoritam (LMA) kojeg su neovisno razvili Kenneth Levenberg [93] i Donald Marquardt [94], a u usporedbi s algoritmom povratnog rasprostiranja pogreške znatno je brži i postiže stabilnu konvergenciju. Dva su glavna razloga zašto algoritam povratnog rasprostiranja pogreške sporo konvergira: prvi je taj što veličina koraka mora biti adekvatna gradijentima. Kada je gradijent strm, veličina koraka mora biti mala kako bi se pronašao globalni minimum u prostoru rješenja (slika 20). Kad se odabere prevelika vrijednost brzine učenja (koraka) velika je vjerojatnost da se ne pronađe odgovarajući minimum. S druge strane, u slučaju kad je gradijent blag i vrijednost brzine učenja mala, trening mreže je vrlo spor. Drugi razlog je taj što zakrivljenost površine pogreške ne mora biti jednaka u svim smjerovima. Npr. u slučaju kad je na površini pogreške prisutna vrlo uska i duga dolina, komponenta gradijenta u smjeru duž dna doline je vrlo mala, dok je komponenta okomita na zidove doline vrlo velika. Zbog toga je kretanje duž dna doline sporo i može uzrokovati sporu konvergenciju.



Slika 20. Proces pretraživanja prostora rješenja s različitim konstantama brzine učenja. Crvenom bojom je prikazan proces pretraživanja s malom vrijednosti konstante brzine učenja koji vodi do spore konvergencije. Plavom bojom je prikazan proces pretraživanja s velikom vrijednosti konstante brzine učenja koji vodi do oscilacije.

Osnovni princip LMA metode je da primjenjuje kombinirani proces treninga mreže. U području površine sa složenom zakrivljenosti koristi se metoda najstrmijeg spusta dok se na dijelu površine koja se može aproksimirati kvadratnom funkcijom, koristi Gauss-Newtonov algoritam kojim se može značajno ubrzati konvergencija. U nastavku će biti opisani pojedini



dijelovi Levenberg-Marquardtovog algoritma, tj. algoritam najstrmijeg spusta, Newtonova metoda i Gauss-Newtonov algoritam [95].

### 2.3.5. Algoritam najstrmijeg spusta

Ovaj algoritam koristi prvu derivaciju funkciju ukupne pogreške (j. 39) kako bi pronašao minimum u prostoru pogreške:

$$E(\mathbf{x}, \mathbf{w}) = \frac{1}{2} \sum_{s=1}^S \sum_{j=1}^J (r_{sj} - y_{sj})^2 \quad (39)$$

pri čemu je  $\mathbf{x}$  vektor ulaznih varijabli,  $\mathbf{w}$  vektor težinskih faktora,  $r$  ciljana izlazna vrijednost i  $y$  trenutna izlazna vrijednost.

Gradijent,  $\mathbf{g}$ , dobiva se deriviranjem jednadžbe (39) po težinskim faktorima:

$$\mathbf{g} = \frac{\partial E(\mathbf{x}, \mathbf{w})}{\partial \mathbf{w}} = \left[ \frac{\partial E}{\partial w_1} \quad \frac{\partial E}{\partial w_2} \quad \dots \quad \frac{\partial E}{\partial w_n} \right]^T \quad (40)$$

Primjenom jednadžbe (40), u slučaju algoritma najstrmijeg spusta, promjena težinskih faktora se izračunava kao:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha \mathbf{g}_k \quad (41)$$

gdje je  $\alpha$  konstanta brzine učenja (veličina koraka). U području blizu minimuma funkcije vrijednosti svih elemenata vektora gradijenta su vrlo male pa je posljedično i promjena težinskih faktora mala.

### 2.3.6. Newtonova metoda

Newtonova metoda pretpostavlja da su sve komponente gradijenta funkcije težinskih faktora te da su svi faktori linearno neovisni (j. 42):

$$\begin{cases} g_1 = F_1(w_1, w_2 \dots w_n) \\ g_2 = F_2(w_1, w_2 \dots w_n) \\ \dots \\ g_n = F_n(w_1, w_2 \dots w_n) \end{cases} \quad (42)$$

gdje su  $F_1, F_2, \dots, F_n$  nelinearne veze između težinskih faktora i povezanih komponenta gradijenta.

Razvojem svake funkcije  $g_i$  ( $i = 1, 2, \dots, n$ ) iz jednadžbe (42) u Taylorov red i primjenom aproksimacije prvog reda dobije se izraz:

$$\left\{ \begin{array}{l} g_1 \approx g_{1,0} + \frac{\partial g_1}{\partial w_1} \Delta w_1 + \frac{\partial g_1}{\partial w_2} \Delta w_2 + \dots + \frac{\partial g_1}{\partial w_n} \Delta w_n \\ g_2 \approx g_{2,0} + \frac{\partial g_2}{\partial w_1} \Delta w_1 + \frac{\partial g_2}{\partial w_2} \Delta w_2 + \dots + \frac{\partial g_2}{\partial w_n} \Delta w_n \\ \dots \\ g_n \approx g_{n,0} + \frac{\partial g_n}{\partial w_1} \Delta w_1 + \frac{\partial g_n}{\partial w_2} \Delta w_2 + \dots + \frac{\partial g_n}{\partial w_n} \Delta w_n \end{array} \right. \quad (43)$$

Koristeći izraz (40), može se pokazati da vrijedi:

$$\frac{\partial g_i}{\partial w_i} = \frac{\partial \left( \frac{\partial E}{\partial w_k} \right)}{\partial w_i} = \frac{\partial^2 E}{\partial w_i \partial w_k} \quad (44)$$

Uvrštavanjem jednadžbe(44) u jednadžbu (43) dobiva se izraz:

$$\left\{ \begin{array}{l} g_1 \approx g_{1,0} + \frac{\partial^2 E}{\partial w_1^2} \Delta w_1 + \frac{\partial^2 E}{\partial w_1 \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_1 \partial w_n} \Delta w_n \\ g_2 \approx g_{2,0} + \frac{\partial^2 E}{\partial w_2 \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_2^2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_2 \partial w_n} \Delta w_n \\ \dots \\ g_n \approx g_{n,0} + \frac{\partial^2 E}{\partial w_n \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_n \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_n^2} \Delta w_n \end{array} \right. \quad (45)$$

Kod Newtonove metode za svaki element vektora gradijenta potrebno je izračunati drugu derivaciju ukupne funkcije pogreške dok se kod algoritma najstrmijeg spusta računa njena prva derivacija. Da bi se odredio minimum ukupne funkcije pogreške, svaki element gradijentnog vektora bi trebao biti nula. Zbog toga jednadžba (45) može se napisati kao:

$$\left\{ \begin{array}{l} 0 \approx g_{1,0} + \frac{\partial^2 E}{\partial w_1^2} \Delta w_1 + \frac{\partial^2 E}{\partial w_1 \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_1 \partial w_n} \Delta w_n \\ 0 \approx g_{2,0} + \frac{\partial^2 E}{\partial w_2 \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_2^2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_2 \partial w_n} \Delta w_n \\ \dots \\ 0 \approx g_{n,0} + \frac{\partial^2 E}{\partial w_n \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_n \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_n^2} \Delta w_n \end{array} \right. \quad (46)$$

a uvrštavanjem jednadžbe (40) u prethodni izraz, dobiva se:

$$\left\{ \begin{array}{l} -\frac{\partial E}{\partial w_1} = -g_{1,0} \approx \frac{\partial^2 E}{\partial w_1^2} \Delta w_1 + \frac{\partial^2 E}{\partial w_1 \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_1 \partial w_n} \Delta w_n \\ -\frac{\partial E}{\partial w_2} = -g_{2,0} \approx \frac{\partial^2 E}{\partial w_2 \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_2^2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_2 \partial w_n} \Delta w_n \\ \dots \\ -\frac{\partial E}{\partial w_n} = -g_{n,0} \approx \frac{\partial^2 E}{\partial w_n \partial w_1} \Delta w_1 + \frac{\partial^2 E}{\partial w_n \partial w_2} \Delta w_2 + \dots + \frac{\partial^2 E}{\partial w_n^2} \Delta w_n \end{array} \right. \quad (47)$$

koji se u matičnom obliku može napisati kao:

$$\begin{bmatrix} -g_1 \\ -g_2 \\ \dots \\ -g_n \end{bmatrix} = \begin{bmatrix} -\frac{\partial E}{\partial w_1} \\ -\frac{\partial E}{\partial w_2} \\ \dots \\ -\frac{\partial E}{\partial w_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_1 \partial w_n} \\ \frac{\partial^2 E}{\partial w_2 \partial w_1} & \frac{\partial^2 E}{\partial w_2^2} & \dots & \frac{\partial^2 E}{\partial w_2 \partial w_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 E}{\partial w_n \partial w_1} & \frac{\partial^2 E}{\partial w_n \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_n^2} \end{bmatrix} \times \begin{bmatrix} \Delta w_1 \\ \Delta w_2 \\ \dots \\ \Delta w_n \end{bmatrix} \quad (48)$$

Jednadžba (48) može se napisati i u skraćenom obliku:

$$-\mathbf{g} = \mathbf{H} \Delta \mathbf{w} \quad (49)$$

gdje se kvadratna matrica  $\mathbf{H}$  naziva Hesse-ova matrica i definirana je kao:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \frac{\partial^2 E}{\partial w_1 \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_1 \partial w_n} \\ \frac{\partial^2 E}{\partial w_2 \partial w_1} & \frac{\partial^2 E}{\partial w_2^2} & \dots & \frac{\partial^2 E}{\partial w_2 \partial w_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 E}{\partial w_n \partial w_1} & \frac{\partial^2 E}{\partial w_n \partial w_2} & \dots & \frac{\partial^2 E}{\partial w_n^2} \end{bmatrix} \quad (50)$$

Preuređenjem izraza (49) dobiva se sljedeća jednadžba:

$$\Delta \mathbf{w} = -\mathbf{H}^{-1} \mathbf{g} \quad (51)$$

Promjena vrijednosti težinskih faktora primjenom Newtonove metode izračunava se prema izrazu:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mathbf{H}_k^{-1} \mathbf{g}_k \quad (52)$$

Kod ove metode promjenu vektora gradijenta određuje Hesse-ova matrica na temelju vrijednosti druge derivacije funkcije ukupne pogreške.

### 2.3.7. Gauss-Newtonov i Levenberg-Marquardtov algoritam

Primjenom Newtonove metode za optimizaciju težinskih faktora potrebno je odrediti Hesse-ovu matricu izračunom druge derivacije funkcije ukupne pogreške. Ovaj izračun može biti složen pa je zbog pojednostavljenja uvedena Jacobijeva matrica,  $\mathbf{J}$  (j. 53).

$$\mathbf{J} = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_1} & \frac{\partial e_{1,1}}{\partial w_2} & \dots & \frac{\partial e_{1,1}}{\partial w_n} \\ \frac{\partial e_{1,2}}{\partial w_1} & \frac{\partial e_{1,2}}{\partial w_2} & \dots & \frac{\partial e_{1,2}}{\partial w_n} \\ \frac{\partial e_{1,1}}{\partial w_1} & \frac{\partial e_{1,1}}{\partial w_2} & \dots & \frac{\partial e_{1,1}}{\partial w_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{s,1}}{\partial w_1} & \frac{\partial e_{s,1}}{\partial w_2} & \dots & \frac{\partial e_{s,1}}{\partial w_n} \\ \frac{\partial e_{s,2}}{\partial w_1} & \frac{\partial e_{s,2}}{\partial w_2} & \dots & \frac{\partial e_{s,2}}{\partial w_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial e_{s,J}}{\partial w_1} & \frac{\partial e_{s,J}}{\partial w_2} & \dots & \frac{\partial e_{s,J}}{\partial w_n} \end{bmatrix} \quad (53)$$

pri čemu je:

$$e_{sj} = r_{sj} - y_{sj} \quad (54)$$

Uvrštavanjem jednadžbe (39) u jednadžbu (40) dobiva se sljedeći izraz:

$$g_i = \frac{\partial E}{\partial w_i} = \frac{\partial \left( \frac{1}{2} \sum_{s=1}^S \sum_{j=1}^J e_{sj}^2 \right)}{\partial w_i} = \sum_{s=1}^S \sum_{j=1}^J \left( \frac{\partial e_{sj}}{\partial w_i} e_{sj} \right) \quad (55)$$

Kombinirajući jednadžbe (55) i (53) dobiva se relacija između Jacobijeve matrice  $\mathbf{J}$  i vektora gradijenta  $\mathbf{g}$  (j. 56):

$$\mathbf{g} = \mathbf{J}\mathbf{e} \quad (56)$$

gdje je  $\mathbf{e}$  vektor pogreške definiran kao što slijedi:

$$\mathbf{e} = \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ \dots \\ e_{1,J} \\ \dots \\ e_{S,1} \\ e_{S,2} \\ \dots \\ e_{S,J} \end{bmatrix} \quad (57)$$

Ukoliko se izrazi (39) i (54) uvrste u Hesse-ovu matricu (j. 50), element matrice  $h_{i,j}$  može se izračunati prema sljedećem izrazu:

$$h_{i,j} = \frac{\partial^2 E}{\partial w_i \partial w_k} = \frac{\partial^2 \left( \frac{1}{2} \sum_{s=1}^S \sum_{j=1}^J e_{sj}^2 \right)}{\partial w_i \partial w_k} = \sum_{s=1}^S \sum_{j=1}^J \frac{\partial e_{sj}}{\partial w_i} \frac{\partial e_{sj}}{\partial w_k} + S_{ik} \quad (58)$$

gdje je  $S_{ik}$  jednak:

$$S_{ik} = \sum_{s=1}^S \sum_{j=1}^J \frac{\partial^2 e_{sj}}{\partial w_i \partial w_k} e_{sj} \quad (59)$$

Kod Gauss-Newtonove metode pretpostavlja se da je vrijednost  $S_{ik} \approx 0$  pa se korelacija između Hesseove i Jacobijeve matrice može napisati kao:

$$\mathbf{H} \approx \mathbf{J}^T \mathbf{J} \quad (60)$$

Kombinirajući jednadžbe (52), (56) i (60), promjena težinskih faktora Gauss-Newtonovim algoritmom izračunava se pomoću sljedećeg izraza:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \left( \mathbf{J}_k^T \mathbf{J}_k \right)^{-1} \mathbf{J}_k \mathbf{e}_k \quad (61)$$

Uvođenjem Jacobijeve matrice nije potrebno računati drugu derivaciju funkcije ukupne pogreške, što je prednost Gauss-Newtonovog algoritma u odnosu na Newtonovu metodu. Međutim, kada je prostor rješenja kompleksan, kod obje metode postoji problem konvergencije. Matematički taj se problem može obrazložiti činjenicom da matrica  $\mathbf{J}^T \mathbf{J}$  može biti singularna.

Kako bi se osiguralo da matrica  $\mathbf{J}^T \mathbf{J}$  bude regularna, Levenberg-Marquardtovim algoritmom uvodi se sljedeća aproksimacija Hesse-ove matrice:

$$H \approx J^T J + \mu I \quad (62)$$

gdje je  $\mu$  kombinacijski koeficijent (uvijek je pozitivan), a  $I$  je matrica identiteta.

Iz izraza (62) vidljivo je da su dijagonalni elementi Hesse-ove matrice uvijek veći od nule pa je samim time i matrica regularna.

Promjena težinskih faktora Levenberg-Marquardtovim algoritmom izračunava se prema izrazu:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \left( \mathbf{J}_k^T \mathbf{J}_k + \mu \mathbf{I} \right)^{-1} \mathbf{J}_k \mathbf{e}_k \quad (63)$$

Budući da navedeni algoritam kombinira metodu najstrmijeg spusta i Gauss-Newtonovu metodu, tijekom procesa treninga mreže ove dvije metode se zamjenjuju. U slučaju kad je vrijednost  $\mu$  vrlo mala (blizu nule), jednadžba (63) se približava izrazu (61) i koristi se Gauss-Newtonov algoritam. U slučaju kada je vrijednost kombinacijskog koeficijenta velika, jednadžba (63) aproksimira izraz (41) i koristi se metoda najstrmijeg spusta. U slučaju kad je kombinacijski koeficijent iz jednadžbe (63) vrlo velik, može se interpretirati kao konstanta brzine učenja algoritma najstrmijeg spusta (j. 64).

$$\alpha = \frac{1}{\mu} \quad (64)$$

Levenberg-Marquardtov algoritam smatra se jednim od najučinkovitijih algoritama za trening srednjih i malih umjetnih neuronskih mreža. Zbog svoje iznimne učinkovitosti preporuča se njegovo korištenje za proces učenja umjetnih neuronskih mreža s do nekoliko stotina težinskih faktora [96].

## § 3. EKSPERIMENTALNI DIO

### 3.1. Materijali

Paracetamol + Tramadol 325 mg + 37,5 mg tablete proizvedene su u pilot pogonu Istraživanja i razvoja u Belupu d.d. U tabletama je maseni udio paracetamola mijenjan u području od 65 do 88 %, a tramadola od 7,5 do 10,1 % i to na način da ne postoji korelacija među količinama (tablica 3). Ostatak formulacije koje sačinjavaju pomoćne tvari nije mijenjan tijekom pripreme tableta.

Tablica 3. Sastav proizvedenih Paracetamol + Tramadol 325 mg + 37,5 mg tableta.

Tableta	Paracetamol		Tramadol	
	w / %	% deklarirane količine	w / %	% deklarirane količine
1	65,0	85	7,9	90
2	68,8	90	10,1	115
3	72,6	95	9,7	110
4	76,5	100	8,8	100
5	80,3	105	7,5	85
6	84,1	110	8,4	95
7	87,9	115	9,3	105

Sve sirovine korištene za proizvodnju tableta bile su Ph.Eur. kvalitete. Tramadol je kupljen od farmaceutske kompanije Chemagis Ltd. (Izrael), a paracetamol od kompanije Mallinckrodt Inc. (SAD). Obje djelatne tvari standardizirane su prema odgovarajućim analitičkim postupcima opisanim u Ph.Eur. monografijama [97] i nakon toga upotrijebljene kao radni standardi.

Za spektrofotometrijsku i kromatografsku analizu korištena su sljedeća otapala: 96 % etanol (Merck, *Reag. Ph Eur*), 85 % fosforna kiselina (Merck, *Reag. Ph Eur*), 37 % klorovodična kiselina (Merck, *Reag. Ph Eur*), acetonitril (Merck, *gradient grade for liquid chromatography, Reag. Ph Eur*), trietilamin (Sigma Aldrich, *for HPLC ≥ 99,5 %*) i deionizirana voda.

Otopina pufera ( $c = 0,025 \text{ mol L}^{-1}$ , pH = 2,5) upotrijebljena za pripremu pokretne faze pripremljena je korištenjem soli kalijevog dihidrogenfosfata (Merck, *Reag. Ph Eur*).

Otopine tableta za spektrofotometrijska mjerenja filtrirane su kroz membranski filter veličine pora 0,45  $\mu\text{m}$  i membrane napravljene od regenerirane celuloze (RC), (Whatman, UK).

Za UHPLC mjerenja otopine standarda i tableta filtrirane su kroz 0,2  $\mu\text{m}$  RC filter (Whatman, UK).

Za kromatografsku analizu kao nepokretna faza korištena je kolona Zorbax SB C18 (Agilent, SAD), dimenzija 50  $\times$  2,1 mm, veličine čestica punila 1,8  $\mu\text{m}$ .

## 3.2. Metode analize

### 3.2.1. Spektrofotometrija

Osnovne otopine standarda pripremljene su otapanjem 50 mg radnog standarda paracetamola i 20 mg radnog standarda tramadola u 50 mL, odnosno 200 mL smjese 96 %-tnog etanola i 0,1 mol L<sup>-1</sup> klorovodične kisline (1:1, volumni omjer). Združene otopine standarda pripremljene su razrjeđivanjem osnovnih otopina u 20 mL prethodno navedene smjese otapala. Koncentracija paracetamola u pripremljenim otopinama bila je od 15 do 37  $\mu\text{g mL}^{-1}$ , a tramadola od 1,7 do 4,3  $\mu\text{g mL}^{-1}$ . Pripremljena su dva niza otopina, jedan za kalibraciju koji se sastojao od 25 otopina i drugi za validaciju koji se sastojao od 19 otopina.

Otopina uzorka tablete pripremljena je tako što je prenesena jedna tableta u odmjernu tikvicu od 100 mL i potom je dodan 1 mL vode da se tableta raspadne. Tikvica je dopunjena do oznake smjesom 96 %-tni etanol / 0,1 mol L<sup>-1</sup> HCl (1:1, volumni omjer) i potom stavljena u ultrazvučnu kupelj da se djelatne tvari otope. Otopina je filtrirana kroz 0,45  $\mu\text{m}$  RC membranski filter nakon čega je filtrat za spektrofotometrijska mjerenja razrijeđen 1 na 100 mL istom smjesom otapala, a za kromatografsku analizu razrijeđen je 1 na 20 mL smjesom 96 %-tni etanol / 0,1 mol L<sup>-1</sup> HCl (1:4, volumni omjer). Analizirano je 10 tableta u kojima je maseni udio paracetamola bio 76,5 %, a tramadola 8,8 %.

Svi UV spektri snimljeni su pomoću dvoznačnog spektrofotometra Cary 100 tvrtke Varian u kivetama duljine optičkog puta 1 cm uz 96 %-tni etanol i 0,1 mol L<sup>-1</sup> HCl (1:1, volumni omjer) kao slijepu probu. Spektri su snimljeni u području od 200 do 320 nm s korakom od 1 nm, uz integracijsko vrijeme 0,5 s.

### 3.2.2. NIR spektroskopija

NIR spektri cijelih tableta snimljeni su uz pomoć FT-NIR spektrometra MPA Bruker opremljenog InGaAs detektorom i vanjskim transmisijskim modulom. Mjerenja su provedena



u području od 12490 do 5793  $\text{cm}^{-1}$  u transmisijskom načinu snimanja s korakom od 8  $\text{cm}^{-1}$ , a spektar svake tablete je izračunan kao prosjek 128 mjerenja. Temperatura pri kojoj je provedena analiza bila je kontrolirana u području od 21 do 23 °C. Ukupno je analizirano 105 tableta od kojih je 70, po 10 tableta za svaki koncentracijski nivo, upotrijebljeno za izradu kalibracijskih modela, dok je za validaciju razvijenih modela korišteno 35 tableta (po 5 tableta za svaki koncentracijski nivo). Nakon provedene NIR analize tablete su analizirane primjenom metode tekućinske kromatografije ultravisoke djelotvornosti.

### 3.2.3. Ramanova spektroskopija

Ramanovi spektri cijelih tableta snimljeni su pomoću spektrometra Bruker Equinox 55 opremljenog s Ramanovim modulom FRA 106-s. Za pobuđivanje korišten je Nd:YAG laser valne duljine 1064 nm i snage 400 mW. Za svaku tabletu mjerenja su provedena u tri položaja s obje strane. U svakom položaju snimljeno je 128 spektara s korakom od 4  $\text{cm}^{-1}$  u području od 3500 do 200  $\text{cm}^{-1}$  te je za svaku tabletu izračunan njen prosječni spektar. Kao i u slučaju analize NIR tehnikom, analizirano je 105 tableta, od kojih je 70 korišteno za kalibraciju, a 35 za validaciju. Nakon mjerenja tablete su analizirane kromatografskom metodom.

### 3.2.4. Tekućinska kromatografija ultravisoke djelotvornosti

Za kromatografsku analizu otopine standarda pripremljene su otapanjem odgovarajuće mase tramadola i paracetamola u smjesi 96 %-tnog etanola i klorovodične kiseline ( $c = 0,1 \text{ mol L}^{-1}$ ) (1:4, volumni omjer) kako bi se dobila koncentracija tramadola 18,7  $\mu\text{g mL}^{-1}$  i paracetamola 162,5  $\mu\text{g mL}^{-1}$ .

Otopine tableta pripremljene su kao što je opisano u poglavlju 3.2.1 za spektrofotometrijska mjerenja.

Otopine standarada i tableta su prije injektiranja u kromatografski sustav filtrirane kroz 0,2  $\mu\text{m}$  RC membranski filter.

UHPLC analiza provedena je pomoću Agilent 1290 Infinity LC sustava opremljenog kvaternom pumpom, *diode array* detektorom, pećnicom i automatskim injektorom. Analiza je provedena na koloni Zorbax SB C18 (50 × 2,1 mm, 1,8  $\mu\text{m}$ ) pri sljedećim uvjetima: izokratna brzina protoka: 0,4  $\text{mL min}^{-1}$ ; detekcija: 274 nm za tramadol i 244 nm za paracetamol; vrijeme analize: 4 min, volumen injektiranja: 3  $\mu\text{L}$ , temperatura kolone: 35 °C. Kao pokretna faza

korištena je smjesa acetonitrila i kalijevog dihidrogenfosfata ( $c = 0,025 \text{ mol L}^{-1}$ ,  $\text{pH} = 2,5$ ) (15:85, volumni omjer) koja je sadržavala 3,5 mL trietilamina.

Metoda je validirana prema trenutno važećim ICH smjernicama o validaciji analitičkih postupaka [63].

### 3.2.5. Obrada podataka

Prikupljanje i obrada kromatografskih podataka, kao i svi izračuni, provedeni su pomoću kromatografskog softvera Empower Enterprise (Waters, SAD).

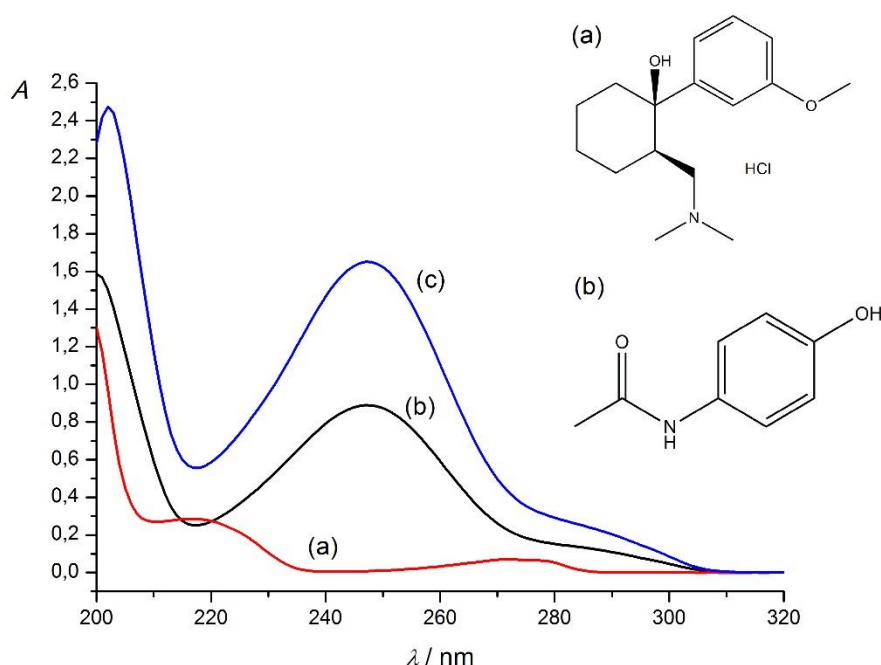
Prilikom obrade podataka UV, NIR i Ramanove spektroskopije primjenom PLS i GA-PLS algoritama korišten je program PLS Toolbox Solo (verzija 7,1, demo) dok je za obradu podataka primjenom umjetnih neuronskih mreža korišten program Matlab (verzija 2015b, trial). Statistička usporedba podataka dobivenih različitim metodama provedena je programom OriginPro (verzija 7.5), koji je korišten i pri grafičkom prikazivanju podataka.

## § 4. REZULTATI I RASPRAVA

### 4.1. Spektrofotometrijska analiza

UV spektri otopine standarda paracetamola, tramadola i smjese otopina tih dviju djelatnih tvari prikazani su na slici 21. Vidljivo je značajan preklapanje spektara tih dvaju lijekova. Dodatno, apsorbancije spektra tramadola značajno su niže od odgovarajućih apsorbancija paracetamola pa zbog toga spektrom združenih otopina značajno dominira spektar paracetamola.

Sastav binarnih smjesa otopina standarda izabran je nasumično kako bi se izbjegla korelacija između koncentracija dviju djelatnih tvari. Od ukupnog broja pripremljenih otopina, njih 25 je nasumično odabrano i korišteno za kalibraciju, dok je njih 19 upotrijebljeno za validaciju razvijenih kemometričkih modela. Sastav tih otopina prikazan je u tablicama 4 i 5. Kalibracijski modeli su razvijeni pomoću metode djelomičnih najmanjih kvadrata [66,67], genetičkih algoritama spregnutih s PLS-om [83–85] i umjetnih neuronskih mreža [26,28] kako bi se istovremeno odredile koncentracije tramadola i paracetamola u pripremljenim otopinama.



Slika 21. UV spektri (a) tramadola ( $\gamma = 10 \mu\text{g mL}^{-1}$ ), (b) paracetamola ( $\gamma = 10 \mu\text{g mL}^{-1}$ ) i (c) smjese tramadola ( $\gamma = 2 \mu\text{g mL}^{-1}$ ) i paracetamola ( $\gamma = 10 \mu\text{g mL}^{-1}$ ) u smjesi 96 %-tni etanol / 0,1 mol L<sup>-1</sup> HCl (1:1, volumeni omjer),  $l = 1$  cm.

Validirani kalibracijski modeli potom su korišteni za izračun količina djelatnih tvari u otopinama tableta. Uzorci Paracetamol + Tramadol 325 mg + 37,5 mg tableta analizirani su primjenom spektrofotometrijske i referentne UHPLC metode. Dobiveni rezultati statistički su uspoređeni upotrebom jednosmjerne analize varijance (ANOVA) [65].

Tablica 4. Sastav otopina za kalibraciju UV spektrofotometrijskom metodom.

Otopina za kalibraciju	$\gamma / \mu\text{g mL}^{-1}$	
	Tramadol	Paracetamol
1	3,10	16,0
2	1,70	18,0
3	3,90	24,0
4	2,10	20,0
5	2,90	31,0
6	3,70	35,0
7	4,20	21,0
8	3,80	35,0
9	1,70	32,0
10	4,30	36,0
11	4,00	28,0
12	1,90	25,0
13	2,20	18,0
14	1,90	33,0
15	4,00	18,0
16	2,50	32,0
17	1,90	17,0
18	2,50	15,0
19	2,60	34,0
20	3,20	26,0
21	4,00	37,0
22	4,00	35,0
23	1,70	32,0
24	4,00	28,0
25	3,80	37,0

#### 4.1.1. Obrada podataka PLS metodom

Kao što je pokazano u Literaturnom pregledu, PLS metoda vrlo se često koristi za obradu spektroskopskih podataka s ciljem kvantitativne analize lijekova u višekomponentnim farmaceutskim proizvodima [20,26,98,99]. Tom metodom se iz cijelog spektra na temelju dekompozicije latentnih varijabli izrađuje kalibracijski model. Algoritam simultano uzima u obzir apsorbancije i koncentracije analita u otopinama za kalibraciju. Postoje dva različita algoritma za izradu modela. S primjenom PLS1 regresije za svaki analit zasebno izrađuje se kalibracijski model koristeći njegov vektor koncentracije dok se PLS2 regresijom

koncentracije svih analita istovremeno koriste za konstrukciju modela [66,67]. U ovom radu korišten je PLS1 algoritam za izradu kalibracijskih modela, a metoda *cross-validacije* izostavljanja jednog uzorka (engl. *leave-one-out*) za odabir optimalnog broja latentnih varijabli [69]. Predviđene koncentracije tramadola i paracetamola u svakom uzorku, a koji nije uključen u izradu modela, su uspoređene sa stvarnim koncentracijama u kalibracijskom setu otopina, te su izračunane vrijednosti sume kvadrata reziduala pogreške predviđanja i korijena srednjeg kvadrata pogreške *cross-validacije*. Da bi se našao optimalni model, pronađen je model s najmanjom RMSECV vrijednosti i s najmanjim brojem latentnih varijabli. Nakon dodatka svakog novog faktora, ponovno je izračunan RMSECV i varijabla je zadržana u modelu ukoliko se vrijednost poboljšala (smanjila) za najmanje 5 %. Postupak je ponovljen za svaku djelatnu tvar zasebno. Za paracetamol pronađeno je da je optimalni broj faktora 2, a za tramadol 6.

Devetnaest združenih otopina standarda pripremljenih za validaciju analizirano je primjenom razvijene metode i za svaku djelatnu tvar izračunan je *recovery* i pogreška predviđanja. Pogreška je izražena kao relativna standardna pogreška (engl. *Relative Standard Error - RSE*) predviđanja koncentracije (j. 65). Rezultati su prikazani u tablici 5.

$$\text{RSE} = \left( \frac{\sum_{i=1}^N (\mathcal{Y}_{(\text{predviđeno})i} - \mathcal{Y}_{(\text{dodano})i})^2}{\sum_{i=1}^N (\mathcal{Y}_{(\text{predviđeno})i})^2} \right) \quad (65)$$

Za tramadol dobivena je srednja vrijednost *recovery*-a 99,0 % i RSE 2,58 %, dok su te vrijednosti za paracetamol 99,5 % i 0,85 %. Na temelju dobivenih vrijednosti može se zaključiti da je dokazana dobra točnost i preciznost rezultata izračunanih primjenom kalibracijskog modela razvijenog pomoću PLS metode.

Tablica 5. Sastav otopina za validaciju i njihove vrijednosti *recovery*-a izračunane primjenom PLS, GA-PLS i ANN metoda.

Otopina za validaciju	$\gamma_{\text{dodano}} / \mu\text{g mL}^{-1}$		<i>Recovery</i> / % (PLS)		<i>Recovery</i> / % (GA-PLS)		<i>Recovery</i> / % (ANN)	
	Tramadol	Paracetamol	Tramadol	Paracetamol	Tramadol	Paracetamol	Tramadol	Paracetamol
1	2,50	17,0	99,2	98,0	100,9	97,0	92,9	94,1
2	3,30	30,0	99,7	99,5	99,4	99,3	99,7	98,3
3	3,00	36,0	98,3	99,9	97,0	99,7	97,8	100,6
4	2,40	37,0	101,5	99,5	99,5	99,6	100,8	99,7
5	3,10	20,0	100,8	99,1	101,9	98,8	98,8	102,5
6	4,10	33,0	99,8	101,2	101,9	100,8	101,4	101,5
7	2,70	36,0	105,1	100,2	102,0	100,5	103,9	100,8
8	3,20	25,0	95,8	99,5	99,2	99,5	94,2	102,8
9	1,70	35,0	105,2	99,4	100,6	99,5	112,5	100,0
10	4,00	28,0	96,9	99,9	100,6	99,4	99,0	100,7
11	2,60	16,0	96,5	99,1	99,3	98,2	96,9	96,3
12	2,30	34,0	95,9	99,3	97,1	99,5	99,5	100,3
13	2,30	16,0	98,6	100,1	98,4	99,6	95,5	100,5
14	3,10	19,0	98,6	98,8	97,4	99,0	96,1	100,5
15	1,80	15,0	94,9	100,4	99,0	100,0	99,8	102,0
16	2,70	29,0	96,2	100,2	96,8	99,9	101,7	99,4
17	3,50	31,0	101,0	100,1	100,0	99,9	94,9	100,0
18	3,90	20,0	99,3	97,1	99,5	97,1	97,4	98,1
19	2,00	23,0	97,9	98,8	96,9	98,7	95,8	96,0
Srednja vrijednost <i>recovery</i> -a			99,0	99,5	99,4	99,3	98,9	99,7
RSE / %			2,58	0,85	1,69	0,90	3,68	1,65

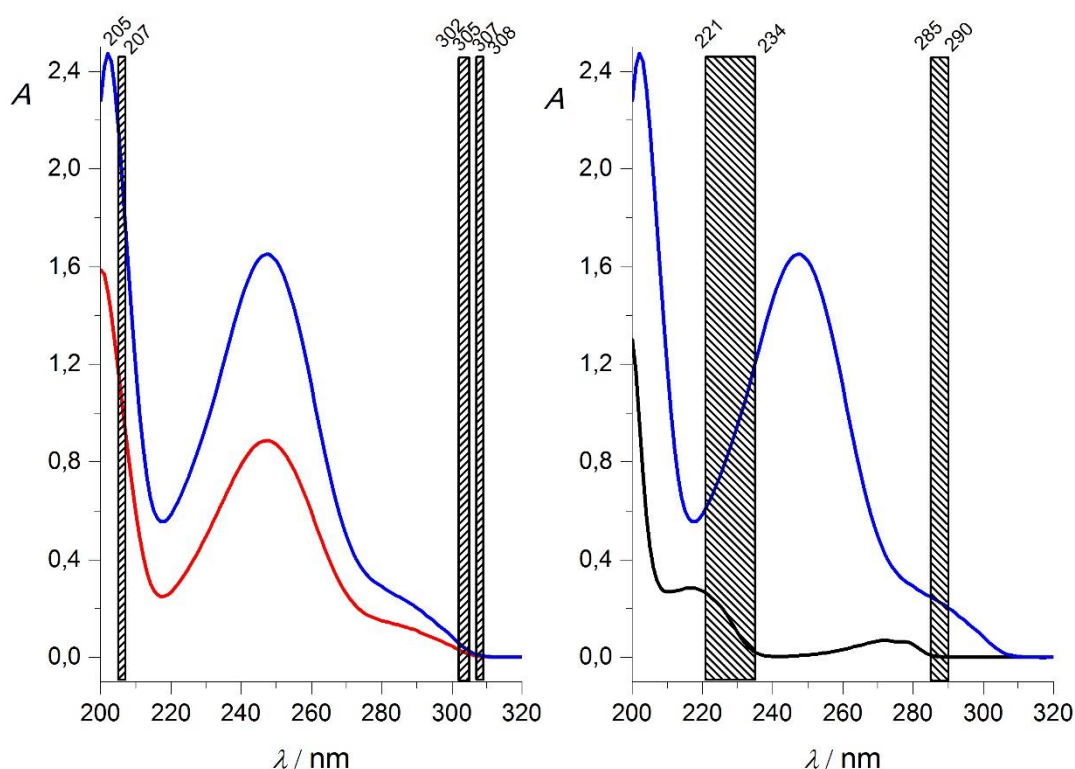
#### 4.1.2. Obrada podataka GA-PLS metodom

Kao što je već rečeno, genetički algoritam je metoda koja se koristi za rješavanje složenih optimizacijskih problema. U Literaturnom pregledu pokazano je da odgovarajuća selekcija optimalnih valnih duljina za konstrukciju kalibracijskog PLS modela može poboljšati njegovu prediktivnu sposobnost. Dodatno, GA metoda može pomoći u razumijevanju koja spektralna područja koreliraju s određenim fizikalno-kemijskim svojstvima sustava koji se istražuje, naročito u slučaju kad je teško ili gotovo nemoguće provesti vizualnu analizu spektara. Ovo je naročito izraženo u višekomponentnim sustavima sa značajnim spektralnim preklapanjem analita ili kada između njih postoji kemijska i fizikalna interakcija [32,86]. U analiziranom sustavu postoji značajno spektralno preklapanje UV spektara tramadola i paracetamola (slika 21) pa je zbog toga upotrijebljen GA algoritam spregnut s PLS metodom.

Podskup odabranih valnih duljina iz čitavog UV spektra (kromosom) prikazan je kao jedan genetički vektor. Svaki kromosom predstavlja moguće rješenje ili je kandidat za rješenje

problema. Pojedinačne vrijednosti kromosoma, tj. geni, su pojedinačne vrijednosti valne duljine, te su kodirani u binarni kod. PLS modeli generirani su korištenjem samo odabranih gena iz genetičkog vektora. Optimalni broj latentnih varijabli za PLS model odabran je *cross-validacijom*. Za svaki kromosom izračunana je PRESS vrijednost i članovi populacije su rangirani prema svojim vrijednostima prikladnosti. Najbolja rješenja su zadržana za daljnje razmnožavanje, dok su ostala rješenja odbačena. Nakon toga, najbolji kromosomi su rekombinirani i mutirani kako bi se dobila nova populacija potomaka. Roditeljska populacija je potom zamijenjena novim setom kreiranih kromosoma. Ponovno, za svaki kromosom izračunana je PRESS vrijednost i cijeli je postupak iterativno ponavljan sve dok nisu pronađeni kromosomi s najboljom vrijednosti prikladnosti. Kreiranje nove populacije kromosoma predstavlja jedan iterativni ciklus genetičkog algoritma i naziva se generacija. GA algoritam se može prekinuti nakon kreiranja kromosoma s predefiniranom vrijednosti prikladnosti ili nakon kreiranja određenog broja generacija kromosoma [28,31,65,83].

Budući da je GA stohastički algoritam, GA-PLS izračuni su ponovljeni 10 puta za 121 varijablu (200 – 320 nm) i maksimalno 10 latentnih varijabli je korišteno za izradu PLS modela. GA je prekinut nakon kreiranja 100 generacija ili kada je 50 % populacije bilo identično u točki konvergencije. Za tramadol odabrane su valne duljine 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 285, 286, 287, 288, 289 i 290 nm, a za paracetamol 205, 206, 207, 302, 303, 304, 305, 307 i 308 nm (slika 22). Za izradu kalibracijskog modela za određivanje tramadola, pronađeno je da su optimalne samo 2 latentne varijable, u usporedbi s njih 6 kada je korišten cijeli spektar. U slučaju paracetamola pronađeno je da je optimalni broj faktora određenih pomoću GA-PLS algoritma 3, dok je za PLS taj broj bio 2. Dobivena je izvrsna točnost i preciznost rezultata primjenom razvijenih modela (tablica 5).



Slika 22. Spektralna područja paracetamola (lijevo) i tramadola (desno) odabrana GA-PLS metodom. UV spektri smjese paracetamola i tramadola (a), paracetamola (b) i tramadola (c).

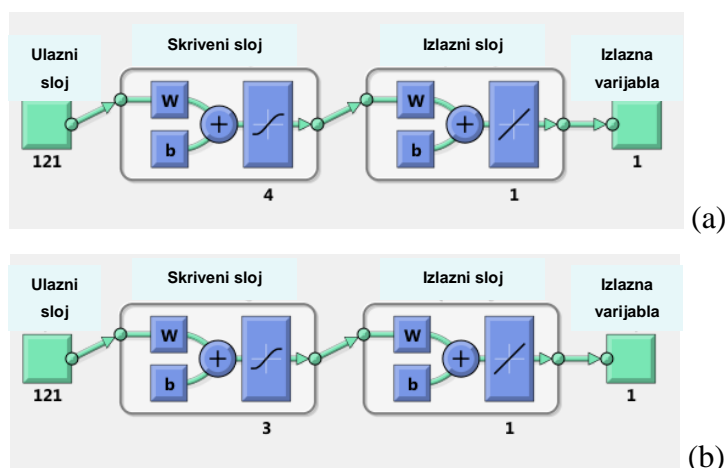
#### 4.1.3. Obrada podataka ANN metodom

ANN je metoda umjetne inteligencije koja odgovara biološkom živčanom sustavu koji ima sposobnost pronaći vezu između ulaznih i izlaznih varijabli. Metoda se najčešće koristi za obradu nelinearnih podataka, međutim postoje aplikacije ove metode i za obradu varijabli među kojim postoji linearna veza [26]. U usporedbi s linearnim metodama ANN algoritam je računski znatno kompleksniji i u većini slučajeva to je glavni nedostatak ove metode jer je potrebno mnogo vremena za njezin trening [27]. U ovom radu korišten je *feed-forward* model umjetne neuronske mreže koja je trenirana primjenom Levenberg-Marquardtovog algoritma [95]. Mreža se sastojala od 3 sloja, ulaznog, skrivenog i izlaznog. Pomoću ulaznog sloja u mrežu su uvedene izmjerene apsorbancije analiziranih otopina. U skrivenom sloju u slučaju tramadola korištena su 4 čvora (slika 23a), a u slučaju paracetamola njih 3 (slika 23b). Za izračun koncentracije oba analita u izlaznom sloju korišten je samo 1 čvor. Kao aktivacijska funkcija u skrivenom sloju korištena je sigmoidalna funkcija tangens hiperbolni (j. 66), dok je u izlaznom sloju upotrijebljena linearna prijenosa funkcija (j. 67).



$$y_j = \frac{2}{1 + e^{-2\alpha_j(\text{Net}_j - \theta_j)}} - 1 \quad (66)$$

$$y_i^{\text{izlazni}} = y_i \quad (67)$$



Slika 23. Shematski prikaz neuronskih mreža upotrijebljenih za izračun koncentracije tramadola (a) i paracetamola (b) u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

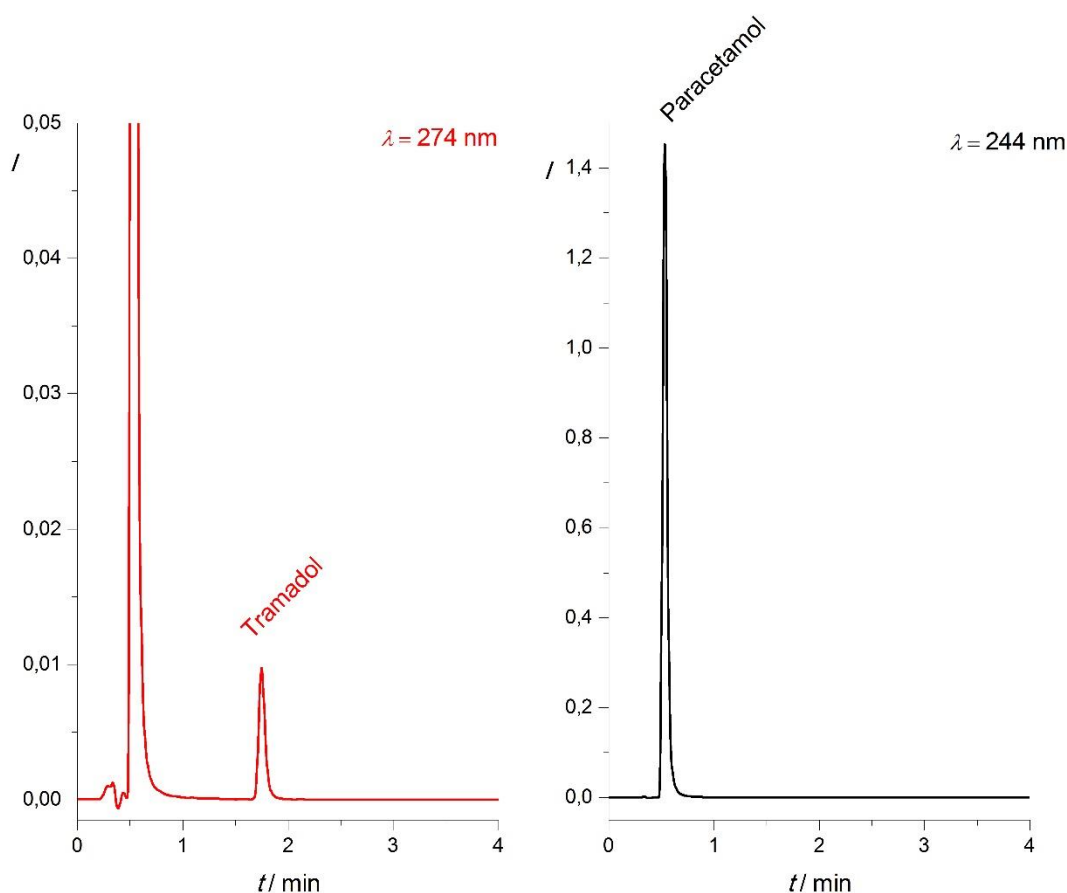
Za trening mreža upotrijebljeno je 25 standardnih otopina za kalibraciju čiji je sastav prikazan u tablici 4. Tijekom procesa treninga, 70 % nasumično odabranih otopina korišteno je za izradu modela, 15 % za njegovu validaciju i odluku o zaustavljanju treninga, te je ostalih 15 % uzoraka iz kalibracijskog skupa otopina upotrijebljeno za provjeru točnosti podataka izračunanih primjenom trenirane mreže. Točnost modela razvijenih primjenom ANN algoritma provjerena je analizom 19 neovisnih združenih otopina standarda tramadola i paracetamola čiji je sastav prikazan u tablici 5. Za svaki lijek izračunan je *recovery* i relativna standardna pogreška predviđanja koncentracije (j. 65). Za tramadol dobivena je srednja vrijednost *recovery*-a 98,9 %, za paracetamol 99,7 %, a izračunana relativna standardna pogreška predviđanja koncentracije tramadola je 3,68 %, a paracetamola 1,65 %. Na temelju izračunanih vrijednosti *recovery*-a može se zaključiti da je točnost svih triju korištenih metoda za obradu podataka, PLS, GA-PLS i ANN metode, za obje djelatne tvari vrlo dobra te se metode međusobno ne razlikuju. Vrijednost relativne standardne pogreške predviđanja koncentracije paracetamola ANN metodom nešto je viša u odnosu na pogreške predviđanja PLS i GA-PLS metoda, čije vrijednosti pogreške su gotovo identične. ANN metoda upotrijebljena za analizu tramadola ima najveću relativnu pogrešku predviđanja, 3,68 %, dok je relativna pogreška

predviđanja GA-PLS metode najmanja, 1,69 %, i gotovo identična relativnoj pogreški predviđanja paracetamola ANN metodom. Veće pogreške predviđanja koncentracije tramadola posljedica su činjenice da je u analiziranim otopinama masena koncentracija tramadola 8,7 puta manja u odnosu na onu paracetamola te postoji značajno preklapanje UV spektara dvaju analita. GA-PLS se u slučaju tramadola pokazao superiorniji u odnosu na druge dvije korištene multivarijatne kemometričke metode što je u skladu s literaturnim podacima [26,27,83]. S druge strane, ANN algoritmom za obje djelatne tvari dobivene su najveće RSE vrijednosti što je pokazatelj najmanje preciznosti rezultata izračunanih pomoću ove metode.

#### 4.1.4. Primjena razvijenih PLS, GA-PLS i ANN metoda

Kako bi se dodatno verificirala upotreba predloženih metoda, 10 pojedinačnih tableta je simultano analizirano primjenom novo-razvijene UV spektrofotometrijske metode i validirane UHPLC metode.

Količine oba lijeka u svakoj tableti određene su UHPLC-om uz standardne otopine tramadola i paracetamola. Tipični kromatogrami snimljeni tijekom UHPLC analize prikazani su na slici 24. Rezultati dobiveni primjenom trima multivarijatnim kemometričkim i UHPLC metodama prikazani su u tablici 6 i 7. UHPLC analizom dobivena je srednja vrijednost od 100,4 % tramadola od njegove deklarirane količine u tableti, a paracetamola 100,3 %. PLS-om je određeno 97,7 % tramadola i 100,3 % paracetamola, GA-PLS-om 100,7 % tramadola i 99,5 paracetamola. ANN algoritmom dobiveno je 99,5 % tramadola i 99,1 % paracetamola od deklarirane vrijednosti. Statistička usporedba podataka dobivenih novim i referentnom metodom provedena je pomoću jednosmjerne ANOVA-e (tablica 8) [65]. U slučaju paracetamola nema signifikantne razlike između metoda pri razini značajnosti 5 %, dok je u slučaju tramadola ta razlika značajna. Zbog toga su provedeni Tukeyev i Fischerov LSD test kako bi se utvrdilo koja od novo-razvijenih metoda, tj. PLS, GA-PLS ili ANN, u slučaju tramadola daje slične rezultate u usporedbi s kromatografskom analizom [65]. Prema rezultatima oba upotrijebljena testa, količina tramadola određena GA-PLS i ANN metodama za obradu UV spektrofotometrijskih podataka statistički se slaže s rezultatom dobivenim referentnom UHPLC metodom. Dobro slaganje količine lijekova određenih različitim tehnikama pokazuje da pomoćne tvari u formulaciji lijeka ne interferiraju tijekom UV spektrofotometrijske analize količina tramadola i paracetamola.



Slika 24. Tipični kromatogrami otopine uzorka Paracetamol + Tramadol 325 mg + 37,5 mg tableta. Detekcija tramadola je na 274 nm (lijevo), a paracetamola na 244 nm (desno).

Tablica 6. UHPLC i UV spektrofotometrijska analiza tramadola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

Tableta	Tramadol							
	UHPLC		PLS		GA-PLS		ANN	
	<i>m</i> / mg	%	<i>m</i> / mg	%	<i>m</i> / mg	%	<i>m</i> / mg	%
1	37,1	98,8	37,6	99,9	37,7	100,6	37,9	100,8
2	37,9	101,0	37,5	99,7	38,1	101,5	38,6	102,5
3	37,9	101,2	37,5	99,7	37,9	101,1	38,8	103,2
4	37,8	100,7	36,3	96,4	37,7	100,6	36,9	98,2
5	38,2	101,8	36,9	98,1	38,2	101,8	38,0	101,0
6	37,9	101,1	36,5	97,2	37,4	99,8	36,4	96,9
7	37,6	100,3	36,1	96,0	37,6	100,2	37,4	99,5
8	37,7	100,7	36,4	96,9	37,8	100,8	36,5	96,9
9	37,9	101,1	37,0	98,5	38,3	102,2	38,3	101,8
10	36,3	96,8	35,7	94,8	36,7	97,9	35,6	94,6
Srednja vrijednost		100,4		97,7		100,7		99,5
SD / %		1,47		1,75		1,20		2,84

Tablica 7. UHPLC i UV spektrofotometrijska analiza paracetamola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

Tableta	Paracetamol							
	UHPLC		PLS		GA-PLS		ANN	
	<i>m</i> / mg	%	<i>m</i> / mg	%	<i>m</i> / mg	%	mg / tbl.	%
1	324,7	99,9	325,9	100,3	322,4	99,2	323,4	99,5
2	328,4	101,1	328,3	101,0	327,5	100,8	327,5	100,7
3	324,3	99,8	322,5	99,2	320,5	98,6	318,4	98,0
4	330,1	101,6	329,8	101,5	326,8	100,6	322,2	99,1
5	328,9	101,2	327,5	100,8	324,9	100,0	323,4	99,5
6	324,0	99,7	323,8	99,6	321,5	98,9	319,7	98,4
7	322,4	99,2	323,5	99,5	320,2	98,5	318,4	98,0
8	329,2	101,3	329,6	101,4	325,6	100,2	324,7	99,9
9	328,8	101,2	329,4	101,3	327,5	100,8	325,1	100,0
10	317,5	97,7	319,7	98,4	316,4	97,3	318,4	98,0
Srednja vrijednost		100,3		100,3		99,5		99,1
SD / %		1,23		0,69		1,15		0,99

Tablica 8. Statistička usporedba količina tramadola i paracetamola određenih predloženim kemometričkim metodama i referentne UHPLC metode primjenom jednosmjerne ANOVA-e.

Izvor varijacije	DF <sup>a</sup>	Tramadol				Paracetamol			
		SS <sup>b</sup>	MS <sup>c</sup>	<i>F</i> <sub>o</sub>	<i>P</i>	SS <sup>b</sup>	MS <sup>c</sup>	<i>F</i> <sub>o</sub>	<i>P</i>
Između metoda	3	5,92	1,97	3,83	0,0177	110,55	36,85	2,83	0,0520
Unutar metoda	36	18,55	0,52			468,57	13,02		
Ukupno	39	24,48				579,12			

<sup>a</sup>DF – broj stupnjeva slobode

<sup>b</sup>SS – suma kvadrata

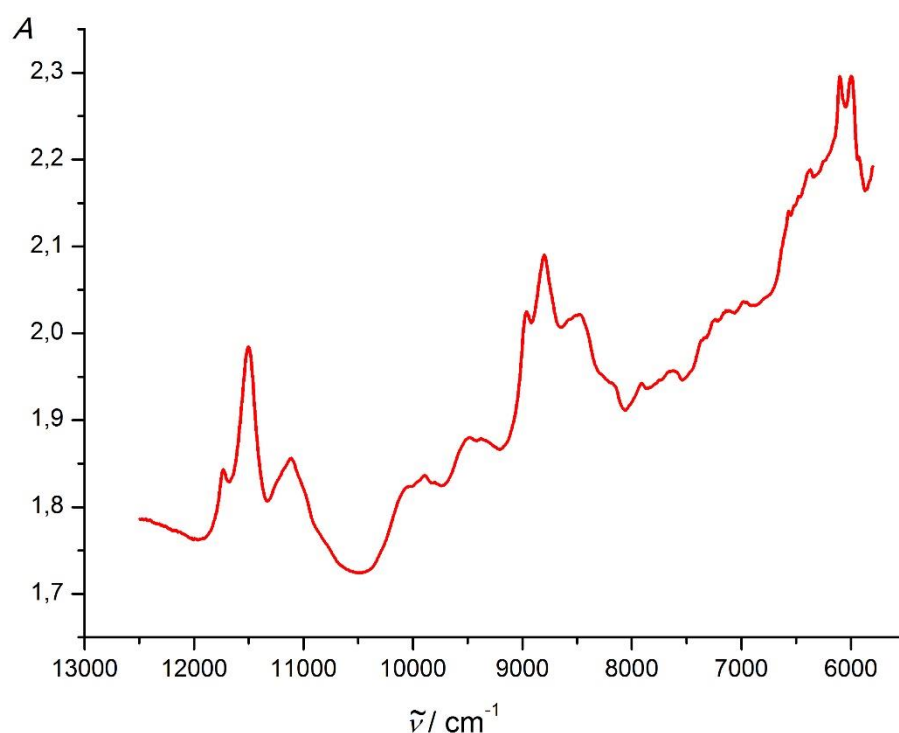
<sup>c</sup>MS – srednji kvadrat

## 4.2. Analiza NIR spektroskopijom

Određivanje količine analita NIR spektroskopijom uobičajeno se razlikuje od ostalih spektroskopskih tehnika, npr. IR ili UV-Vis spektroskopije. Za razliku od navedenih „klasičnih“ spektroskopskih metoda, uzorke analizirane NIR spektroskopijom potrebno je analizirati odgovarajućom referentnom metodom [100]. Dio analiziranih uzoraka koristi se za kalibraciju, a dio za provjeru razvijenih kemometričkih modela. Uzorci upotrijebljeni za izradu kalibracijskih modela trebaju pokrivati koncentracijska područja analita koja se očekuju tijekom analize budućih uzoraka. Snimljeni NIR spektri kalibracijskih uzoraka se potom kemometričkim metodama koreliraju s količinama analita određenih referentnom metodom [101]. U Literaturnom pregledu pokazano je da se NIR spektroskopija u sprezi s

kemometričkim izračunima koristi za kvantitativne analize kako djelatnih tako i pomoćnih tvari u različitim farmaceutskim oblicima.

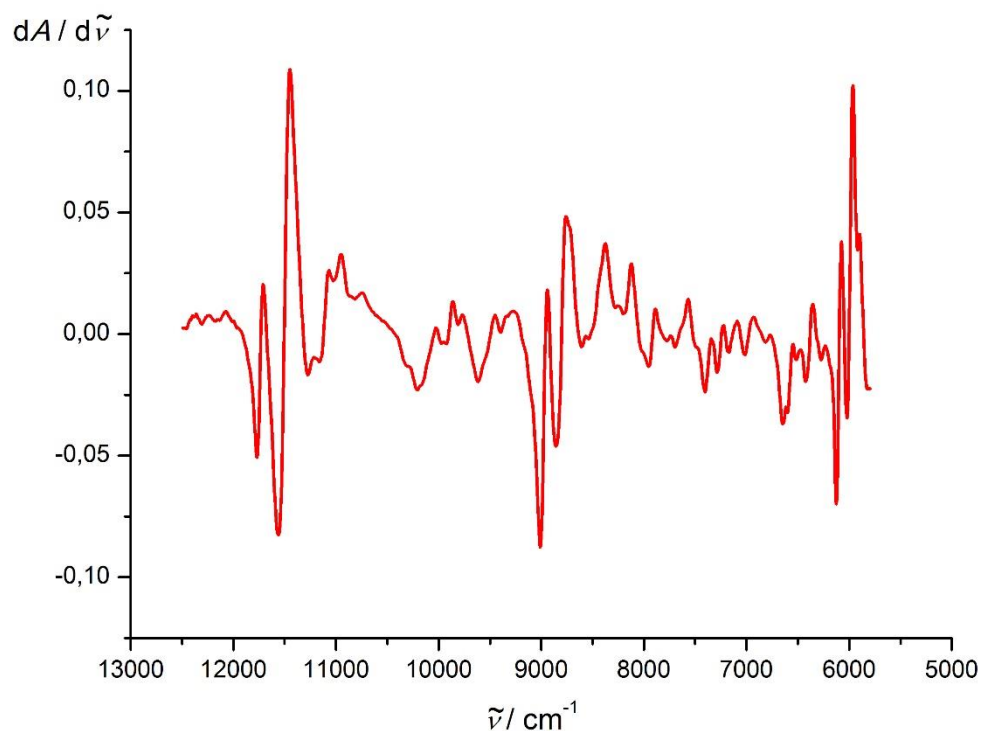
U ovom radu, za kalibraciju su proizvedene tablete u kojima je udio paracetamola mijenjan u području od 65 do 88 %, a tramadola od 7,5 do 10,1 % (tablica 3). Količine lijekova su mijenjane na takav način da je korelacija među njima bila minimalna. Za pripremu svih uzoraka korišten je isti proizvodni postupak. Način pripreve kalibracijskih uzoraka, kao i koncentracijsko područje tih uzoraka, odabrani su tako da bi pokrili maksimalni očekivani raspon količina djelatnih tvari u tabletama i varijabilnosti tijekom standardnog proizvodnog procesa. Proizvedene tablete analizirane su primjenom NIR spektroskopije (slika 25) i referentne kromatografske metode na način kao što je opisano u dijelu 3.2. Snimljeni NIR spektri potom su obrađeni primjenom PLS, GA-PLS i ANN multivarijantnih kemometričkih metoda s ciljem kvantitativne analize djelatnih tvari u tabletama. Način obrade spektroskopskih podataka, kao i dobiveni rezultati primjenom predloženih metoda, bit će prikazani u nastavku ovog rada.



Slika 25. Primjer NIR spektra Paracetamol + Tramadol 325 mg + 37,5 mg tablete snimljen transmisivskim načinom.

#### 4.2.1. Obrada podataka PLS metodom

Obrada NIR spektara provedena je kao i u slučaju obrade spektrofotometrijskih podataka, primjenom PLS algoritma kojim su istovremeno uzeti u obzir intenziteti NIR spektara snimljenih u transmisijskom načinu i koncentracije analita u tabletama određene primjenom UHPLC metode. Za svaku djelatnu tvar postupak izrade kalibracijskog modela proveden je zasebno te je i u ovom slučaju korištena PLS1 vrsta algoritma [66]. Ukupno je upotrijebljeno 70 tableta za kalibraciju, po 10 tableta za svaki koncentracijski nivo (tablica 3). Za predobradu izmjerenih spektara korišten je Savitzky–Golayev algoritam [102] pomoću kojega je izračunana prva derivacija, a potom je provedena njegova transformacija primjenom *standard normal variate* metode [103] u slučaju paracetamola (slika 26), a u slučaju tramadola korištena je *multiplicative scatter* korekcija [14]. Ove metode predobrade podataka upotrijebljene su kako bi se iz NIR spektara uklonile multiplikativne interferencije signala bazne linije, raspršenog zračenja, veličine čestica, tvrdoće tableta, itd. Odabir metoda za predobradu spektroskopskih podataka i optimalnog broja faktora kemometričkih modela izvršen je primjenom *cross-validacije* izostavljanja jednog uzorka [69]. Na temelju predviđenih koncentracija djelatnih tvari u svakom uzorku kalibracijskog niza tableta koji nije uključen u izradu modela i koncentracija koje su određene referentnom metodom izračunane su vrijednosti PRESS-a i RMSECV-a. Optimalnim modelom proglašen je onaj s najmanjom RMSECV vrijednosti i s najmanjim brojem latentnih varijabli. Vrijednost RMSECV (j. 10) izračunana je nakon dodatka svake nove latentne varijable, koja je zadržana u modelu ukoliko se vrijednost smanjila za najmanje 5 %. Ovakvim postupkom pronađeno je da je optimalni broj latentnih varijabli za paracetamol 5, a za tramadol 8.



Slika 26. Prva derivacija NIR spektra Paracetamol + Tramadol 325 mg + 37,5 mg tableta izračunana primjenom Savitzky–Golayevog algoritma. Spektar je normaliziran primjenom SNV metode.

Kako bi se provjerila točnost i preciznost rezultata dobivenih razvijenim kemometričkim modelima analizirano je 35 neovisnih validacijskih uzoraka, po 5 tableta za svaki koncentracijski nivo. Za procjenu kvalitete (moći predviđanja) predloženih metoda upotrijebljene su sljedeće statističke veličine: koeficijent korelacije,  $R$  [65], korijen srednjeg kvadrata pogreške predviđanja, RMSEP [101], *bias* [28], i devijacija reziduala predviđanja (engl. *Residual Prediction Deviation* - RPD) [104]. *Bias* je sustavna pogreška izmjerenih (izračunanih) vrijednosti u odnosu na one određene referentnom metodom te je izračunan pomoću sljedećeg izraza:

$$bias = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (68)$$

gdje je:  $y_i$  – količina djelatne tvari određena referentnom metodom za  $i$ -ti uzorak

$\hat{y}_i$  – količina djelatne tvari izračunana iz kalibracijskog modela za  $i$ -ti uzorak

$n$  – ukupan broj analiziranih uzoraka.

Devijacija reziduala predviđanja je omjer standardne devijacije i standardne pogreške predviđanja, a izračunana je primjenom sljedeće jednadžbe:

$$RPD = \frac{SD}{SEP} \quad (69)$$

pri čemu je: SD – standardna devijacija količina djelatnih tvari analiziranih UHPLC metodom

SEP – standardna pogreška predviđanja djelatne tvari (j. 70).

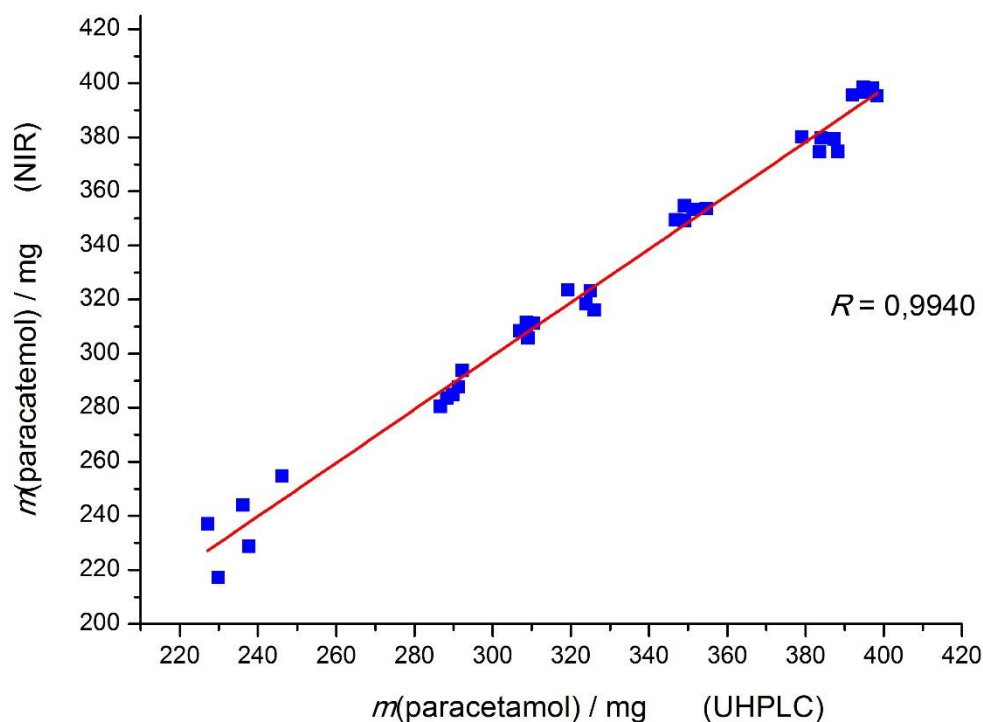
$$SEP = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i - bias)^2}{n-1}} \quad (70)$$

U tablici 9 prikazane su količine paracetamola određene u validacijskom nizu Paracetamol + Tramadol 325 mg + 37,5 mg tableta primjenom razvijene PLS i referentne UHPLC metode. Srednji kvadrat pogreške predviđanja razvijene NIR metode iznosi 5,86 mg, koeficijent korelacije 0,9940 (slika 27), devijacija reziduala predviđanja 9,13 i *bias* 1,17 mg.



Tablica 9. UHPLC i NIR analiza paracetamola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

Tableta	Paracetamol			
	UHPLC	PLS	GA-PLS	ANN
	<i>m</i> / mg	<i>m</i> / mg.	<i>m</i> / mg	<i>m</i> / mg
1	346,8	349,5	352,5	351,3
2	349,1	354,7	354,6	353,5
3	351,3	353,3	353,7	350,2
4	354,6	353,7	353,6	355,8
5	349,1	349,2	354,4	357,3
6	227,2	237,0	234,1	232,0
7	229,9	217,3	217,4	223,3
8	236,2	244,0	239,0	237,7
9	246,2	254,8	250,7	247,3
10	237,7	228,8	225,5	232,4
11	384,0	379,8	381,6	382,7
12	388,3	374,8	380,1	383,2
13	387,3	379,3	379,4	382,2
14	379,0	380,1	380,9	384,5
15	383,6	374,7	377,6	384,3
16	319,3	323,6	325,0	324,8
17	325,5	320,0	321,0	314,9
18	325,0	323,2	324,4	324,1
19	323,8	318,4	319,5	325,4
20	326,0	316,1	319,4	323,7
21	392,1	395,6	394,9	396,8
22	394,8	398,5	396,8	396,9
23	394,9	396,6	396,3	397,1
24	397,2	398,2	394,7	397,2
25	398,3	395,4	395,2	396,7
26	307,1	308,4	308,7	304,2
27	308,8	310,3	309,4	309,1
28	308,7	311,5	313,6	311,1
29	309,0	305,7	306,2	308,5
30	310,5	311,1	308,4	309,4
31	289,8	284,8	283,3	289,2
32	288,3	283,6	283,5	282,1
33	292,3	293,7	290,7	291,0
34	291,3	287,8	287,2	291,4
35	286,7	280,4	283,5	281,8
<i>R</i>		0,9940	0,9954	0,9956
RMSEP / mg		5,86	5,19	5,06
RPD		9,13	10,30	10,52
<i>bias</i> / mg		1,17	1,11	0,70

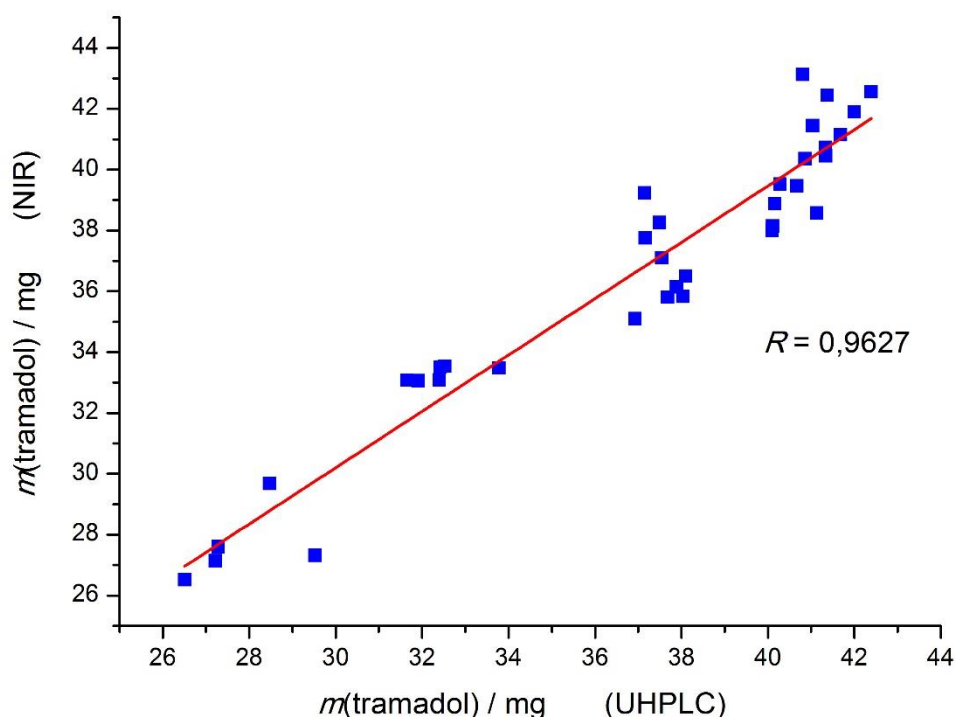


Slika 27. Količine paracetamola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene NIR metode. Za obradu podataka korištena je PLS metoda.

Količine tramadola izračunane primjenom kemometričkog modela razvijenog PLS algoritmom i one određene UHPLC-om prikazane su u tablici 10. Iz dobivenih podataka izračunana je  $R$  vrijednost 0,9627 (slika 28), RMSEP 1,33 mg, RPD 3,69 i *bias* 0,30 mg. Na temelju ovih vrijednosti može se zaključiti da je razvijena PLS metoda za analizu tramadola lošija u usporedbi s odgovarajućom metodom za određivanje paracetamola. Isto je zamijećeno i kod predložene UV spektrofotometrijske metode. Znatna razlika u koncentracijama djelatnih tvari u tabletama i u slučaju NIR spektroskopije ima za posljedicu da su količine tramadola određene pomoću modela kreiranog primjenom PLS-a znatno nepouzdanije od količina paracetamola izračunanih upotrebom iste kemometričke metode. Prema izračunanim vrijednostima  $R$ , RMSEP-a, RPD-a i *bias*-a NIR metoda spregnuta s PLS algoritmom za predviđanje količine paracetamola može se koristiti u rutinskoj kontroli kvalitete proučavanog lijeka. S druge strane, količina tramadola analizirana pomoću predložene NIR metode je nepouzdana te metoda nema potencijal za upotrebu u rutinskoj kontroli.

Tablica 10. UHPLC i NIR analiza tramadola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

Tableta	Tramadol			
	UHPLC	PLS	GA-PLS	ANN
	<i>m</i> / mg	<i>m</i> / mg	<i>m</i> / mg	<i>m</i> / mg
1	31,9	33,1	33,7	33,4
2	31,7	33,1	32,1	31,3
3	32,4	33,1	32,6	32,0
4	32,5	33,5	33,0	31,7
5	32,4	33,5	32,3	34,0
6	28,5	29,7	29,4	29,0
7	29,5	27,3	29,4	27,4
8	26,5	26,5	28,1	26,3
9	27,2	27,1	27,7	27,6
10	27,3	27,6	27,7	27,6
11	37,5	38,3	37,2	40,0
12	38,0	35,8	36,8	39,3
13	37,6	37,1	36,6	37,7
14	37,2	39,2	37,7	39,1
15	37,9	36,1	36,4	38,9
16	37,7	35,8	38,1	38,4
17	33,8	33,5	33,1	37,3
18	36,9	35,1	36,1	36,1
19	38,1	36,5	37,8	38,9
20	37,2	37,8	35,4	39,5
21	40,9	40,4	41,6	41,2
22	41,0	41,5	41,5	41,2
23	41,3	40,7	41,1	41,1
24	41,4	42,4	41,8	40,9
25	41,7	41,2	41,2	41,1
26	40,1	38,1	39,4	37,9
27	40,2	38,9	39,2	38,5
28	40,1	38,0	39,2	39,0
29	40,3	39,5	40,2	40,1
30	40,7	39,5	40,3	40,3
31	42,0	41,9	41,6	41,2
32	41,3	40,5	41,3	40,6
33	40,8	43,1	41,2	40,9
34	42,4	42,6	41,9	41,5
35	41,1	38,6	40,0	40,5
<i>R</i>		0,9627	0,9879	0,9657
RMSEP / mg		1,33	0,80	1,14
RPD		3,69	6,11	4,13
<i>bias</i> / mg		0,30	0,13	-0,04

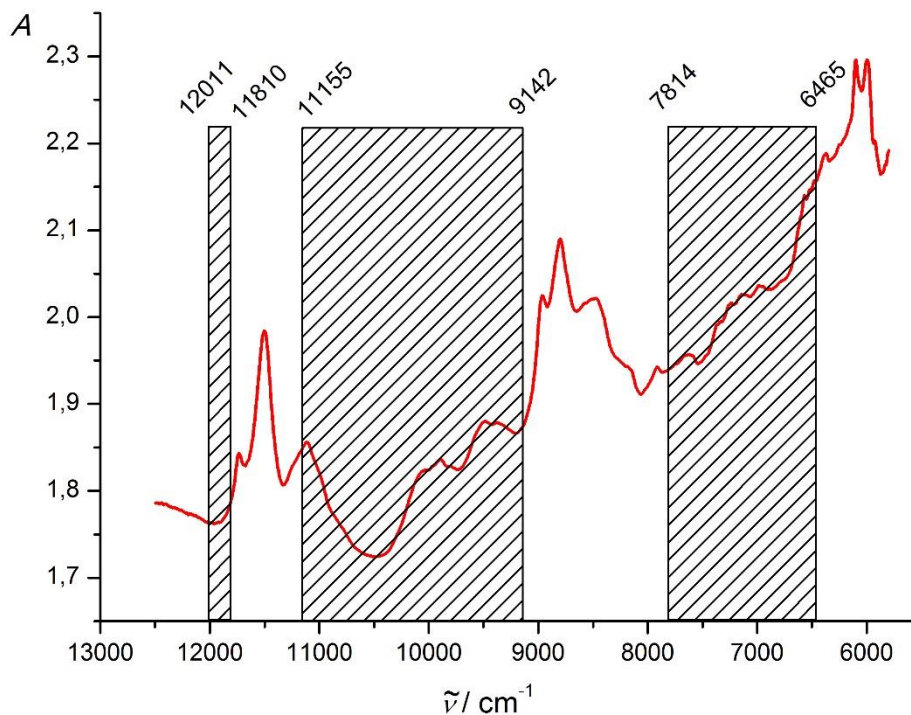


Slika 28. Količine tramadola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i predložene NIR metode. Za kemometričku obradu spektroskopskih podataka korištena je PLS metoda.

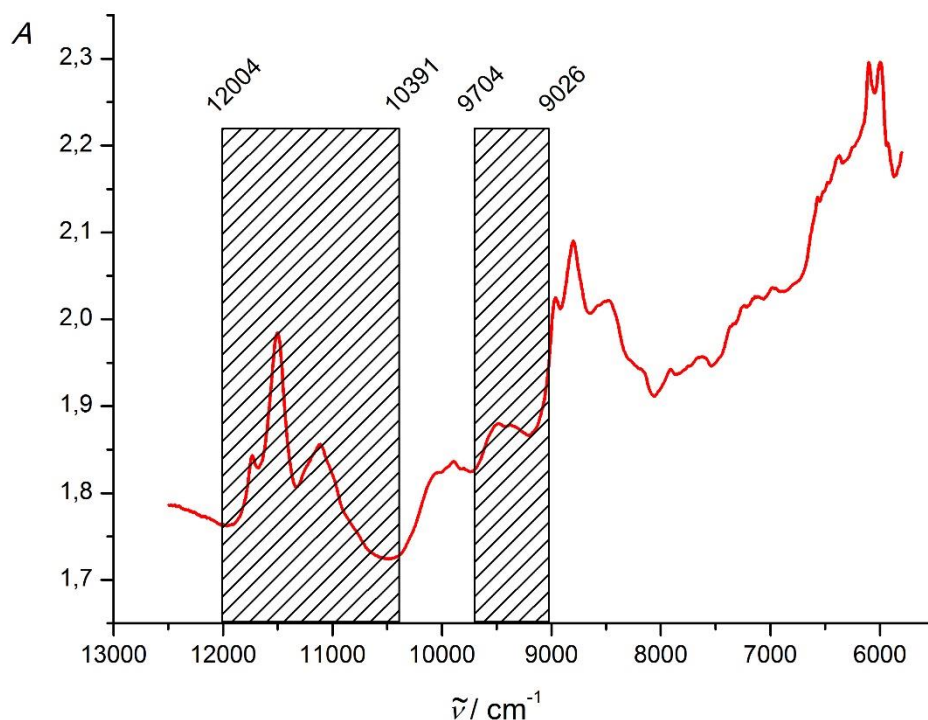
#### 4.2.2. Obrada podataka GA-PLS metodom

NIR spektri kalibracijskog niza tableta te odgovarajuće količine djelatnih tvari određene UHPLC-om upotrijebljeni su za izradu kalibracijskih modela primjenom genetičkog algoritma spregnutog s PLS metodom. Prije provedbe GA smanjen je broj spektralnih podataka uprosječivanjem izmjerenih intenziteta četiriju uzastopnih varijabli. Budući da su vrpce NIR spektara uobičajeno vrlo široke te postoji značajna autokorelacija između spektralnih varijabli ovaj korak je uveden kako bi se dodatno smanjila raznolikost među dobivenim modelima [70]. Kao i u slučaju obrade UV spektrofotometrijskih podataka, GA-PLS izračuni su za svaki analit ponovljeni 10 puta i maksimalno 10 latentnih varijabli je korišteno za izradu PLS modela. Algoritam je zaustavljen nakon kreiranja 100 generacija ili kada je 50 % populacije bilo identično u točki konvergencije. Tako su za paracetamol odabrana spektralna područja 12011 – 11810, 1115 – 9142 i 7814 – 6465  $\text{cm}^{-1}$  (slika 29), a za tramadol 12004 – 10391 i 9704 – 9026  $\text{cm}^{-1}$  (slika 30). Primjenom Savitzky–Golayevog algoritma izračunana je prva derivacija

odabranih dijelova spektara, a potom je provedena njihova transformacija upotrebom SNV metode u slučaju paracetamola, a u slučaju tramadola MSC korekcije.

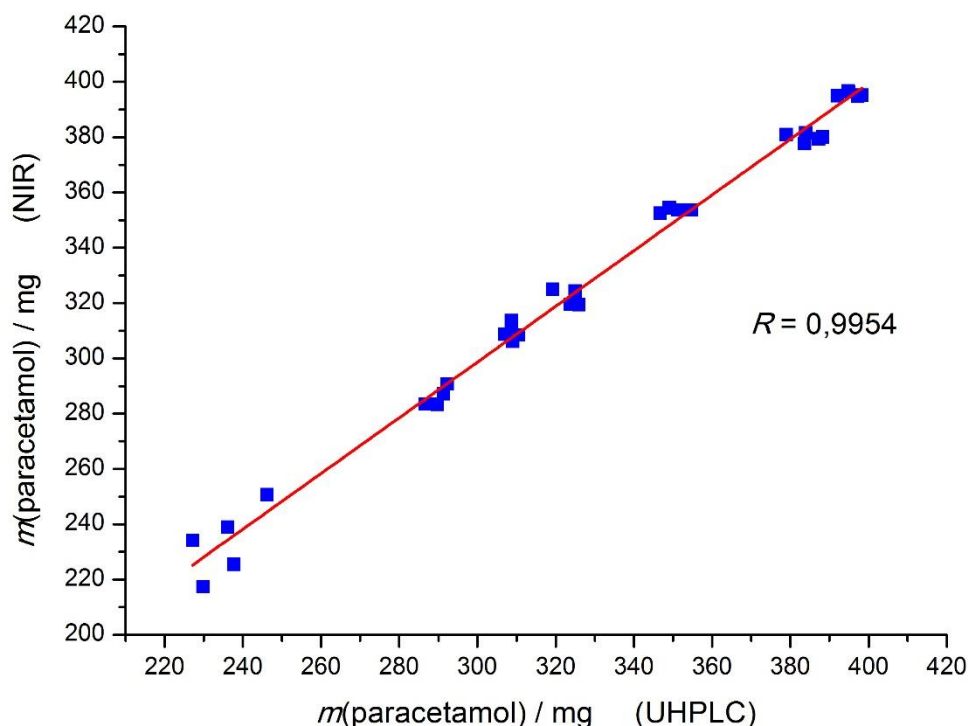


Slika 29. NIR spektralna područja odabrana GA-PLS metodom za predviđanje koncentracije paracetamola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.



Slika 30. NIR spektralna područja odabrana GA-PLS metodom za predviđanje koncentracije tramadola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

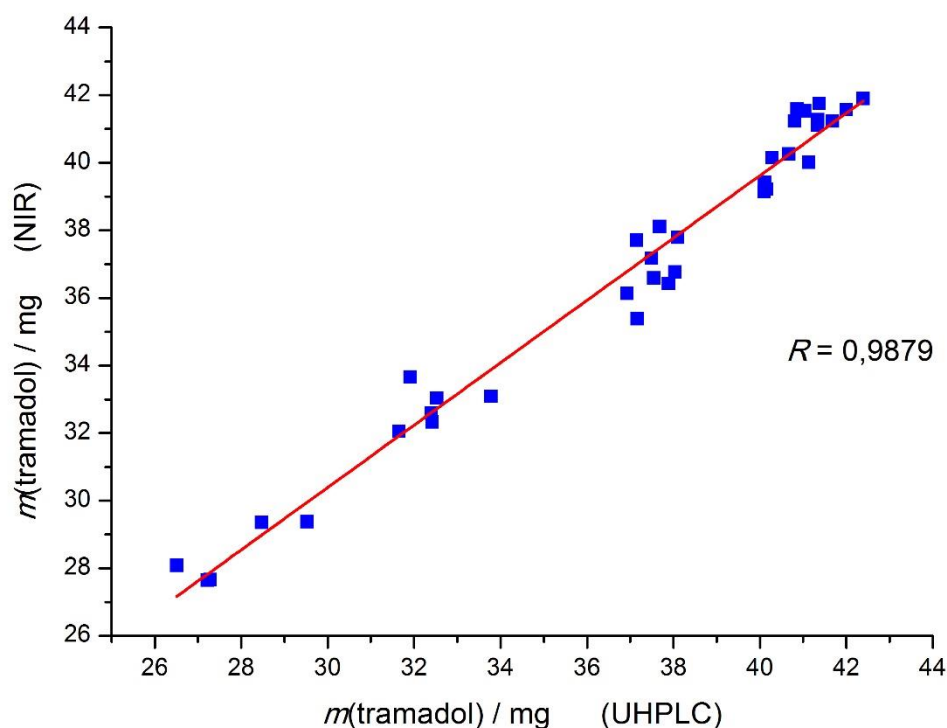
Kao i kod obrade podatka PLS metodom, SNV i MSC algoritmi upotrijebljeni su kako bi se iz spektara uklonilo raspršeno zračenje, utjecaj veličine čestica, tvrdoće tableta i signala bazne linije. Primjenom *cross-validacije* izostavljanja jednog uzorka izračunane su RMSECV vrijednosti na temelju kojih je određen optimalni broj latentnih varijabli u modelima za predikciju koncentracije djelatnih tvari. U slučaju oba analita pronađeno je da je optimalni broj latentnih varijabli 4. Primjenom razvijenih modela izračunane su količine paracetamola i tramadola u validacijskom skupu tableta (tablice 9 i 10). Na temelju tih vrijednosti i količina određenih referentnom metodom izračunane su vrijednosti  $R$  0,9954 (slika 31), RMSEP 5,19 mg, RPD 10,3 i *bias* 1,11 mg za paracetamol, dok je za tramadol dobiveno 0,9879 ( $R$ ) (slika 32), 0,80 mg (RMSEP), 6,11 (RPD) i 0,13 mg (*bias*).



Slika 31. Količine paracetamola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene NIR metode. Za kemometričku obradu spektroskopskih podataka korištena je GA-PLS metoda.

Primjenom GA-PLS-a smanjena je kompleksnost razvijenih modela u odnosu na PLS metodu. U slučaju analize paracetamola PLS-om utvrđeno je da je optimalni broj latentnih varijabli 5, dok je genetičkim algoritmom pronađeno da je taj broj 4. Osim smanjenja broja latentnih faktora, dobiveno je i bolje slaganje između rezultata dobivenih primarnom i predloženom NIR metodom. Tako je vrijednost koeficijenta korelacije izračunanog za koncentracije paracetamola određene PLS-om i UHPLC-om 0,9940, a vrijednost istog parametra određena za koncentracije analizirane primjenom GA-PLS-a i primarne metode je 0,9954. Također, u slučaju GA-PLS metode smanjena je vrijednost korijena srednjeg kvadrata pogreške predviđanja koncentracije djelatne tvari. Za količine paracetamola izračunane primjenom GA-PLS-a utvrđena je RMSEP vrijednost od 5,19 mg, dok je za odgovarajuće koncentracije izračunane PLS metodom dobivena vrijednost 5,86 mg. Budući da RPD predstavlja veličinu koja govori o tome koliko dobro neki kalibracijski model predviđa analizirano svojstvo, ta je vrijednost prikladna za međusobnu usporedbu performansi modela. Kemometrički modeli za koje su vrijednosti RPD-a veće od 5 mogu se upotrebljavati za

kontrolu kvalitete proizvoda [105,106]. RPD vrijednost modela za predviđanje količine paracetamola razvijenog primjenom PLS-a je 9,13, a primjenom GA-PLS-a 10,30. Za obje primijenjene statističke metode RPD je značajno veći od 5 te se obje mogu koristiti u rutinskoj kontroli proučavanog lijeka. Budući da je vrijednost RPD-a za model koji je razvijen upotrebom GA-PLS algoritma veća od odgovarajuće vrijednosti za model razvijen PLS-om, može se zaključiti da su rezultati dobiveni ovom metodom pouzdaniji. Također, izračunane vrijednosti *bias*-a u slučaju paracetamola pokazuju da je sustavna pogreška GA-PLS metode manja (tablica 9). Na temelju svih izračunanih statističkih vrijednosti vidljivo je da je genetički algoritam spregnut s PLS metodom superiorniji u odnosu na PLS metodu.



Slika 32. Količine tramadola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene NIR metode. Za kemometričku obradu spektroskopskih podataka korištena je GA-PLS metoda.

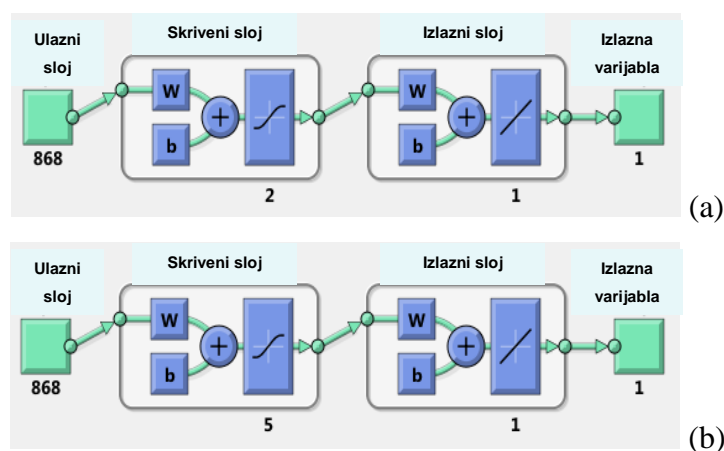
Smanjenje kompleksnosti kemometričkog modela za kvantitativnu analizu tramadola razvijenog upotrebom GA-PLS-a u odnosu na PLS još je značajnije nego u slučaju paracetamola. Tako je analizom tramadola PLS-om utvrđeno da je optimalni broj faktora 8, a GA-PLS algoritmom 4. Također, primjenom GA znatno je poboljšana korelacija rezultata NIR



i UHPLC metode. Koeficijent korelacije izračunan za koncentracije tramadola određene PLS algoritmom i primarnom metodom iznosi 0,9627, a onaj određen za rezultate dobivene GA-PLS-om i referentne metode je 0,9879. Vrijednost korijena srednjeg kvadrata pogreške predviđanja količine tramadola je primjenom GA-PLS-a smanjen s 1,33 mg na 0,80 mg. Odabirom pogodnih spektralnih područja značajno je povećana pouzdanost rezultata NIR metode, što je jasno vidljivo i iz izračunanih RPD vrijednosti. Tako u slučaju PLS-a RPD vrijednost iznosi 3,69, a u slučaju GA-PLS-a ta je vrijednost 6,11. Budući da je za model razvijen primjenom GA-PLS-a RPD veći od 5, ta statistička metoda, za razliku od PLS-a, ima potencijal za upotrebu u kontroli kvalitete Paracetamol + Tramadol 325 mg + 37,5 mg tableta.

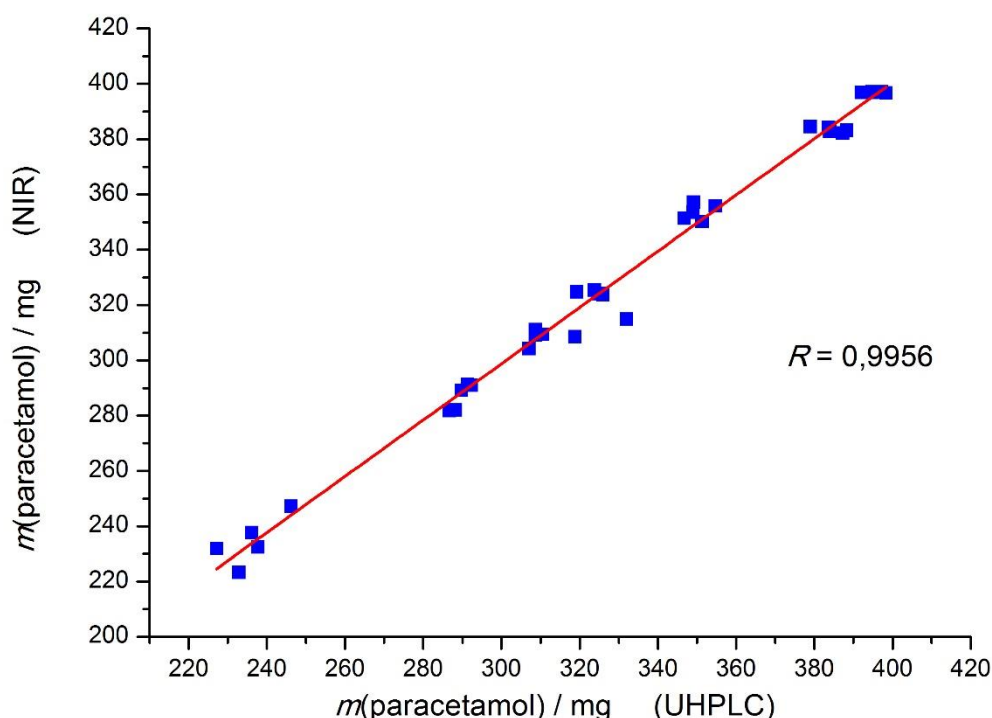
#### 4.2.3. Obrada podataka ANN metodom

NIR spektri kalibracijskog niza tableta te pripadajuće koncentracije djelatnih tvari određene primjenom referentne metode upotrijebljeni su za izradu kalibracijskih modela primjenom ANN metode. Za predviđanje koncentracije oba analita upotrijebljen je *feed-forward* tip neuronskih mreža koje su se, kao i u slučaju obrade spektrofotometrijskih podataka, sastojale od tri sloja, ulaznog, skrivenog i izlaznog. Izmjereni intenziteti NIR spektara kalibracijskih uzoraka uvedeni su u mrežu pomoću ulaznog sloja. Optimalni modeli razvijeni su tako da su u skrivenom sloju u slučaju paracetamola upotrijebljena 2 čvora (slika 33a), a u slučaju tramadola njih 5 (slika 33b). U izlaznom sloju za oba analita korišten je jedan čvor. Kao aktivacijska funkcija u skrivenom sloju upotrijebljen je tangens hiperbolni (j. 66), dok je u izlaznom sloju korištena linearna prijenosa funkcija (j. 67). Trening mreža proveden je primjenom Levenberg-Marquardtovog algoritma.

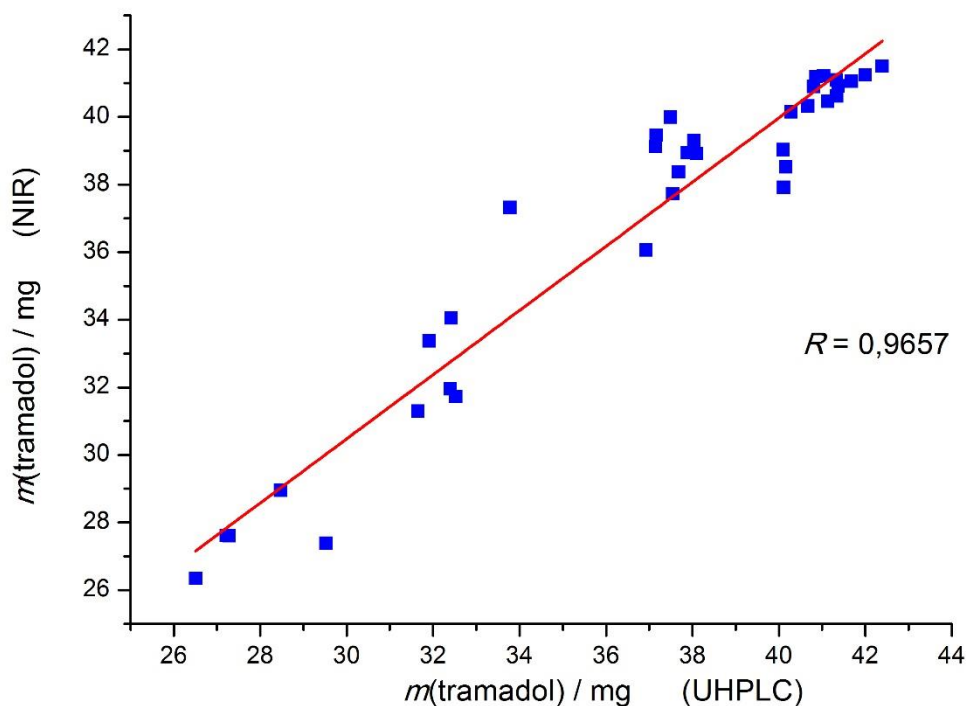


Slika 33. Shematski prikaz neuronskih mreža upotrijebljenih za predviđanje koncentracije paracetamola (a) i tramadola (b).

Tijekom treninga, 70 % nasumično odabranih kalibracijskih uzoraka tableta korišteno je za izradu modela, 15 % za njegovu validaciju i odluku o zaustavljanju treninga, a preostalih 15 % uzoraka iz kalibracijskog skupa upotrijebljeno je za provjeru točnosti podataka izračunanih primjenom trenirane mreže. Performanse razvijenih modela provjerene su analizom 35 tableta iz validacijskog niza uzoraka, kao i u slučaju PLS i GA-PLS metoda, a vrijednosti izračunanih statističkih veličina koje su upotrijebljene za procjenu njihove kvalitete prikazane su u tablici 9 i 10. Tako je za paracetamol dobivena vrijednost  $R$  0,9956 (slika 34), RMSEP 5,06 mg, RPD 10,52 i *bias* 0,70 mg, a odgovarajuće vrijednosti određene za tramadol su 0,9657 ( $R$ ) (slika 35), 1,14 mg (RMSEP), 4,13 (RPD) i - 0,04 mg (*bias*).



Slika 34. Količine paracetamola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene NIR metode. Za kemometričku obradu spektroskopskih podataka korištena je ANN metoda.



Slika 35. Količine tramadola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene NIR metode. Za kemometričku obradu spektroskopskih podataka korištena je ANN metoda.

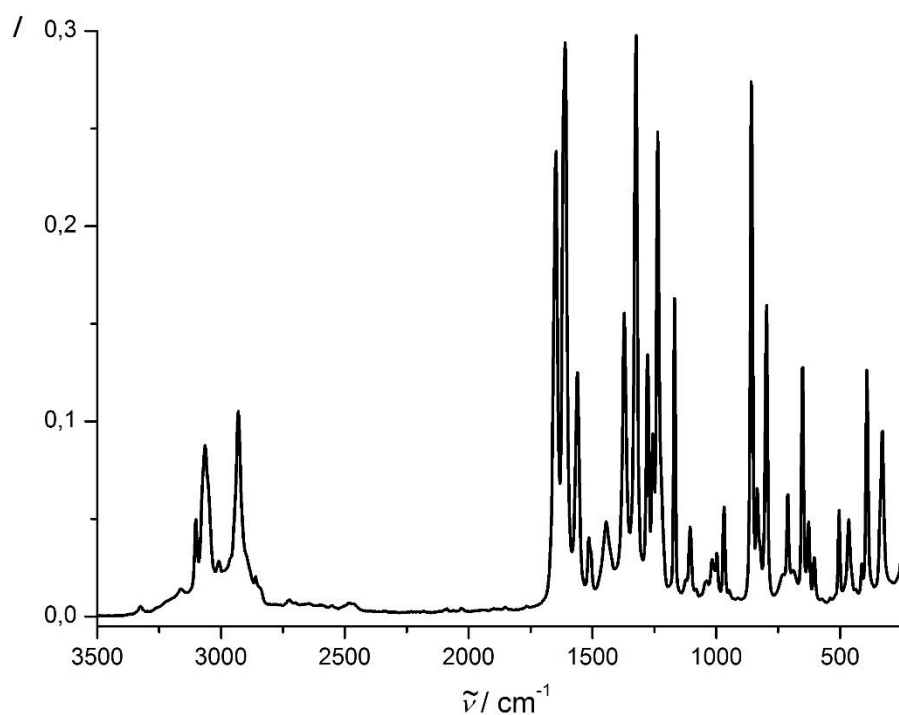
ANN algoritam je za analizu paracetamola superiorniji u odnosu na druge dvije primijenjene regresijske metode, međutim u slučaju GA-PLS-a ta razlika je vrlo mala. Također, rezultati izračunani pomoću modela kreiranog PLS metodom vrlo su pouzdani, ali usporedbi s onima dobivenima pomoću druge dvije metode su najlošiji. Iz toga proizlazi da između koncentracija paracetamola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama i odgovarajućih NIR spektara postoji linearna, ali i slaba nelinearna veza [107]. Prema RPD vrijednostima sva tri modela mogu se koristiti za kvantitativnu kontrolu kvalitete tableta. S druge strane, u slučaju analize tramadola rezultati dobiveni GA-PLS algoritmom su prema svim statističkim pokazateljima značajno bolji u odnosu na one izračunane primjenom PLS i ANN metode. Prema RPD vrijednostima samo model razvijen GA-PLS-om ima odgovarajuće performanse za kvantitativnu analizu. Zbog značajno manje količine tramadola u formulaciji analiziranih tableta u odnosu na onu paracetamola, odabir pogodnih spektralnih područja primjenom genetičkog algoritma spregnutog s PLS metodom ima znatan utjecaj na točnost i

preciznost rezultata izračunanih primjenom ove metode. Ovakva opažanja su u skladu s nalazima opisanim u literaturi.

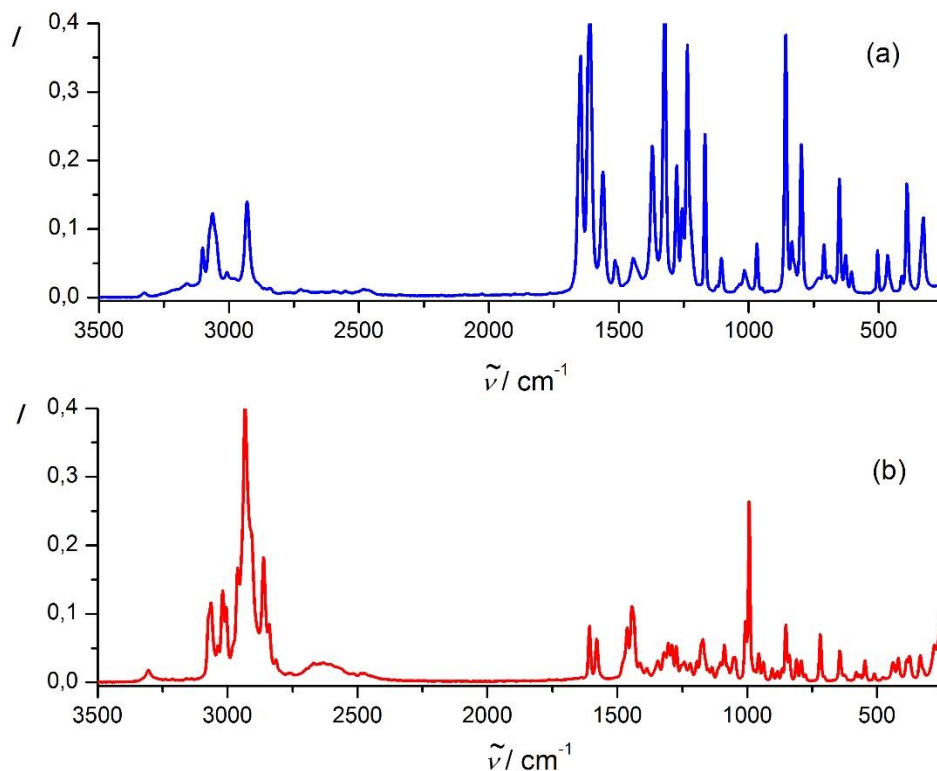
### 4.3. Analiza Ramanovom spektroskopijom

Zbog vremenskog odmaka u analizi za Ramanovu spektroskopiju upotrijebljen je drugi niz kalibracijskih i validacijskih tableta od onih korištenih za razvoj NIR metode. Kao i u slučaju analize NIR spektroskopijom korišteno je 70 uzoraka za izradu kalibracijskih modela, po 10 tableta za svaki koncentracijski nivo. Isto tako, za validaciju je upotrijebljeno 35 tableta, po 5 tableta za svaku koncentraciju djelatnih tvari. Maseni udio paracetamola u korištenim uzorcima bio je u području od 65 do 88 %, a tramadola od 7,5 do 10,1 %. Spektri cijelih tableta snimljeni su u tri položaja s obje strane kako bi se poboljšala njihova kvaliteta. Poznato je da kod analize krutih heterogenih uzoraka Ramanovom spektroskopijom u refleksijskom načinu mjerenja postoji problem poduzorkovanja (engl. *subsampling*) [21,22]. Naime, zbog vrlo malog promjera laserske zrake (od 50  $\mu\text{m}$  do 3 mm) površina uzorka na kojoj se vrši mjerenje je također vrlo mala. Ako su čestice sirovina koje su ugrađene u proizvod veće od površine laserske zrake snimljeni spektar može sadržavati samo informacije o čestici koja je bila izložena zračenju. S druge strane, ako je veličina čestica sirovina manja u odnosu na promjer zrake lasera, snimljeni spektar će biti prosjek svih osvjetljenih čestica te će sadržavati informacije koje odgovaraju ukupnom sastavu proizvoda. Zbog toga je za svaku analiziranu tabletu iz ukupno 768 snimljenih spektara (snimljeno je 128 spektara po jednom položaju) izračunan prosječni spektar. Prosječni spektri i odgovarajuće količine djelatnih tvari određene UHPLC-om upotrijebljene su za statističke izračune promjenom PLS, GA-PLS i ANN metoda. Kao što se može vidjeti na slici 36, Ramanovim spektrom tableta u potpunosti dominira spektar paracetamola. Iako se spektri sirovina razlikuju (slike 37a i 37b), zbog znatno veće količine paracetamola u odnosu na onu tramadola, vrpce tramadola u spektrima tableta uopće nisu vidljive.

Kvaliteta i performanse modela razvijenih primjenom statističkih metoda opisane su u nastavku rada.



Slika 36. Ramanov spektar Paracetamol + Tramadol 325 mg + 37,5 mg tableta. Za pobuđivanje korišteno je lasersko zračenje valne duljine 1064 nm i snage 400 mW.



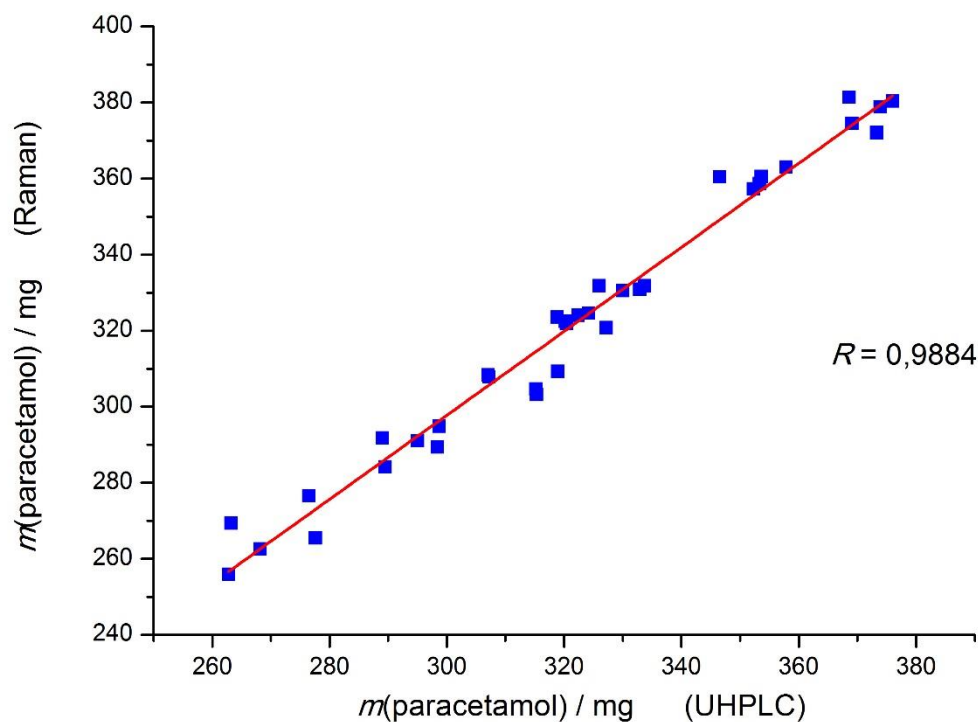
Slika 37. Ramanovi spektri sirovina paracetamola (a) i tramadola (b). Za pobuđivanje korišteno je lasersko zračenje valne duljine 1064 nm i snage 400 mW.

#### 4.3.1. Obrada podataka PLS metodom

Primjenom već opisanog PLS1 algoritma na temelju intenziteta Ramanovih spektara i količina tramadola i paracetamola analiziranih primarnom metodom izrađeni su kvantitativni modeli za predviđanje koncentracije djelatnih tvari u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama. Za predobradu izmjenjenih spektara za oba analita upotrijebljena je MSC korekcija. Ta metoda transformacije upotrijebljena je kako bi se iz spektara uklonila varijabilnost zbog raspršenog zračenja, promjene intenziteta lasera, šuma, itd. Odabir metode za predobradu i optimalnog broja latentnih varijabli izvršen je na temelju prethodno opisane metode *cross-validacije*. Kao i u slučaju obrade NIR podataka, model s najmanjom vrijednosti RMSECV-a i najmanjim brojem latentnih varijabli proglašen je optimalnim. Za paracetamol i tramadol pronađeno je da je model sa 6 faktora optimalan. Kvaliteta modela je i u ovom slučaju procijenjena analizom validacijskog niza tableta te izračunom koeficijenta korelacije, korijena srednjeg kvadrata pogreške predviđanja, *bias*-a i devijacije reziduala predviđanja. Količine paracetamola određene u validacijskim uzorcima primjenom PLS-a i UHPLC-a prikazane su u tablici 11. Iz koncentracija određenih tim dvjema metodama izračunane su sljedeće vrijednosti:  $R$  0,9884 (slika 38), RMSEP 6,43 mg, RPD 5,01 i *bias* - 0,04 mg.

Tablica 11. Analiza paracetamola UHPLC-om i Ramanovom spektroskopijom u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

Tableta	Paracetamol			
	UHPLC	PLS	GA-PLS	ANN
	<i>m</i> / mg	<i>m</i> / mg	<i>m</i> / mg	<i>m</i> / mg
1	262,8	255,9	254,9	258,7
2	276,5	276,6	279,0	285,3
3	327,2	320,8	330,7	323,6
4	332,9	330,9	328,2	327,2
5	330,0	330,5	327,8	329,7
6	326,0	331,8	327,0	322,4
7	333,7	331,8	330,3	321,0
8	263,2	269,4	272,5	269,8
9	277,6	265,5	276,9	281,0
10	268,2	262,6	273,8	270,1
11	352,3	357,3	351,0	353,7
12	346,5	360,5	352,3	354,8
13	353,3	358,7	350,6	360,3
14	353,6	360,6	352,0	356,3
15	357,9	363,1	353,6	364,8
16	320,4	321,8	321,2	324,6
17	324,2	324,6	324,5	330,2
18	320,2	322,5	323,8	314,4
19	322,4	324,0	323,7	328,0
20	318,8	323,6	323,0	325,7
21	373,3	372,1	374,9	374,0
22	376,0	380,4	376,1	375,5
23	373,9	378,9	373,1	383,5
24	368,6	381,4	377,8	377,5
25	369,1	374,5	374,6	376,1
26	318,9	309,3	319,5	315,2
27	315,2	304,7	313,5	321,4
28	307,2	307,8	312,8	307,5
29	307,0	308,4	315,8	301,3
30	315,3	303,2	313,2	306,3
31	289,0	291,8	291,5	289,9
32	295,0	291,0	294,7	290,9
33	298,7	294,8	298,3	295,5
34	298,4	289,4	291,5	296,9
35	289,5	284,2	279,3	279,1
<i>R</i>		0,9884	0,9904	0,9859
RMSEP / mg		6,43	4,56	5,95
RPD		5,01	7,12	5,39
<i>bias</i> / mg		- 0,04	0,59	0,84



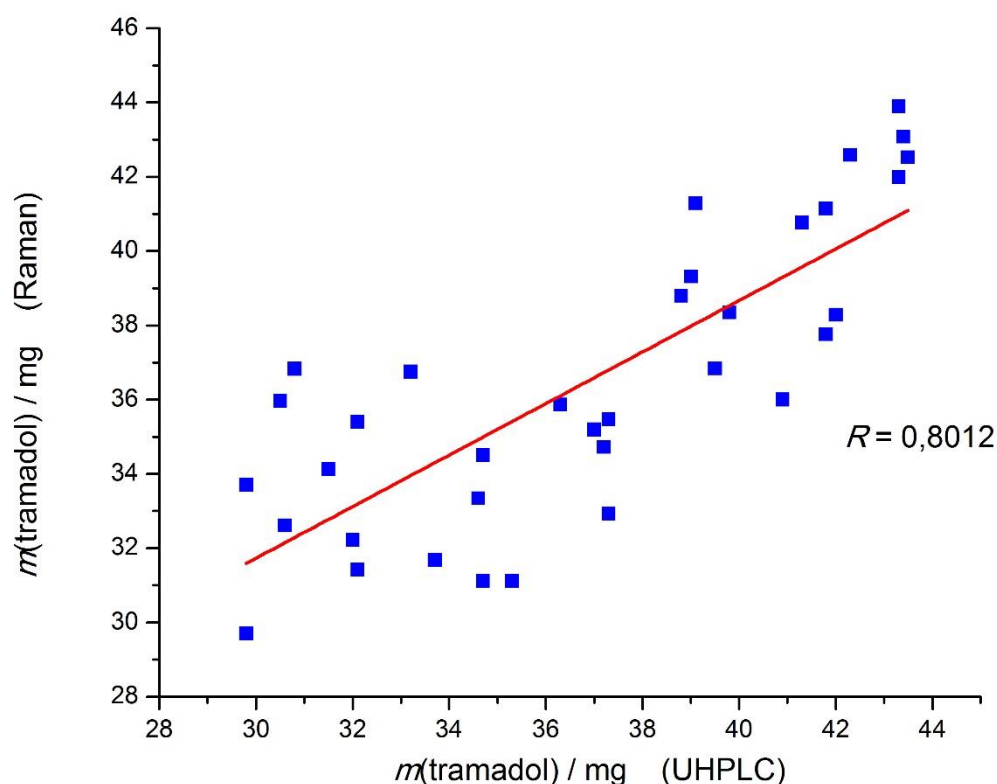
Slika 38. Količine paracetamola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene Ramanove metode. Za kemometričku obradu podataka korištena je PLS metoda.

Na slici 39 prikazane su količine tramadola određene PLS i referentnom metodom u validacijskim uzorcima. Izračunane statističke vrijednosti (tablica 12) iz dobivenih podataka su 0,8012 ( $R$ ), 2,71 (RMSEP), 1,66 ( $bias$ ) i 0,37 (RPD). Na temelju tih vrijednosti može se zaključiti da PLS metoda nije pogodna za kvantitativnu analizu tramadola, dok za paracetamol daje zadovoljavajuće rezultate koji su u vrlo dobrom slaganju s onima dobivenima kromatografskom analizom. Kao što je vidljivo na slici 36, Ramanov spektar Paracetamol + Tramadol 325 mg + 37,5 mg tableta gotovo je identičan spektru paracetamola čiji je maseni udio u formulaciji najveći. Dodatno, intenzitet vrpci u spektru sirovine tramadola je manji u odnosu na onaj paracetamola (slike 37a i 37b). Iako je spektar svake tablete snimljen pri 6 različitih položaja, on ne sadrži dovoljno informacija za zadovoljavajuću kvantitativnu analizu tramadola. Najvjerojatniji uzrok loše kvalitete podataka leži u činjenici da su Ramanovi spektri snimani refleksijskim načinom.



Tablica 12. Analiza tramadola UHPLC-om i Ramanovom spektroskopijom u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

Tableta	Tramadol			
	UHPLC	PLS	GA-PLS	ANN
	<i>m</i> / mg	<i>m</i> / mg	<i>m</i> / mg	<i>m</i> / mg
1	33,2	36,8	35,4	35,2
2	29,8	29,7	30,2	35,8
3	30,8	36,8	34,5	34,2
4	32,1	35,4	36,0	32,9
5	31,5	34,1	34,7	31,4
6	32,0	32,2	33,8	31,6
7	32,1	31,4	32,3	32,3
8	29,8	33,7	32,6	29,7
9	30,6	32,6	29,8	33,1
10	30,5	36,0	33,2	34,3
11	34,6	33,3	33,9	36,4
12	33,7	31,7	33,1	35,6
13	34,7	34,5	35,7	37,9
14	34,7	31,1	32,9	36,1
15	35,3	31,1	32,7	36,3
16	37,3	32,9	32,3	32,4
17	37,2	34,7	34,4	36,4
18	37,3	35,5	34,2	34,3
19	37,0	35,2	34,5	35,6
20	36,3	35,9	35,7	38,5
21	39,1	41,3	40,1	37,2
22	39,8	38,4	39,0	37,0
23	39,5	36,8	38,0	36,2
24	38,8	38,8	39,8	37,1
25	39,0	39,3	39,0	37,4
26	41,8	37,8	35,2	36,7
27	41,8	41,1	38,9	39,8
28	40,9	36,0	34,7	37,4
29	41,3	40,8	39,0	38,8
30	42,0	38,3	36,1	37,8
31	42,3	42,6	42,6	41,2
32	43,3	43,9	42,8	42,2
33	43,5	42,5	42,3	39,8
34	43,4	43,1	42,8	40,7
35	43,3	42,0	43,7	44,0
<i>R</i>		0,8012	0,8032	0,8130
RMSEP / mg		2,71	2,74	2,69
RPD		1,66	1,68	1,65
<i>bias</i> / mg		0,37	0,69	- 0,49

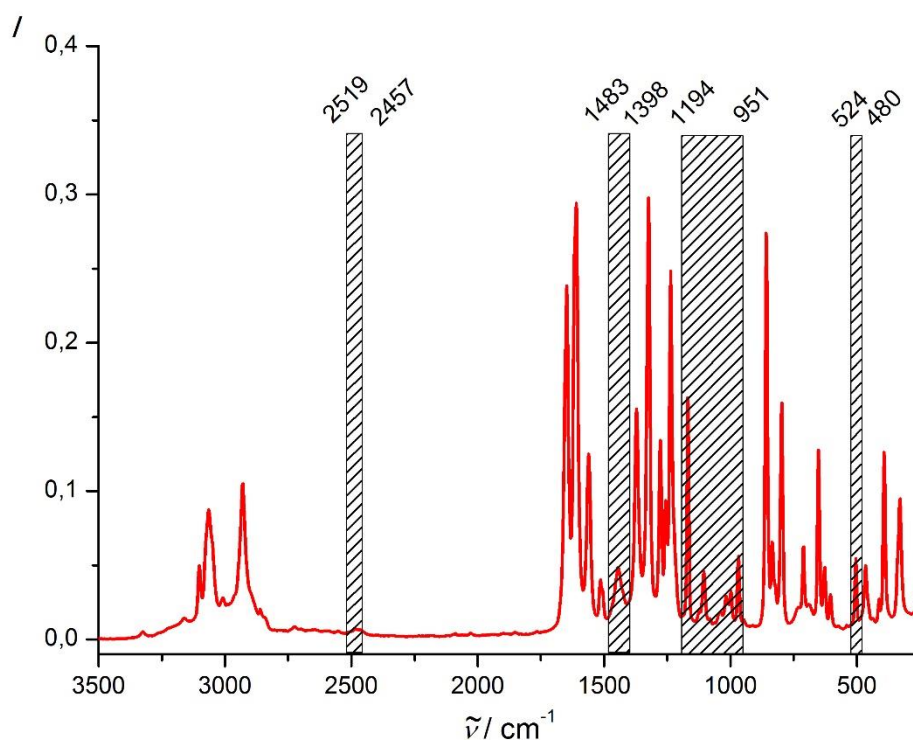


Slika 39. Količine tramadola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene Ramanove metode. Za kemometričku obradu podataka korištena je PLS metoda.

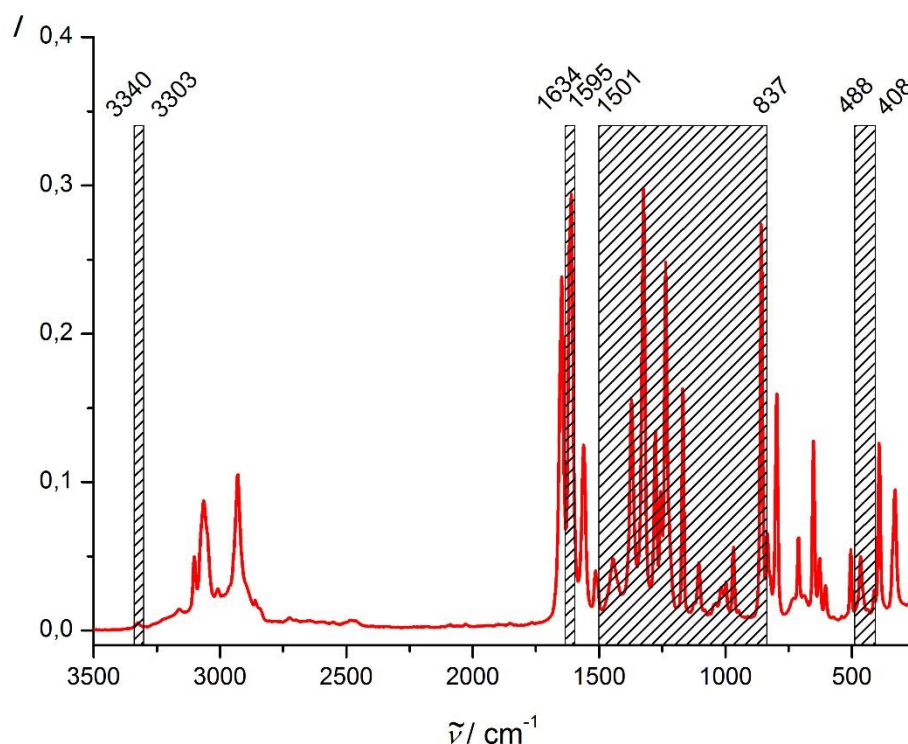
#### 4.3.2. Obrada podataka GA-PLS metodom

Snimljeni Ramanovi spektri obrađeni su također i primjenom GA-PLS metode. Prije provedbe računa smanjen je broj varijabli u spektrima tako da su uprosječeni intenziteti koji su izmjereni pri devet uzastopnih valnih brojeva. Zbog velikog broja varijabli u Ramanovom spektru u odnosu na broj analiziranih uzoraka tableta, GA algoritam odabire mnogo različitih spektralnih područja te se tijekom provedenih 100 generacija kreira mnogo različitih modela. Zbog toga umjesto konvergencije prema optimalnom rješenju, odnosno odabiru najboljih spektralnih područja, metoda odabire područja koja ne sadrže relevantne informacije te se u model uključuju dijelovi spektra koji sadrže šum. Posljedica toga je da kreirani kalibracijski modeli imaju vrlo malu vrijednost RMSECV-a, ali kad se primjene za izračun validacijskog niza

uzoraka RMSEP pogreška je velika. GA-PLS izračun je ponovljen 10 puta za svaki analiti zasebno, te je pri tome upotrijebljeno maksimalno 10 latentnih varijabli za izradu PLS modela. Kao i slučaju obrade UV i NIR podataka, algoritam je zaustavljen nakon kreiranja 100 generacija ili kad je 50 % populacije identično u točki konvergencije. Za paracetamol odabrana su područja 2519 – 2457, 1483 – 1398, 1194 – 951 i 524 – 480  $\text{cm}^{-1}$  (slika 40), a za tramadol 3340 – 3303, 1634 – 1595, 1501 – 837 i 488 – 438  $\text{cm}^{-1}$  (slika 41). Odabrana spektralna područja potom su transformirana primjenom MSC korekcije.



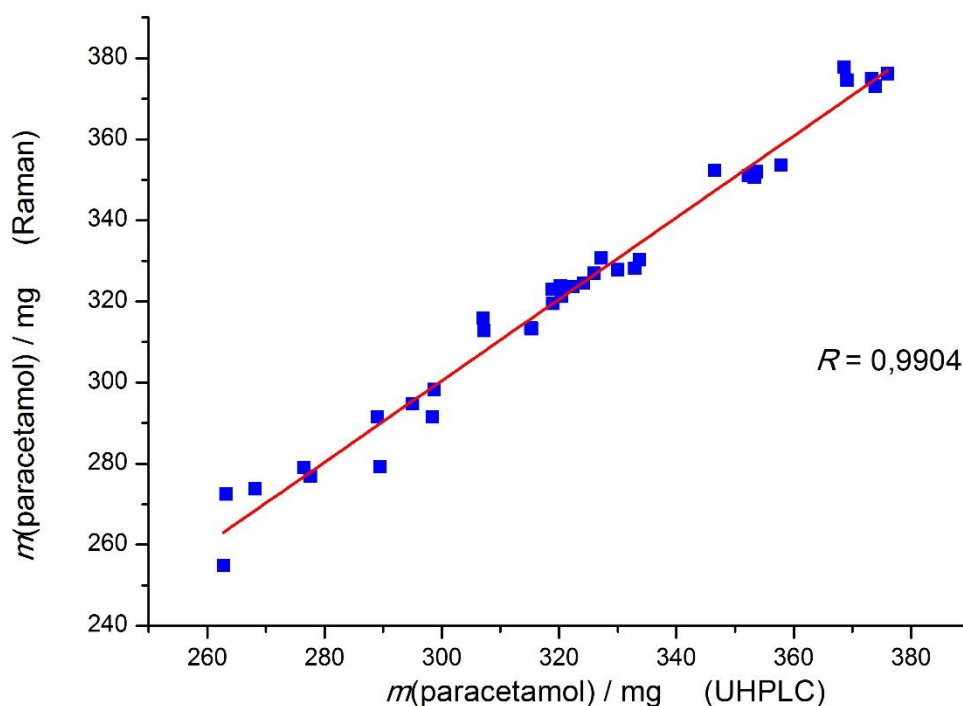
Slika 40. Područja Ramanovog spektra odabrana GA-PLS metodom za predviđanje koncentracije paracetamola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.



Slika 41. Područja Ramanovog spektra odabrana GA-PLS metodom za predviđanje koncentracije tramadola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama.

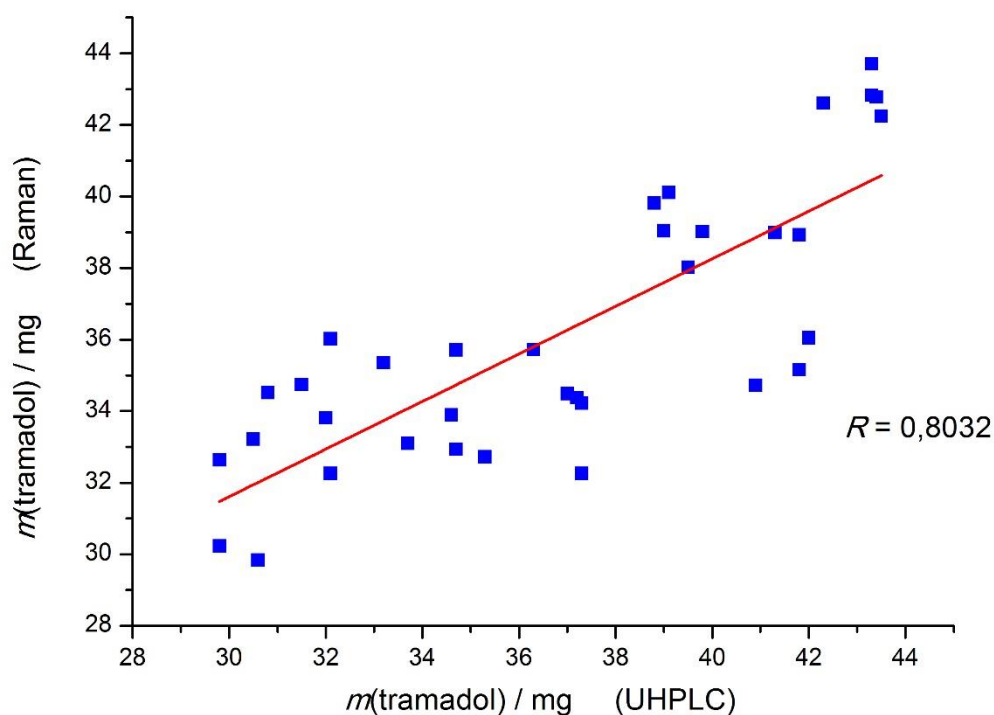
Optimalni kalibracijski modeli određeni su primjenom *cross-validacije* izostavljanja jednog uzorka i izračunane vrijednosti RMSECV-a (j. 10). Za paracetamol pronađeno je da je optimalni broj faktora 2, a za tramadol 4. Na temelju vrijednosti količina djelatnih tvari izračunanih primjenom razvijenih modela i referentnih vrijednosti u validacijskom nizu tableta određene su sljedeće vrijednosti za paracetamol:  $R$  0,9904 (slika 42), RMSEP 4,56 mg, RPD 7,12 i *bias* 0,59 mg. Za tramadol izračunano je  $R$  0,8032 (slika 43), RMSEP 2,74 mg, RPD 1,68 i *bias* - 0,69 mg

GA-PLS metoda se ponovno u slučaju analize paracetamola pokazala superiornija u odnosu na PLS algoritam. Za model kreiran primjenom GA-PLS-a upotrijebljene su dvije latentne varijable, u usporedbi s njih 6 korištenih za PLS metodu. Osim smanjenja kompleksnosti modela, na temelju vrijednosti RMSEP-a može se zaključiti da su rezultati dobiveni GA-PLS metodom pouzdaniji u odnosu na one kada je za izračun korišten cijeli spektar. RPD vrijednosti za obje metode su veće od 5 (tablica 11) pa se bi se obje mogle koristiti u kontroli kvalitete lijeka, međutim performanse GA-PLS metode su bolje.



Slika 42. Količine paracetamola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene Ramanove metode. Za kemometričku obradu podataka korištena je GA-PLS metoda.

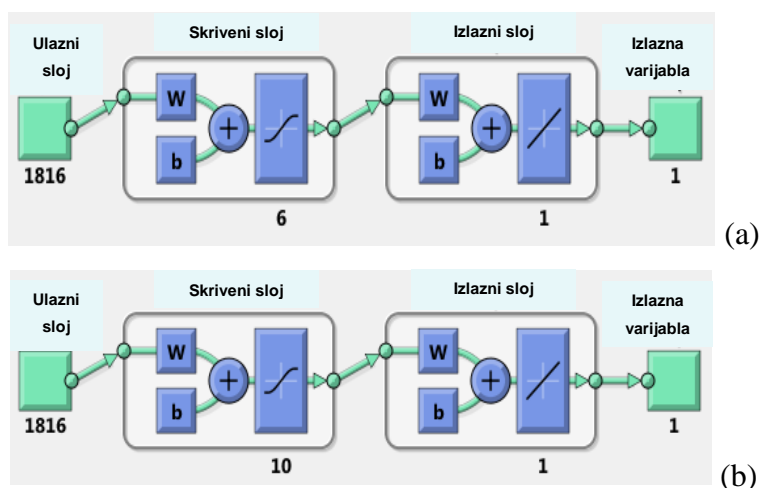
U slučaju analize tramadola, Ramanova spektroskopija, kod koje su spektri snimljeni mjerenjem raspršenog zračenja s površine, nije pogodna za kvantitativno određivanje tog lijeka u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama. Zbog znatno manje količine tramadola u odnosu na onu paracetamola snimljeni spektri ne sadrže dovoljno informacija o tramadolu pa bez obzira na primijenjenu statističku metodu nije moguće dobiti zadovoljavajuću korelaciju između intenziteta vrpca snimljenih spektara i koncentracije djelatne tvari (slika 43). Rezultati izračunani primjenom PLS i GA-PLS metoda međusobno se ne razlikuju značajno i nisu zadovoljavajuće pouzdanosti. Vrijednosti korijena srednjeg kvadrata pogreške predviđanja koncentracije za obje primijenjene statističke metode su velike, 2,71 i 2,74 mg, a koeficijenti korelacije mali, 0,8012 i 0,8032. RPD vrijednosti daleko su manje od 5 što dodatno potvrđuje slabe performanse razvijenih modela. Ramanove spektre trebalo bi snimiti mjerenjem raspršenog zračenja koje je prošlo kroz uzorak kako bi se utvrdilo je li Ramanova spektroskopija pogodna za kvantitativnu analizu tramadola u analiziranim uzorcima.



Slika 43. Količine tramadola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene Ramanove metode. Za kemometričku obradu podataka korištena je GA-PLS metoda.

#### 4.3.3. Obrada podataka ANN metodom

Za obradu Ramanovih spektara Paracetamol + Tramadol 325 mg + 37,5 mg tableta upotrijebljen je *feed-forward* tip neuronskih mreža. Mreža se, kao i slučaju obrade spektrofotometrijskih i NIR podataka, sastojala od ulaznog, skrivenog i izlaznog sloja (slike 44a. i 44b.). Intenziteti Ramanovih spektara uvedeni su u mrežu pomoću ulaznog sloja. U skrivenom sloju za predviđanje koncentracije paracetamola upotrijebljeno je 6 čvorova (slika 44a), a za tramadol njih 10 (slika 44b). Za analizu obje djelatne tvari korišten je jedan čvor u izlaznom sloju. Tangens hiperbolni je korišten kao aktivacijska funkcija, a linearna funkcija kao prijenosna u izlaznom sloju. Mreža je i u ovom slučaju trenirana upotrebom Levenberg-Marquardtovog algoritma. Nasumično je odabrano 70 % uzoraka iz kalibracijskog niza tableta za izradu modela, 15 % za validaciju izračunanih modela i odluku o zaustavljanju treninga, a preostalih 15 % tableta iz kalibracijskog niza koristilo se za provjeru točnosti podataka izračunanih primjenom trenirane mreže.

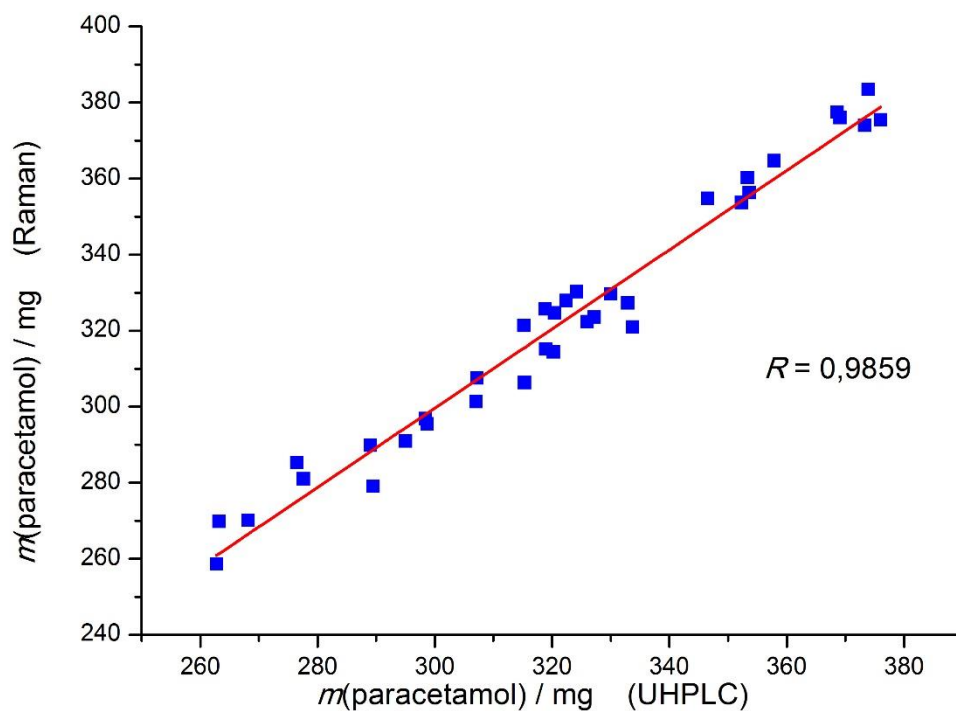


Slika 44. Shematski prikaz neuronskih mreža upotrijebljenih za predviđanje koncentracije paracetamola (a) i tramadola (b).

Provjera trenirane mreže provedena je analizom 35 uzorka validacijskog niza tableta. Rezultati dobiveni ANN analizom prikazani su u tablicama 11 i 12. Na temelju količina djelatnih tvari određenih razvijenom metodom i referentnih vrijednosti dobivenih UHPLC-om izračunane su vrijednosti koeficijenta korelacije, korijena srednjeg kvadrata pogreške predviđanja, *bias*-a i devijacije reziduala predviđanja koncentracije. Za paracetamol je izračunana vrijednost  $R$  0,9859 (slika 45), RMSEP 5,95 mg, RPD 5,39 i *bias* 0,84 mg, a za tramadol  $R$  0,8130 (slika 46), RMSEP 2,69 mg, RPD 1,65 i *bias* - 0,49 mg.

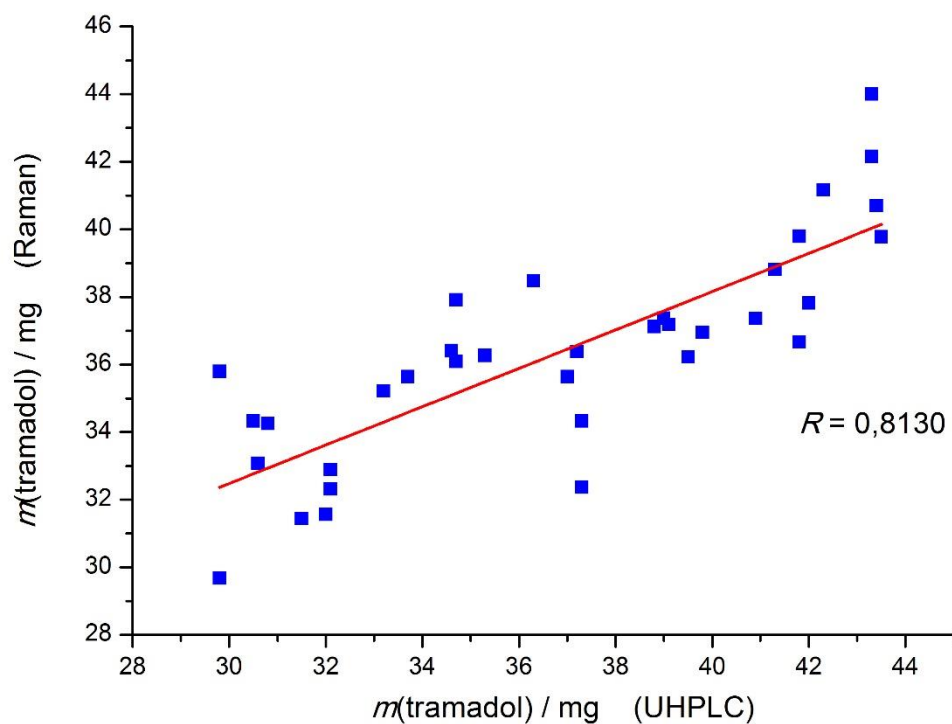
Koncentracije paracetamola određene primjenom ANN algoritma vrlo su slične onima dobivenima PLS metodom. Koeficijenti korelacije, pogreške predviđanja te RPD vrijednosti tih dviju metoda vrlo su slične, što sugerira da su im i performanse slične. S druge strane, GA-PLS metoda ponovno ima najmanju pogrešku predviđanja, te rezultati dobiveni tom metodom imaju najveću točnost. Prema RPD vrijednostima sva tri modela mogu se koristiti za kvantitativnu kontrolu kvalitete tableta.

U slučaju analize tramadola rezultati dobiveni ANN metodom se prema svim statističkim pokazateljima značajno ne razlikuju u odnosu na one izračunane primjenom PLS-a i GA-PLS-a. Prema RPD vrijednostima (tablica 12) niti jedan od modela nema odgovarajuće performanse za kvantitativnu analizu. Kako je već utvrđeno prilikom obrade podataka upotrebom PLS i GA-PLS metoda, Ramanovi spektri snimljeni mjerenjem raspršenog zračenja s površine tableta zbog malog masenog udjela tramadola u formulaciji ne sadrže dovoljno informacija o tom analitu. Zbog toga, bez obzira na upotrijebljenu statističku metodu korelacija između izmjerenih spektara i koncentracije tramadola nije zadovoljavajuća.



Slika 45. Količine paracetamola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene Ramanove metode. Za kemometričku obradu podataka korištena je ANN metoda.





Slika 46. Količine tramadola u validacijskim uzorcima tableta određene primjenom referentne UHPLC i razvijene Ramanove metode. Za kemometričku obradu podataka korištena je ANN metoda.

## § 5. ZAKLJUČAK

U ovom su radu razvijeni novi analitički postupci temeljeni na spektroskopskim metodama i nizu kemometričkih modela za kvantitativno određivanje količine tramadola i paracetamola u Paracetamol + Tramadol 325 mg + 37,5 mg tabletama. Referente količine djelatnih tvari u analiziranim uzorcima tableta određene su kromatografijom ultravisoke djelotvornosti koja je validirana u skladu sa ICH smjernicom o validaciji analitičkih postupaka namijenjenih za kontrolu kvalitete farmaceutskih proizvoda.

Nove UV spektrofotometrijske metode razvijene primjenom djelomičnih najmanjih kvadrata, genetičkih algoritama spregnutih s PLS-om, te umjetnih neuronskih mreža za analizu podataka uspješno su upotrijebljene za istovremeno kvantitativno određivanje djelatnih tvari u tabletama. Za oba lijeka dobiveno je vrlo dobro slaganje rezultata između nove i UHPLC metode, pri čemu se GA-PLS pristup obradi podataka pokazao superiorniji u odnosu na PLS i ANN algoritme. U slučaju tramadola, genetički algoritam spregnut s metodom djelomičnih najmanjih kvadrata znatno je smanjio kompleksnost kalibracijskog modela te poboljšao točnost i relativnu standardnu pogrešku rezultata. Predložene metode su jednostavne, točne i ne zahtijevaju dugotrajnu pripremu uzorka, kao ni upotrebu skupocjene analitičke opreme. Zbog toga se mogu jednostavno koristiti za pouzdano istovremeno određivanje tramadola i paracetamola u tabletama.

Djelomični najmanji kvadrati, GA-PLS i ANN algoritmi uspješno su upotrijebljeni za kvantitativnu obradu NIR spektara za određivanje količina djelatnih tvari u tabletama. U slučaju paracetamola, sva tri razvijena kemometrička modela daju rezultate koji se vrlo dobro slažu s onima dobivenima referentom metodom. I u ovom slučaju primjena PLS metode za obradu cijelih spektara daje lošije rezultate u odnosu na druga dva korištena statistička pristupa. Modeli konstruirani upotrebom umjetnih neuronskih mreža i genetičkih algoritama spregnutih s PLS-om imaju tek neznatno različite performanse. S druge strane, kod kvantitativne analize tramadola razlike u točnosti i preciznosti rezultata izračunanih pomoću različitih kemometričkih metoda značajne su. Odabir valnih duljina genetičkim algoritmom i u ovom slučaju se pokazao superiorniji u odnosu na modele kreirane PLS i ANN metodama. GA-PLS postupak smanjio je broj latentnih varijabli u kalibracijskom modelu i znatno povećao pouzdanost određenih količina tramadola u tabletama. Također, rezultati određeni pomoću

ANN metode točniji su u odnosu na one dobivene PLS-om. Na temelju izračunanih statističkih vrijednosti ( $R$ , RMSEP, RPD i *bias*) može se zaključiti da su sve razvijene NIR metode pogodne za kvantitativnu analizu paracetamola dok se u slučaju tramadola jedino genetički algoritam spregnut s PLS-om može upotrijebiti za kontrolu kvalitete lijeka. Predložene NIR metode su brze, ne zahtijevaju dugotrajnu pripremu uzorka i upotrebu kemikalija pa su zbog toga u usporedbi s drugim tehnikama koje se koriste u farmaceutskoj industriji, npr. tekućinskom kromatografijom visoke djelotvornosti, znatno brže, imaju niže troškove analize i znatno su manje štetne za okoliš. Novi analitički postupci temeljeni na NIR spektroskopiji i kemometričkoj obradi podataka mogu se primijeniti i za određivanje koncentracije djelatnih tvari u proizvodnji tijekom samog proizvodnog procesa tableta kao dio procesne analitičke tehnologije.

Ramanova spektroskopija uspješno je primijenjena za kvantitativnu analizu paracetamola u tabletama. Na temelju izračunanih statističkih vrijednosti pokazano je da su sve tri statističke metode korištene za obradu spektara dale usporedive rezultate s onima određenima UHPLC-om. Razlike u performansama razvijenih modela vrlo su male, iako se pokazalo da su najpouzdaniji rezultati izračunani upotrebom genetičkih algoritama spregnutih s djelomičnim najmanjim kvadratima, što je zamijećeno i u slučaju predloženih UV spektrofotometrijskih i NIR postupaka. S druge strane, Ramanova spektroskopija nije se pokazala pouzdanom za kvantitativnu analizu tramadola. Bez obzira na primijenjene kemometričke postupke za izradu kalibracijskih modela, dobiveni podaci nisu usporedivi s onima određenima kromatografijom ultravisoke djelotvornosti. Zbog mjerenja raspršenog zračenja s površine tablete snimljeni spektri ne sadrže dovoljno informacija o tramadolu zbog njegove relativno male količine u proučavanoj formulaciji. Valjalo bi napraviti mjerenja raspršenog zračenja koje je prošlo kroz uzorak kako bi se smanjio problem poduzorkovanja i dobili spektri koji sadrže više informacija. Također, vrijedilo bi istražiti upotrebu umjetnih neuronskih mreža spregnutih s genetičkim algoritmima ili metodom glavnih komponenta za kvantitativnu obradu prikupljenih spektroskopskih podataka.

## § 6. LITERATURNI IZVORI

- [1] ICH, *Specifications: Test procedures and acceptance criteria for new drug substances and new drug products: chemical substances Q6A*, Guidance, 2000, pp. 1–21.
- [2] Y.C. Shen, *Terahertz pulsed spectroscopy and imaging for pharmaceutical applications: A review*, *Int. J. Pharm.* **417** (2011) 48–60.
- [3] B. Van Eerdenbrugh, L.S. Taylor, *Application of mid-IR spectroscopy for the characterization of pharmaceutical systems*, *Int. J. Pharm.* **417** (2011) 3–16.
- [4] J. Luypaert, D.L. Massart, Y. Vander Heyden, *Near-infrared spectroscopy applications in pharmaceutical analysis*, *Talanta* **72** (2007) 865–883.
- [5] ICH, *Pharmaceutical development Q8(R2)*, Guidance, 2009, pp. 1–28.
- [6] K.E. Bakev (Ed.), *Process analytical technology: Spectroscopic tools and implementation strategies for the chemical and pharmaceutical industries*, 2nd ed., Wiley, Chippingham, Wiltshire, 2010.
- [7] M. Inés Toral, J. Rivas, M. Saldías, C. Soto, S. Orellana, *Simultaneous determination of acetaminophen and tramadol by second derivative spectrophotometry*, *J. Chil. Chem. Soc.* **53** (2008) 1543–1548.
- [8] T.A. Bamigbade, C. Davidson, R.M. Langford, J.A. Stamford, *Actions of tramadol, its enantiomers and principal metabolite, O-desmethyltramadol, on serotonin (5-HT) efflux and uptake in the rat dorsal raphe nucleus*, *Br. J. Anaesth.* **79** (1997) 352–356.
- [9] S. Grond, A. Soblotzki, *Clinical Pharmacology of Tramadol*, *Clin. Pharmacokinet.* **46** (2004) 879–923.
- [10] B.J. Anderson, *Paracetamol (Acetaminophen): Mechanisms of action*, *Paediatr. Anaesth.* **18** (2008) 915–921.
- [11] M. Jozwiak-Bebenista, J.Z. Nowak, *Paracetamol: Mechanism of action, applications and safety concern*, *Acta Pol. Pharm.* **71** (2014) 11–23.
- [12] S. Dhillon, *Tramadol / Paracetamol Fixed-Dose Combination A Review of its Use in the Management of Moderate to Severe Pain*, *Clin. Drug Investig.* **30** (2010) 711–733.
- [13] E. Ziémons, J. Mantanus, P. Lebrun, E. Rozet, B. Evrard, P. Hubert, *Acetaminophen determination in low-dose pharmaceutical syrup by NIR spectroscopy*, *J. Pharm. Biomed. Anal.* **53** (2010) 510–516.

- [14] L.K.H. Bittner, N. Heigl, C.H. Petter, M.F. Noisternig, U.J. Griesser, G.K. Bonn, C.W. Huck, *Near-infrared reflection spectroscopy (NIRS) as a successful tool for simultaneous identification and particle size determination of amoxicillin trihydrate*, *J. Pharm. Biomed. Anal.* **54** (2011) 1059–1064.
- [15] D. Mainali, J. Li, P. Yehl, N. Chetwyn, *Development of a comprehensive near infrared spectroscopy calibration model for rapid measurements of moisture content in multiple pharmaceutical products*, *J. Pharm. Biomed. Anal.* **95** (2014) 169–175.
- [16] Y. Sulub, R. LoBrutto, R. Vivilecchia, B.W. Wabuye, *Content uniformity determination of pharmaceutical tablets using five near-infrared reflectance spectrometers: A process analytical technology (PAT) approach using robust multivariate calibration transfer algorithms*, *Anal. Chim. Acta* **611** (2008) 143–150.
- [17] D.M. Koller, A. Posch, G. Hörl, C. Voura, S. Radl, N. Urbanetz, S.D. Fraser, W. Tritthart, F. Reiter, M. Schlingmann, J.G. Khinast, *Continuous quantitative monitoring of powder mixing dynamics by near-infrared spectroscopy*, *Powder Technol.* **205** (2011) 87–96.
- [18] C. Bodson, W. Dewé, P. Hubert, L. Delattre, *Comparison of FT-NIR transmission and UV-vis spectrophotometry to follow the mixing kinetics and to assay low-dose tablets containing riboflavin*, *J. Pharm. Biomed. Anal.* **41** (2006) 783–790.
- [19] F.J. Muzzio, P. Robinson, C. Wightman, D. Brone, *Sampling practices in powder blending*, *Int. J. Pharm.* **155** (1997) 153–178.
- [20] S.M. Tawakkol, M. Farouk, O.A. Elaziz, A. Hemdan, M.A. Shehata, *Comparative study between univariate spectrophotometry and multivariate calibration as analytical tools for simultaneous quantitation of Moexipril and Hydrochlorothiazide*, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **133** (2014) 300–306.
- [21] M.C. Hennigan, A.G. Ryder, *Quantitative polymorph contaminant analysis in tablets using Raman and near infra-red spectroscopies*, *J. Pharm. Biomed. Anal.* **72** (2013) 163–171.
- [22] R. Szostak, S. Mazurek, *FT-Raman quantitative determination of ambroxol in tablets*, *J. Mol. Struct.* **704** (2004) 229–233.
- [23] M.P. Freitas, L.M. Silva, F. M. Giannotti, A. Couto, E. Tonhi, R.S. Medeiros, G.L. Coco, A. Sabadin, V.F.T. Russo, J.A. Martins, *Prediction of drug dissolution profiles from tablets using NIR diffuse reflectance spectroscopy: A rapid and nondestructive method*, *J. Pharm. Biomed. Anal.* **39** (2005) 17–21.

- [24] H.M. Abdou, *Dissolution, bioavailability and bioequivalence*, Mack Pub. Co., Easton, PA, 1983.
- [25] United States Pharmacopeia, 39th revision 2016, NF34
- [26] H.W. Darwish, S.A. Hassan, M.Y. Salem, B.A. El-Zeany, *Different approaches in Partial Least Squares and Artificial Neural Network models applied for the analysis of a ternary mixture of Amlodipine, Valsartan and Hydrochlorothiazide*, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **122** (2014) 744–750.
- [27] M.R. Khoshayand, H. Abdollahi, M. Shariatpanahi, A. Saadatfard, A. Mohammadi, *Simultaneous spectrophotometric determination of paracetamol, ibuprofen and caffeine in pharmaceuticals by chemometric methods*, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **70** (2008) 491–499.
- [28] M.M. Elkhoudary, R.A.A. Salam, G.M. Hadad, *Comparative artificial neural network and partial least squares models for analysis of metronidazole, diloxanide, spiramycin and cliquinol in pharmaceutical preparations*, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **130** (2014) 222–229.
- [29] Y. Dou, T. Zou, T. Liu, N. Qu, Y. Ren, *Calibration in non-linear NIR spectroscopy using principal component artificial neural networks*, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **68** (2007) 1201–1206.
- [30] Y. Huan, G. Feng, B. Wang, Y. Ren, Q. Fei, *Quantitative analysis of cefalexin based on artificial neural networks combined with modified genetic algorithm using short near-infrared spectroscopy*, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **109** (2013) 308–312.
- [31] L. Sratthaphut, N. Ruangwises, *Genetic algorithms-based approach for wavelength selection in spectrophotometric determination of vitamin B12 in pharmaceutical tablets by partial least-squares*, *Procedia Eng.* **32** (2012) 225–231.
- [32] M.J. Arcos, M.C. Ortiz, B. Villahoz, L.A. Sarabia, *Genetic-algorithm-based wavelength selection in multicomponent spectrometric determinations by PLS: application on indomethacin and acetaminophen mixture*, *Anal. Chim. Acta* **339** (1997) 63–77.
- [33] Z. Bouhsain, S. Garrigues, M. de la Guardia, *Flow injection–Fourier transform infrared spectrometric determination of paracetamol in pharmaceuticals*, *Analyst* **121** (1996) 635–639.
- [34] M.A. Mallah, *Simultaneous quantification of ibuprofen and paracetamol in tablet formulations using transmission Fourier transform infrared spectroscopy*, *Am. J. Anal. Chem.* **3** (2012) 503–511.

- [35] Z. Wang, B. Xiang, *Application of artificial neural network to determination of active principle ingredient in pharmaceutical quality control based on near infrared spectroscopy*, *Microchem. J.* **89** (2008) 52–57.
- [36] M.M. Said, S. Gibbons, A.C. Moffat, M. Zloh, *Near-infrared spectroscopy (NIRS) and chemometric analysis of Malaysian and UK paracetamol tablets: A spectral database study*, *Int. J. Pharm.* **415** (2011) 102–109.
- [37] M. Baptistao, W.F.D.C. Rocha, R.J. Poppi, *Quality control of the paracetamol drug by chemometrics and imaging spectroscopy in the near infrared region*, *J. Mol. Struct.* **1002** (2011) 167–171.
- [38] F.A. Mohamed, M.A. AbdAllah, S.M. Shammat, *Selective spectrophotometric determination of p-aminophenol and acetaminophen*, *Talanta* **44** (1997) 61–68.
- [39] O.A. Ismaiel, M.M. Hosny, *Development and validation of a spectrophotometric method for the determination of tramadol in human urine using liquid-liquid extraction and ion pair formation*, *Int. J. Instrum. Sci.* **1** (2012) 34–40.
- [40] J.C.L. Alves, R.J. Poppi, *Pharmaceutical analysis in solids using front face fluorescence spectroscopy and multivariate calibration with matrix correction by piecewise direct standardization*, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **103** (2013) 311–318.
- [41] M.S. Bloomfield, *A sensitive and rapid assay for 4-aminophenol in paracetamol drug and tablet formulation, by flow injection analysis with spectrophotometric detection*, *Talanta* **58** (2002) 1301–1310.
- [42] A.P. Dewani, B.B. Barik, V.D. Chipade, R.L. Bakal, A.V. Chandewar, S.K. Kanungo, *RP-HPLC-DAD method for the determination of phenylephrine, paracetamol, caffeine and chlorpheniramine in bulk and marketed formulation*, *Arab. J. Chem.* **7** (2012) 811–816.
- [43] F. Musshoff, B. Madea, F. Stuber, U.M. Stamer, *Enantiomeric determination of tramadol and O-desmethyltramadol by liquid chromatography- mass spectrometry and application to postoperative patients receiving tramadol*, *J. Anal. Toxicol.* **30** (2006) 463–467.
- [44] B.N. Patel, N. Sharma, M. Sanyal, P.S. Shrivastav, *An accurate, rapid and sensitive determination of tramadol and its active metabolite O-desmethyltramadol in human plasma by LC-MS/MS*, *J. Pharm. Biomed. Anal.* **49** (2009) 354–366.

- [45] S.T. Ho, J.J. Wan, W.J. Liaw, C.M. Ho, J.H. Li, *Determination of tramadol by capillary gas chromatography with flame ionization detection Application to human and rabbit pharmacokinetic studies*, *J. Chromatogr. B. Biomed. Sci. Appl.* **736** (1999) 89–96.
- [46] Q. Tao, D.J. Stone, M.R. Borenstein, V. Jean-Bart, E.E. Codd, T.P. Coogan, D. Desai-Krieger, S. Liao, R.B. Raffa, *Gas chromatographic method using nitrogen-phosphorus detection for the measurement of tramadol and its O-desmethyl metabolite in plasma and brain tissue of mice and rats*, *J. Chromatogr. B. Biomed. Sci. Appl.* **763** (2001) 165–171.
- [47] Y.F. Sha, S. Shen, G.L. Duan, *Rapid determination of tramadol in human plasma by headspace solid-phase microextraction and capillary gas chromatography-mass spectrometry*, *J. Pharm. Biomed. Anal.* **37** (2005) 143–147.
- [48] H.J. Leis, G. Fauler, W. Windischhofer, *Synthesis of d1-N-ethyltramadol as an internal standard for the quantitative determination of tramadol in human plasma by gas chromatography-mass spectrometry*, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **804** (2004) 369–374.
- [49] J. Li, H. Ju, *Simultaneous determination of ethamsylate, tramadol and lidocaine in human urine by capillary electrophoresis with electrochemiluminescence detection*, *Electrophoresis* **27** (2006) 3467–3474.
- [50] S. Zhao, W. Bai, H. Yuan, D. Xiao, *Detection of paracetamol by capillary electrophoresis with chemiluminescence detection*, *Anal. Chim. Acta* **559** (2006) 195–199.
- [51] C. Engin, S. Yilmaz, G. Saglikoglu, S. Yagmur, M. Sadikoglu, *Electroanalytical investigation of paracetamol on glassy carbon electrode by voltammetry*, *Int. J. Electrochem. Sci.* **10** (2015) 1916–1925.
- [52] C. Radovan, C. Cofan, D. Cinghita, *Simultaneous determination of acetaminophen and ascorbic Acid at an unmodified boron-doped diamond electrode by differential pulse voltammetry in buffered media*, *Electroanalysis* **20** (2008) 1346–1353.
- [53] E.M.P.J. Garrido, J.M.P.J. Garrido, F. Borges, C. Delerue-Matos, *Development of electrochemical methods for determination of tramadol--analytical application to pharmaceutical dosage forms*, *J. Pharm. Biomed. Anal.* **32** (2003) 975–981.
- [54] G. Burgot, F. Auffret, J.L. Burgot, *Determination of acetaminophen by thermometric titrimetry*, *Anal. Chim. Acta* **343** (1997) 125–128.



- [55] D. Easwaramoorthy, Y.C. Yu, H.J. Huang, *Chemiluminescence detection of paracetamol by a luminol-permanganate based reaction*, *Anal. Chim. Acta* **439** (2001) 95–100.
- [56] E. Dinç, A. Ozdemir, D. Baleanu, *An application of derivative and continuous wavelet transforms to the overlapping ratio spectra for the quantitative multiresolution of a ternary mixture of paracetamol, acetylsalicylic acid and caffeine in tablets*, *Talanta* **65** (2005) 36–47.
- [57] E. Dinç, *The spectrophotometric multicomponent analysis of a ternary mixture of ascorbic acid, acetylsalicylic acid and paracetamol by the double divisor-ratio spectra derivative and ratio spectra-zero crossing methods*, *Talanta* **48** (1999) 1145–1157.
- [58] E. Dinç, D. Baleanu, G. Ioele, M. De Luca, G. Ragno, *Multivariate analysis of paracetamol, propiphenazone, caffeine and thiamine in quaternary mixtures by PCR, PLS and ANN calibrations applied on wavelet transform data*, *J. Pharm. Biomed. Anal.* **48** (2008) 1471–1475.
- [59] M.R. Khoshayand, H. Abdollahi, A. Ghaffari, M. Shariatpanahi, H. Farzanegan, *Simultaneous spectrophotometric determination of paracetamol, phenylephrine and chlropheniramine in pharmaceuticals using chemometric approaches*, *Daru* **18** (2010) 292–297.
- [60] R.M. Kamble, S.G. Singh, *Stability-indicating RP-HPLC method for analysis of paracetamol and tramadol in a pharmaceutical dosage form*, *E-Journal Chem.* **9** (2012) 1347–1356.
- [61] F. Ghorbani-Bidkorbbeh, S. Shahrokhian, A. Mohammadi, R. Dinarvand, *Simultaneous voltammetric determination of tramadol and acetaminophen using carbon nanoparticles modified glassy carbon electrode*, *Electrochim. Acta* **55** (2010) 2752–2759.
- [62] T. Zhu, L. Ding, X. Guo, L. Yang, A. Wen, *Simultaneous Determination of Tramadol and Acetaminophen in Human Plasma by LC-ESI-MS*, *Chromatographia* **66** (2007) 171–178.
- [63] ICH, *Validation of a analytical Procedures: text and methodology Q2(R1)*, Guidance, 2005, pp. 1–17.
- [64] J. Emer, J. Miller (Eds.), *Method validation in pharmaceutical analysis: A guide to best practice*, 1st ed., Wiley-VCH, Weinheim, 2006.
- [65] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Elsevier, Amsterdam, 1997.

- [66] D.M. Haaland, E. V. Thomas, *Partial least-squares methods for spectral analyses 1. Relation to other quantitative calibration methods and the extraction of qualitative information*, *Anal. Chem.* **60** (1988) 1193–1202.
- [67] H. Abdi, *Partial Least Square Regression PLS-Regression*, in: N.J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*, Thousands Oaks, CA, 2007, pp. 1–13.
- [68] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, *Chemometrics tools used in analytical chemistry: An overview*, *Talanta* **123** (2014) 186–199.
- [69] Richard Kramer, *Chemometric techniques for quantitative analysis*, Marcel Dekker, New York, Basel, 1998.
- [70] R. Leardi, *Genetic algorithm-PLS as a tool for wavelength selection in spectral data sets*, in: R. Leardi (Ed.), *Nature-inspired Methods in Chemometrics: Genetic Algorithms and Artificial Neural Networks*, 1st ed., Elsevier, Amsterdam, Boston, 2003, pp. 169–194.
- [71] S.H. Chen (Ed.), *Genetic Algorithms and Genetic Programming in Computational Finance*, Kluwer Academic Publishers, Boston, 2002.
- [72] P. Corté, J. Larrañeta, L. Onieva, J.M. García, M.S. Caraballo, *Genetic algorithm for planning cable telecommunication networks*, *Appl. Soft Comput.* **1** (2001) 21–23.
- [73] H. Nazif, L.S. Lee, *Optimised crossover genetic algorithm for capacitated vehicle routing problem*, *Appl. Math. Model.* **36** (2012) 2110–2117.
- [74] P. Sethi, V. Kapoor, *A proposed novel architecture for information hiding in image steganography by using genetic algorithm and cryptography*, *Procedia Comput. Sci.* **87** (2016) 61–66.
- [75] H.-H. Liu, C.-S. Ong, *Variable selection in clustering for marketing segmentation using genetic algorithms*, *Expert Syst. Appl.* **34** (2008) 502–510.
- [76] H.-T. Kim, K.-J. Kim, *Hybrid of Rule-based Systems Using Genetic Algorithm to Improve Platform Game Performance*, *Procedia Comput. Sci.* **24** (2013) 114–120.
- [77] Ö. Gündoğdu, *Optimal seat and suspension design for a quarter car with driver model using genetic algorithms*, *Int. J. Ind. Ergon.* **37** (2007) 327–332.
- [78] C. West, A. Montazeri, S.D. Monk, C.J. Taylor, *A genetic algorithm approach for parameter optimization of a 7DOF robotic manipulator*, *IFAC-PapersOnLine* **49** (2016) 1261–1266.
- [79] A. Supady, V. Blum, C. Baldauf, *First-Principles Molecular Structure Search with a Genetic Algorithm*, *J. Chem. Inf. Model.* **55** (2015) 2338–2348.

- [80] C. Huang, X. Yang, Z. He, *Protein folding simulations of 2D HP model by the genetic algorithm based on optimal secondary structures*, *Comput. Biol. Chem.* **34** (2010) 137–142.
- [81] A. Velásco-Mejía, V. Vallejo-Becerra, A.U. Chávez-Ramírez, J. Torres-González, Y. Reyes-Vidala, F. Castañeda-Zaldivar, *Modeling and optimization of a pharmaceutical crystallization process by using neural networks and genetic algorithms*, *Powder Technol.* **292** (2016) 122–128.
- [82] X.G. Chen, X. Li, L. Kong, J.Y. Ni, R.H. Zhao, H.F. Zou, *Application of uniform design and genetic algorithm in optimization of reversed-phase chromatographic separation*, *Chemom. Intell. Lab. Syst.* **67** (2003) 157–166.
- [83] R. Leardi, *Application of genetic algorithm–PLS for feature selection in spectral data sets*, *J. Chemom.* **14** (2000) 643–655.
- [84] J. Ghasemi, A. Niazi, R. Leardi, *Genetic-algorithm-based wavelength selection in multicomponent spectrophotometric determination by PLS: Application on copper and zinc mixture*, *Talanta* **59** (2003) 311–317.
- [85] U. Depczynski, V.J. Frost, K. Molt, *Genetic algorithms applied to the selection of factors in principal component regression*, *Anal. Chim. Acta* **420** (2000) 217–227.
- [86] T.E.M. Nordling, J. Koljonen, J.T. Alander, P. Geladi, *Genetic algorithms as a tool for wavelength selection*, In J.T. Alander, P. AlaSiuru and H. Hyötyniemi, editors, *The Proceedings of the 11th Finnish Artificial Intelligence Conference STeP 2004*, pages 99–113, Vantaa (Finland), 1–3 Sep. 2004. FAIS Helsinki
- [87] J. Zupan, *Basics of artificial neural networks*, in: R. Leardi (Ed.), *Nature-inspired Methods in Chemometrics: Genetic Algorithms and Artificial Neural Networks*, 1st ed., Elsevier, Amsterdam, Boston, 2003, pp. 199–229.
- [88] W.S. McCulloch, W. Pitts, *A logical calculus nervous activity*, *Bull. Math. Biol.* **52** (1990) 99–115.
- [89] D.O. Hebb, *The Organization of Behavior: A neuropsychological theory*, Wiley, New York, 1949.
- [90] N. Yadav, *History of Neural Networks*, in: N. Yadav, A. Yadav, M. Kumar, *An introduction to neural network methods for differential equations*, Springer, Heidelberg, New York, London, 2015, pp. 13–15.
- [91] M. Minsky, S. Paper, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, Cambridge, 1969.

- [92] D.E. Rumelhart, J.L. McClelland, *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: foundations*, MIT Press, Cambridge, 1986.
- [93] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Q. Appl. Math.* **2** (1944) 164–168.
- [94] D.W. Marquardt, *An algorithm for least-squares estimation of nonlinear parameters*, *J. Soc. Ind. Appl. Math.* **11** (1963) 431–441.
- [95] H. Yu, B.M. Wilamowski, *Levenberg-Marquardt training*, in: J.D. Irwin, B.M. Wilamowski (Eds.), *Intelligent Systems*, 2nd ed., CRC Press, Boca Raton, 2011, pp. 1–16.
- [96] M.T. Hagan, M.B. Menhaj, *Training feedforward networks with the Marquardt algorithm*, *IEEE Trans. Neural Netw. Learn. Syst.* **5** (1994) 989–993.
- [97] European Pharmacopoeia, 8th ed., Council of Europe, Strasbourg, 2014
- [98] M. De Luca, F. Oliverio, G. Ioele, G. Ragno, *Multivariate calibration techniques applied to derivative spectroscopy data for the analysis of pharmaceutical mixtures*, *Chemom. Intell. Lab. Syst.* **96** (2009) 14–21.
- [99] H.C. Goicoechea, A.C. Olivieri, *Chemometric assisted simultaneous spectrophotometric determination of four-component nasal solutions with a reduced number of calibration samples*, *Anal. Chim. Acta* **453** (2002) 289–300.
- [100] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, *A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies*, *J. Pharm. Biomed. Anal.* **44** (2007) 683–700.
- [101] P. Gemperline (Ed.), *Practical guide to chemometrics*, 2nd ed., CRC Press, Boca Raton, London, New York, 2006.
- [102] A. Savitzky, M.J.E. Golay, *Smoothing and differentiation of data by simplified least squares procedures*, *Anal. Chem.* **36** (1964) 1627–1639.
- [103] R.J. Barnes, M.S. Dhanoa, S.J. Lister, *Standard normal variate transformation and detrending of near-infrared diffuse reflectance spectra*, *Appl. Spectrosc.* **43** (1989) 772–777.
- [104] C. De Bleye, P.-F. Chavez, J. Mantanus, R. Marini, P. Hubert, E. Rozet, E. Ziemons, *Critical review of near-infrared spectroscopic methods validations in pharmaceutical applications*, *J. Pharm. Biomed. Anal.* **69** (2012) 125–132.
- [105] M.H. Ferreira, J.W.B. Braga, M.M. Sena, *Development and validation of a chemometric method for direct determination of hydrochlorothiazide in pharmaceutical samples by diffuse reflectance near infrared spectroscopy*, *Microchem. J.* **109** (2013) 158–164.

- [106] B. Igne, C.R. Hurburgh, *Local chemometrics for samples and variables: Optimizing calibration and standardization processes*, *J. Chemom.* **24** (2010) 75–86.
- [107] N. Wangdong, L. Nørgaard, M. Mørup, *Non-linear calibration models for near infrared spectroscopy*, *Anal. Chim. Acta* **813** (2014) 1–14.

## § 7. ŽIVOTOPIS

Siniša Glavanović rođen je 25. kolovoza 1980. godine u Koprivnici. Nakon završene osnovne škole, 1995. godine upisao se u Gimnaziju Fran Galović u Koprivnici gdje je maturirao 1999. godine.

Školovanje je nastavio na Kemijskom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu. Diplomski rad „Praćenje nastajanja proliektrolitno-proteinskih višeslojeva metodom optičke reflektometrije“ izradio je u Zavodu za fizikalnu kemiju Kemijskog odsjeka Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu pod vodstvom prof. dr. sc. Davora Kovačevića. Diplomirao je 2004. godine. U sektoru Istraživanja i razvoja u Belupu zaposlio se 2005. godine.

Usmenim priopćenjem sudjelovao je na Simpoziju studenata doktorskih studija PMF-a (Zagreb, Hrvatska, 2016.). Kao autor na posterskim priopćenjima sudjelovao je na međunarodnom simpoziju „Central European Symposium on Pharmaceutical Technology“ (Dubrovnik, Hrvatska, 2012.) i 5. hrvatskom kongresu farmacije s međunarodnim sudjelovanjem (Rovinj, Hrvatska, 2015.), a kao koautor na međunarodnom simpoziju „Symposium on Macroion Complexation: fundamentals and applications“ (Wageningen, Nizozemska, 2005.).

Doktorsku disertaciju izradio je u Belupu d.d. u Koprivnici, u Zavodu za fizikalnu kemiju Kemijskog odsjeka Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu i u Laboratoriju pokušališta Sveučilišta u Zagrebu Agronomskog fakulteta pod vodstvom prof. dr. sc. Vladislava Tomišića i izv. prof. dr. sc. Marine Vranić.

## Popis publikacija

1. D. Kovačević, S. Glavanović, N. Peran, *Application of optical reflectometry for characterization of polyelectrolyte–protein multilayers*, *Colloids Surf. A Physicochem Eng. Asp.* **277** (2006) 177-182.
2. S. Glavanović, M. Glavanović, V. Tomišić, *Simultaneous quantitative determination of paracetamol and tramadol in tablet formulation using UV spectrophotometry and chemometric methods*, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **157** (2016) 258-264.
3. M. Vranić, K. Bošnjak, S. Glavanović, M. Vinceković, D. Jareš, A. Cundić, *Prediction of Chemical Composition from Semi- natural Grassland by NIR Spectroscopy*, *Agric. Conspec. Sci.* **81** (2016) 35-41.

## Posterska priopćenja na znanstvenim skupovima

1. D. Kovačević, S. Glavanović, N. Peran, *Monitoring of formation of polyelectrolyte-protein multilayers using optical reflectometry*, *XXI. Symposium on Macroion Complexation: fundamentals and applications*, *Book of Abstracts / De Vries, Renko (ur.); Wageningen: Wageningen university, 2005.*
2. M. Jagar, S. Glavanović, M. Bećirević-Laćan, *Metoda in vitro oslobađanja u razvoju oblika za dermalnu primjenu ibuprofena*, *4. hrvatski kongres farmacije, Knjiga sažetaka / Zorc, Branka (ur.); Opatija, Hrvatsko farmaceutsko društvo, 2010.*
3. S. Glavanović, M. Glavanović, V. Tomišić, *Determination of tramadol and paracetamol in their combined tablet formulation using chemometrics*, *The 9th Central European Symposium on Pharmaceutical Technology Book of Abstracts / Lovrić, Jasnima; Pepić, Ivan; Filipović-Grčić, Jelena; Mrhar, Aleš (ur.); Dubrovnik, Faculty of Pharmacy and Biochemistry University of Zagreb, 2012.*
4. S. Glavanović, J. Smetiško, *Determination of ibuprofen release from oral suspension using second-order derivative UV spectrophotometry*, *5. hrvatski kongres farmacije, Knjiga sažetaka / Zorc, Branka (ur.); Rovinj, Hrvatsko farmaceutsko društvo, 2015.*