

# Primjena neuronskih mreža za procjenu koncentracije otopine ksilometazolin hidroklorida u *n*-butanolu primjenom ATR-FTIR spektroskopije *in situ*

T. Herceg,<sup>a</sup> Ž. Ujević Andrijić,<sup>a\*</sup> M. Gavran,<sup>a</sup> J. Sacher,<sup>a</sup> I. Vrban<sup>b</sup> i N. Bolf<sup>a</sup>

Ovo djelo je dano na korištenje pod  
Creative Commons Attribution 4.0  
International License



<sup>a</sup> Sveučilište u Zagrebu Fakultet kemijskog inženjerstva i tehnologije, Zavod za mjerenja i automatsko vođenje procesa, Savska c. 16/5A, 10 000 Zagreb

<sup>b</sup> PLIVA Hrvatska d. o. o., Prilaz baruna Filipovića 25, 10 000 Zagreb

## Sažetak

Procesna analitička tehnologija (PAT) sve se češće primjenjuje u procesu kristalizacije za kontinuirano praćenje neke od ključnih procesnih veličina i značajki kvalitete proizvoda. Vrlo bitne procesne varijable za kristalizaciju hlađenjem su temperatura i koncentracija otopine. Kontinuirano mjerenje koncentracije omogućeno je naprednim *in situ* spektroskopskim instrumentima. U takve metode spada i prigušena ukupna reflektancija Fourier transformacijske infracrvene spektroskopije (engl. *Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy*, ATR-FTIR), primijenjena u ovom istraživanju. U ovom radu razvijeni su kemometrijski modeli za kontinuiranu procjenu koncentracije otopine u stvarnom vremenu metodama strojnog učenja. Cilj ovog rada bio je razviti i ispitati modele parcijalne regresije metodom najmanjih kvadrata (engl. *Partial Least Squares Regression*, PLSR) i modele neuronskih mreža (engl. *Neural Networks*, NN) za modeliranje ovisnosti koncentracije djelatne tvari, ksilometazolin hidroklorid, u *n*-butanolu, o temperaturi i spektralnim podacima dobivenim mjerenjima ATR-FTIR spektrometrom. U radu je provedena predobrada prikupljenih podataka tehnikom MSC (engl. *Multiplicative Scatter Correction*) te je provedeno skaliranje podataka prema normalizacijama *Min-Max* i *Z-score*. Analiziran je broj neurona u prvom i drugom skrivenom sloju, broj skrivenih slojeva, vrsta primijenjenog algoritma učenja (ADAM, NADAM, RMSprop) kao i utjecaj vrste prijenosne funkcije (*ReLU*, *sigmoid*, *tanh*) na kvalitetu razvijenih neuronskih mreža. Razvijeni modeli na sva četiri skupa podataka dali su iznimno dobre rezultate s obzirom na iznose koeficijenta determinacije i srednje kvadratne pogreške. Modelom neuronske mreže dobiveni su koeficijenti determinacije u rasponu vrijednosti od 0,9979 do 0,9989, dok je srednja kvadratna pogreška od 0,0020 do 0,0011. PLSR modelom dobiveni su koeficijenti determinacije od 0,9990 do 0,9995, odnosno srednja kvadratna pogreška od 0,0009 do 0,0005. Dobiveni rezultati pokazali su da predobrada podataka i dodavanje drugoga skrivenog sloja neuronske mreže u ovoj studiji nisu imali velik utjecaj na konačne rezultate. Takav način praćenja i kontrole procesa rezultira efikasnijom proizvodnjom uz smanjenu vjerojatnost pogreške, što u konačnici farmaceutskoj industriji omogućuje brže plasiranje proizvoda na tržište.

## Ključne riječi

Kristalizacija, procesna analitička tehnologija, ATR-FTIR spektroskopija, strojno učenje, ksilometazolin hidroklorid

## 1. Uvod

Kristalizacija je jedinična operacija koja je vrlo bitna za farmaceutsku industriju budući da se tim procesom, najčešće iz otopine, dobiva nova čvrsta faza, odnosno djelatna tvar u kristalnoj formi. *Ostwald-Miersovim* dijagramom definirana su tri koncentracijska područja. Jedno od tih područja je stabilna, nezasićena zona u kojoj kristalizacija nije moguća, a zatim metastabilna zona omeđena krivuljom topljivosti i krivuljom taloženja. U tom području proces kristalizacije moguć je samo uz dodatak cjepiva te se odigrava pri niskom stupnju prezasićenosti da bi se izbjegla velika brzina nukleacije koja rezultira nastajanjem velikog broja nukleusa, što eventualno može pogodovati ugradnji nečistoća u kristalnu rešetku. Iznad krivulje prezasićenosti sustav je labilan i u tom području dolazi do spontane nukleacije.<sup>1</sup> S obzirom na navedeno, može se zaključiti da je precizno mjerenje koncentracije otopine, u bilo kojem trenutku u procesu, preduvjet za vođenje procesa kristalizacije u smjeru dobivanja kristala zahtijevane kvalitete s obzirom na raspodjelu veličina čestica, oblik, čistoću,

polimorfnu formu itd. Uobičajena mjerenja koncentracije temelje se na principima gravimetrije, indeksa loma i kromatografije. No postoje brojna ograničenja za primjenu tih tehnika u *on-line* mjerenjima koncentracije matične otopine. *Off-line* analize podrazumijevaju ručno uzorkovanje i pripremu uzorka, a posljedično i moguću kontaminaciju uzoraka te promjenu koncentracije otopine uslijed temperaturnih promjena za vrijeme analize. Razvojem naprednih spektroskopskih instrumenata s ATR procesnim sondama omogućeno je mjerenje koncentracije sustava u stvarnom vremenu, bez uzorkovanja i naknadne pripreme uzorka.

### 1.1. Procesne analitičke tehnologije

S obzirom na strogu kontrolu proizvodnih procesa i zahtjeve za visokom kvalitetom farmaceutskih proizvoda, Američka agencija za hranu i lijekove (engl. *Food and Drug Administration*, FDA) 2004. godine predlaže inovativan pristup istraživanju i razvoju, proizvodnji i kontroli kvalitete farmaceutskih proizvoda primjenom procesne analitičke tehnologije (engl. *Process Analytical Technology*, PAT).<sup>2</sup> PAT ujedinjuje znanstveni i inženjerski pristup tako da uključuje razvoj i implementaciju analitičkih metoda, napred-

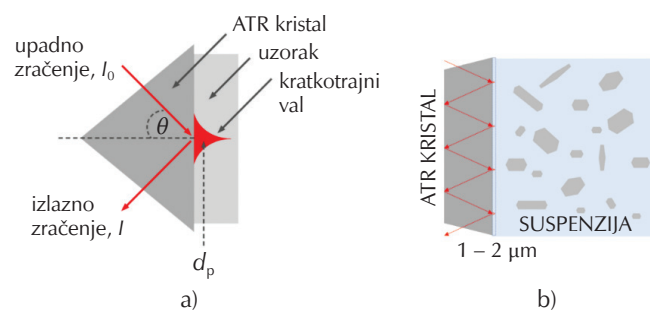
\* Autor za dopisivanje: doc. dr. sc. Željka Ujević Andrijić  
e-pošta: [zujevic@fkit.unizg.hr](mailto:zujevic@fkit.unizg.hr)

nih tehnika i alata za mjerenje ključnih značajki kvalitete i vođenje ključnih procesnih varijabli u stvarnom vremenu tijekom proizvodnje farmaceutskih i drugih proizvoda. Cilj PAT-a je optimirati proizvodni proces i poboljšati kvalitetu i reproducibilnost svojstava konačnog proizvoda. PAT podrazumijeva primjenu naprednih analitičkih tehnika i alata kao što je spektroskopija za kontinuirano praćenje i kontrolu ključnih procesnih varijabli (engl. *Critical Process Parameters*, CPP). Na taj način postiže se bolje razumijevanje procesa, rana identifikacija i rješavanje potencijalnih problema u stvarnom vremenu, čime se otklanja mogućnost neuspjelih šarži, smanjuje se otpad i potrošnja energije, a povećava se djelotvornost procesa i kvaliteta proizvoda uz istodobno ispunjavanje regulatornih zahtjeva.<sup>3</sup>

Kontinuirano mjerenje koncentracije moguće je uz napredne analitičke instrumente koji na temelju svjetlosnog spektra kvantitativno i/ili kvalitativno karakteriziraju određeno svojstvo. Jedna od najčešće primjenjivanih spektroskopskih tehnika je prigušena ukupna reflektancija Fourier transformacijske infracrvene spektroskopije (engl. *Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy*, ATR-FTIR), primijenjena u ovom radu.<sup>4</sup> Posljednjih je godina primjena ATR-FTIR spektroskopije za mjerenje koncentracije matične otopine tijekom procesa kristalizacije u porastu.<sup>5–9</sup>

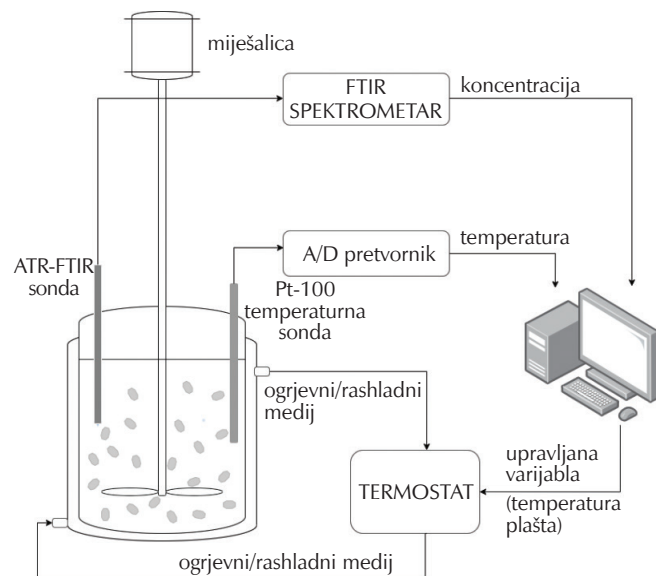
ATR-FTIR spektrometar i optička vlakna osjetljivi su na temperaturne promjene tijekom mjerenja, što otežava precizno određivanje koncentracije otopine. Međutim, primjenom odgovarajućih predobrada spektara kao što su normalizacija, korekcija višestrukog raspršenja (engl. *Multiplicative Scatter Correction*),<sup>10</sup> zaglađivanje (engl. *smoothing*) i derivacija<sup>11</sup> kalibracijskih modela moguće je nadvladati taj problem.

Prigušena potpuna refleksija (ATR) neinvazivna je tehnika uzorkovanja te se najčešće kombinira s infracrvenom (IR) spektroskopijom da bi se dobile kvalitativne i kvantitativne informacije o uzorku. Princip mjerenja temelji se na pojavi kratkotrajnog, tzv. evanescentnog, vala na granici uzorka i ATR kristala s različitim indeksima loma (slika 1). Upadno zračenje prolazi kroz ATR kristal u kojem se reflektira, dok mali dio zračenja u vidu kratkotrajnog vala prodire u optič-



Slika 1 – a) Shematski prikaz pojave kratkotrajnog vala kao rezultata potpune unutarnje refleksije<sup>13</sup> i b) shematski prikaz višestruke refleksije, čime se pojačava intenzitet signala  
Fig. 1 – a) Schematic representation of the appearance of a short-lived wave as a result of complete internal reflection,<sup>13</sup> and b) schematic representation of multiple reflection, which increases the intensity of the signal

ki manje gust materijal tj. uzorak, koji ga djelomično apsorbira, uzrokujući slabljenje potpune refleksije. Reflektirana zraka sadrži informacije o apsorpciji uzorka i prenosi ih dalje do detektora. Dubina prodiranja vala u uzorak iznosi svega  $1 - 2 \mu\text{m}$ , stoga je, neovisno o nehomogenosti kristalizacionog sustava, moguće dobiti informaciju isključivo o otopini.<sup>12</sup>



Slika 2 – Shema praćenja procesa kristalizacije primjenom procesne analitičke tehnologije

Fig. 2 – Scheme for crystallization process monitoring using process analytical technology

Da bi se ATR-FTIR spektroskopija mogla primjenjivati za mjerenje koncentracije u otopini, ključni korak je izrada kalibracijskog (kemometrijskog) modela.<sup>4</sup> Po Lambert-Beerovom zakonu, apsorbanca otopine izravno je proporcionalna koncentraciji apsorbirajućih vrsta u otopini i duljini puta svjetlosti kroz otopine, te je stoga razvojem kalibracijskih modela moguće procijeniti koncentraciju otopine. U zadnje vrijeme popularan je razvoj empirijskih modela tehnikama strojnog učenja (engl. *Machine Learning*). S obzirom na to da su mnogi odnosi između koncentracija spojeva i spektara nelinearni, pored PLSR modela u ovom radu su primijenjene i umjetne neuronske mreže (engl. *Artificial Neural Networks*). Umjetne neuronske mreže su kao nelinearni empirijski modeli pogodni i popularni za primjenu u različitim područjima zbog relativno lake primjene i adaptivnosti, a dobrih sposobnosti predviđanja.<sup>14</sup>

## 1.2. Strojno učenje

Algoritmi strojnog učenja "uče" informacije i odnose među varijablama izravno iz podataka ne oslanjajući se na fundamentalne zakonitosti kao što su temeljni zakoni očuvanja mase, energije i količine gibanja. Pri tome se modeli unaprjeđuju porastom broja uzoraka dostupnih za učenje. Predobrada podataka može biti ključan korak pri razvoju modela strojnog učenja.<sup>15</sup> Ona podrazumijeva pronalaženje ekstremnih vrijednosti (engl. *outliers*), njihovo uklanja-

nje, zamjenu nedostajućih vrijednosti, filtriranje podataka, generiranje dodatnih vrijednosti varijabli, odabir utjecajnih varijabli te skaliranje podataka. Prije samog procesa učenja podatke je potrebno podijeliti na skup za učenje i optimiranje modela te skup za provjeru (validaciju) modela. Skup podataka za učenje trebao bi sadržavati cjelokupan raspon mogućih vrijednosti podataka da bi se moglo dobivenim modelom predviđati njihovo vladanje. Skup za validaciju sadrži podatke nepoznate dobivenom modelu te se njima provjerava kako se model vlada na nepoznatom skupu podataka. Pomoću njega ocjenjuju se generalizacijska svojstva modela.<sup>15</sup> Umjetna neuronska mreža (engl. *neural network*, NN) je skup međusobno povezanih jednostavnih procesnih elemenata (neurona) čija se funkcionalnost temelji na biološkom neuronu te služi distribuiranoj usporednoj obradi podataka.<sup>13</sup> Strukturu umjetne neuronske mreže čine ulazni, skriveni i izlazni sloj. Ulazni sloj prima informacije, ulazne signale, i prosljeđuje ih mreži. Svaki ulazni signal množi se težinskim koeficijentom,  $w_i$ . Tako otežani ulazni signali se zbrajaju, a njihov zbroj se uspoređuje s pragom osjetljivosti neurona,  $T_i$  (engl. *threshold*). Skriveni sloj (engl. *hidden layer*) prima informacije iz ulaznog sloja i obrađuje ih tako što zbraja otežane ulaze pomoću neke funkcije sumiranja i tako stvara vlastitu internu aktivaciju. Ako je zbroj otežanih signala veći od praga osjetljivosti neurona, nelinearna aktivacijska funkcija  $f$  generira izlazni signal neurona iznosa  $b_j$ . Prijenosna funkcija može biti diskontinuirana skokomična funkcija ili neka kontinuirana funkcija, kao npr. sigmoidna ili tangens-hiperbolna funkcija. Ovisno o vrsti prijenosne funkcije koja se primjenjuje, podatke je potrebno prethodno skalirati na vrijednosti s kojima funkcija može računati. Izlazni sloj prima obrađene informacije i daje rezultate. Specificiranje umjetne neuronske mreže podrazumijeva definiranje strukture mreže (shema povezivanja neurona, broj skrivenih slojeva i broj neurona u slojevima), prijenosne funkcije i pravila učenja. Model neuronske mreže definiran je parametrima modela koji se podešavaju tijekom učenja (težinski koeficijenti i pristranost u neuronskim mrežama) te hiperparametrima koji se biraju na temelju prethodno provedenih eksperimenata (prijenosna funkcija, broj neurona, broj skrivenih slojeva, algoritam učenja). Učenje neuronske mreže podrazumijeva iteracijski postupak podešavanja težinskih koeficijenata tako da se dobije minimalna vrijednost pogreške između izlaza dobivenim modelom i stvarne vrijednosti izlazne veličine. Jedno od pravila učenja je pravilo širenja unatrag ili algoritam unatragne propagacije izlazne pogreške (engl. *back propagation*). Pravilo učenja širenja unatrag radi na principu da se na temelju pogreške između stvarne i predviđene vrijednosti, ispravljaju vrijednosti težinskih faktora metodom unatragne propagacije. Postupak se ponavlja sve dok nije zadovoljen zadani kriterij uspješnosti. Parametar koji je važno zadati jest koeficijent učenja, jer on diktira brzinu promjene težinskih koeficijenata s obzirom na vrijednost pogreške. Premala vrijednost može rezultirati presporim učenjem, dok prevelika vrijednost nikad ne može pronaći minimum funkcije pogreške. Da bi se spriječilo pretreniranje mreže, može se primjenjivati tehnika ranog zaustavljanja (engl. *early stopping*). Ona omogućuje zaustavljanje učenja onda kad pogreška na skupu za učenje i dalje pada, a pogreška na skupu za provjeru dosegne minimum.<sup>17</sup>

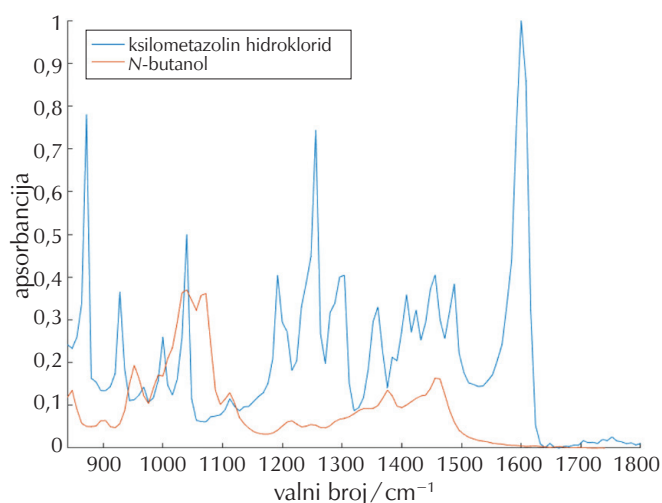
Još jedna tehnika strojnog učenja je parcijalna regresija metodom najmanjih kvadrata (engl. *Partial Least Squares Regression*, PLSR). Ta tehnika povezuje značajke metode glavnih komponenti (engl. *Principal Component Analysis*, PCA) i viševeličinske linearne regresije (engl. *Multiple Linear Regression*, MLR). Primjenjuje se kad je potrebno definirati skup zavisnih varijabli iz vrlo velikog skupa nezavisnih varijabli. Iako je prvotno razvijena za primjenu u društvenim znanostima (ekonomija Herman Wold 1966.), svoju popularnost stekla je u kemometriji. Cilj PLSR metode je, kao i kod PCA, reducirati broj komponenti danih varijabli na manji skup komponenti koje nisu međusobno korelirane. Ono što je specifično za PLSR je da, za razliku od PCA, uzima u obzir i zavisnu ( $X$ ) i nezavisnu varijablu ( $Y$ ), odnosno maksimizira kovarijancu između tih dviju varijabli. Ono što povezuje PLSR s MLR jest regresija. Najprije se PLSR metodom reducira broj komponenti danih varijabli te se s njima (umjesto s originalnim) ulazi u regresijski model.<sup>18</sup>

U ovome radu, primjenom PLSR metode i umjetnih neuronskih mreža, razvijen je kalibracijski model za kontinuirano praćenje koncentracije ksilometazolin hidroklorida u procesu kristalizacije.

## 2. Eksperimentalni dio

### 2.1. Materijali

Za izradu kalibracijskoga modela iz prikupljenih spektralnih podataka upotrijebljena je djelatna tvar ksilometazolin hidroklorid (molekulska formula:  $C_{16}H_{25}ClN_2$ , molekulska masa:  $280,8 \text{ g mol}^{-1}$ ) otopljena u *n*-butanolu (molekulska formula:  $C_4H_9OH$ , molekulska masa:  $74,123 \text{ g mol}^{-1}$ ). Na slici 3 prikazani su spektri ksilometazolin hidroklorida i *n*-butanola da bi se razlikovala spektroskopska slika



Slika 3 – Usporedba *offline* spektra ksilometazolin hidroklorida sa spektrom *n*-butanola

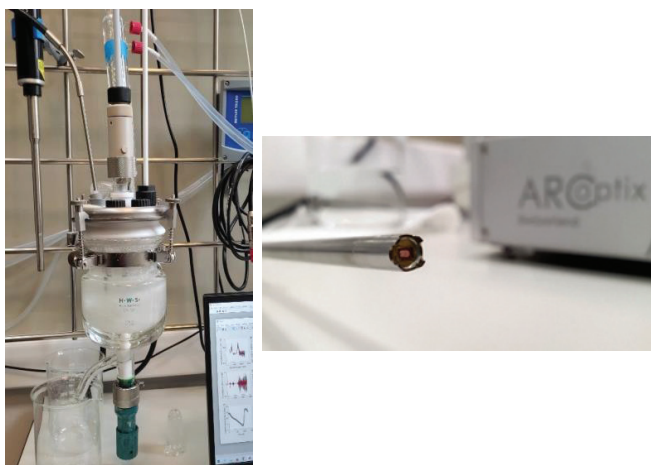
Fig. 3 – Comparison of *offline* xylometazolin hydrochloride spectrum and *n*-butanol spectrum



molekule ksilometazolin hidroklorida i otapala. *Off-line* ATR-FTIR spektroskopijom (PerkinElmer, Spectrum One, Waltham, MA, SAD) provedena je karakterizacija kristala te djelatne tvari da bi se potvrdili literaturni podatci o karakterističnim apsorpcijskim vrpčama te, shodno tome, potvrdila primjenjivost mjerne metode za otopinu ksilometazolin hidroklorida u *n*-butanolu.<sup>19</sup>

## 2.2. Uređaji

Aparatura za provedbu eksperimenta prikazana je na slici 4 te se sastoji od staklenoga kristalizatora s plaštem volumena 500 ml, temperaturnoga osjetila Pt-100 i termostata za regulaciju temperature (Julabo, Magio MS-1000F, Seelbach, Njemačka) te stupne miješalice s turbinskim miješalom od teflona. Za prikupljanje spektralnih podataka upotrijebljen je FTIR spektrometar (ArcOptix, FC-120-4TE, Neuchatel, Švicarska) s dijamantnom ATR sondom uronjenom u otopinu.



Slika 4 – Aparatura: kristalizator s plaštem, miješalom, Pt-100 temperaturnom sondom i ATR FTIR sondom (lijevo), ATR FTIR spektrometar (desno)

Fig. 4 – Apparatus: jacketed crystallizer, stirrer, Pt-100 temperature probe and ATR FTIR probe (left), ATR FTIR spectrometer (right)

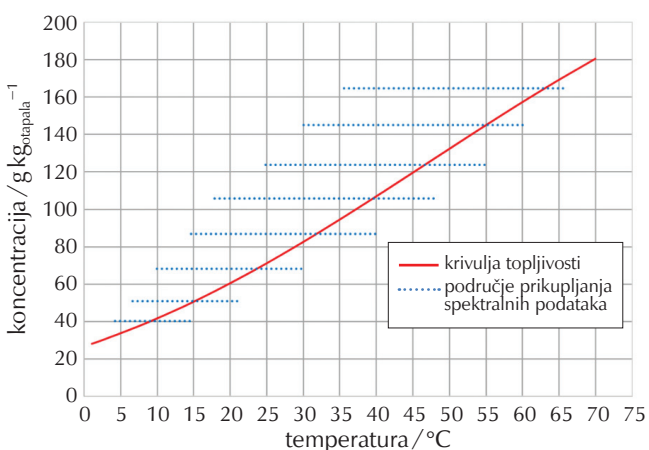
## 2.3. Postupak rada

### 2.2.1. Određivanje krivulje topljivosti

Prvi korak u provođenju eksperimenta bio je određivanje topljivosti ksilometazolin hidroklorida u *n*-butanolu. *N*-butanol je u reaktoru ohlađen na 0 °C te je dodano 40 g ksilometazolin hidroklorida. Nakon dodatka mase od 3 g po potrebi se povećala temperatura u suspenziji do potpunoga otapanja kristala, odnosno za 1 do 4 °C. Topljivost ksilometazolin hidroklorida u *n*-butanolu određena je pri različitim temperaturama te prikazana na slici 5.

### 2.2.2. Priprema otopina i prikupljanje temperatura

Drugi korak u provođenju eksperimenta bio je priprema otopina poznatih koncentracija. Na temelju prethodno određene topljivosti ksilometazolin hidroklorida u *n*-butanolu odabrano je osam koncentracija koje se protežu kroz raspon temperatura u metastabilnoj i nezasićenoj zoni. Otopine koncentracija 40,1, 55,4, 71,1, 88,5, 107,0, 125,1, 144,3 i 164,8 g kg<sup>otapala</sup><sup>-1</sup> pripremljene su otapanjem definiranih masa ksilometazolin hidroklorida u *n*-butanolu. Otopine su zatim zagrijavane do maksimalno 10 °C iznad temperature topljivosti da bi se osiguralo potpuno otapanje djelatne tvari. Nakon toga otopine su hladene konstantnom brzinom hlađenja od 0,3 °C min<sup>-1</sup>. Tijekom hlađenja bilježeni su spektralni podatci s promjenom temperature.



Slika 5 – Krivulja topljivosti i temperaturno područje snimanja spektralnih podataka u otopinama zadanih koncentracija

Fig. 5 – Solubility curve and temperature range of obtained spectral data in solutions of given concentrations

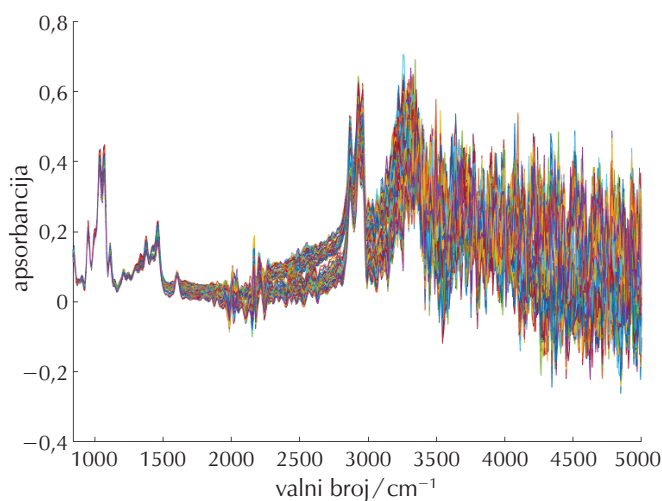
### 2.2.3. Prikupljanje i predobrada spektralnih podataka

Prikupljanje spektralnih podataka iz nezasićene i metastabilne zone ksilometazolin hidroklorida u *n*-butanolu provedeno je *in situ* ATR-FTIR spektroskopijom. Za odabrano temperaturno područje pri definiranoj koncentraciji dobiva se spektar s rasponom valnoga broja od 840 do 5000 cm<sup>-1</sup>. Prije razvoja modela snimljeni spektralni podatci predobrađeni su tehnikom MSC (engl. *Multiplicative Scatter Correction*).<sup>20,21</sup> Svrha te tehnike predobrade podataka jest smanjivanje raspršenja u spektralnim podacima kao i eliminacija odstupanja (engl. *outliers*). Za vizualizaciju i predobradu podataka upotrijebljen je programski jezik *Python* i pripadajući paketi *Scikit learn*, *Pandas*, *Numpy*, *Seaborn* i *Matplotlib* (*Python Software Foundation*, verzija 3.9.7, Delaware, SAD).

### 2.2.4. Razvoj kalibracijskog modela

Za razvoj modela neuronske mreže primijenjeni su snimljeni spektralni podatci u rasponu valnoga broja 840 – 1800 cm<sup>-1</sup> s pripadajućim temperaturama. Snimlje-

ni spektralni podatci otopine ksilometazolin hidroklorida u *n*-butanolu iznad  $1800\text{ cm}^{-1}$  sadržavali su šum te nisu primijenjeni za razvoj modela (slika 6).



Slika 6 – Neobrađeni FTIR spektar ksilometazolin hidroklorida u *n*-butanolu

Fig. 6 – Raw FTIR spectrum of xylometazoline hydrochloride in *n*-butanol

U ovome je radu za razvoj i ispitivanje modela neuronske mreže upotrijebljen programski jezik *Python*, odnosno njegovo programsko okruženje *Spyder* (engl. *Scientific Python Development Environment (Python Software Foundation, verzija 3.9.7, Delaware, SAD)*). Za razvoj modela upotrijebljeni su skupovi podataka prikazani u tablici 1, dimenzija matrice je  $354 \times 122$ , koja obuhvaća podatke prikupljenih temperatura veličine  $354 \times 1$ , koncentracija  $354 \times 1$  te apsorbancija  $354 \times 120$ . Za određivanje valjanosti razvijenih modela primijenjeni su kriterij srednje kvadratne pogreške (engl. *mean squared error, MSE*) i koeficijent determinacije (engl. *coefficient of determination, R<sup>2</sup>*). Pri tome su se navedeni kriteriji pratili na setu za učenje modela te

na setu za validaciju modela. U svrhu dobivanja što točnijega modela provjeravao se utjecaj broja skrivenih slojeva, utjecaj broja neurona u prvome i drugom skrivenom sloju, vrsta aktivacijske funkcije. U ovome radu primijenjeni su optimizacijski algoritmi, tzv. optimizatori koji se danas najčešće primjenjuju kod razvoja neuronskih mreža, a temelje se na stohastičkom algoritmu gradijentnog spusta kao što su ADAM, NADAM i RMSprop. Sva tri algoritma temelje se na adaptivnim stopama učenja, što znači da prilagođavaju stopu učenja trenutačnim vrijednostima težinskih faktora neuronske mreže na temelju gradijenata funkcije gubitka. Svi navedeni optimizatori dizajnirani su da bi prevladali izazove tradicionalnih algoritama s gradijentnim spustom, kao što je spora konvergencija i zaglavlivanje u lokalnim minimumima. NADAM predstavlja proširenje ADAM optimizatora, a uključuje Nesterovu metodu (engl. *Nesterov accelerated gradient*) ubrzavanja konvergencije uzimajući u obzir moment gradijenata. RMSprop optimizator za kontrolu brzine ažuriranja prosjeka gradijenata primjenjuje parametar zaboravljanja (engl. *decay*), dok ADAM i NADAM primjenjuje kombinaciju eksponencijalnih pomičnih prosjeka za procjenu prvog i drugog momenta gradijenata. Pokazalo se da NADAM daje bolje rezultate od ADAM optimizatora na nekim skupovima podataka, dok primjerice RMSprop optimizator daje zadovoljavajuće rezultate na dinamičkim RNN i LSTM mrežama.<sup>22–24</sup> Potrebno je zadati i broj epoha (broj prolaska svih podataka kroz mrežu) te veličinu grupe podataka (engl. *batch size*) koja predstavlja broj uzoraka sljedova podataka koji se obrađuju kao cjelina u jednom prolazu kroz neuronsku mrežu. Broj epoha može imati značajan utjecaj na ishod učenja neuronske mreže. Ako je broj epoha premali, model možda neće imati dovoljno vremena da nauči temeljne obrasce u podacima, što dovodi do loše izvedbe modela. S druge strane, ako je broj epoha prevelik, model se može pretrenirati, što dovodi do loše generalizacije modela na novim podacima.<sup>25</sup> Pored praćenja pogreške na oba skupa podataka s povećanjem broja epoha u praksi je uobičajeno primjenjivati i tehnike poput ranog zaustavljanja a bi se pronašao optimalan broj epoha za određeni zadatak koji može varirati ovisno o varijabilnosti podataka. U ovom radu ispitivale su se prijenosne funkcije koje su pogodne za

Tablica 1 – Skupovi podataka za optimiranje modela neuronske mreže

Table 1 – Data sets for neural network models optimisation

Skup podataka Data set	Dimenzije matrice Dimensions of matrix	Predobrada Preprocessing	Broj skrivenih slojeva Number of hidden layers	Broj neurona u skrivenom sloju Number of neurons in hidden layer	Prijenosne funkcije Transfer functions	Optimizatori Optimisers
D1	$354 \times 122$	/	1	2-112, korak 10	<i>sigmoid</i>	Nadam, Adam, RMSprop
					<i>tanh</i>	
					<i>ReLU</i>	
D2	$354 \times 122$	MSC	1	2-112, korak 10	<i>sigmoid</i>	Nadam, Adam, RMSprop
					<i>tanh</i>	
					<i>ReLU</i>	
D3	$354 \times 122$	/	2	2-112, korak 10	<i>tanh</i>	Adam
D4	$354 \times 122$	MSC	2	2-112, korak 10	<i>ReLU</i>	Nadam

viševeličinske regresijske probleme kao što su: sigmoidna, tangens-hiperbolna i *ReLU* funkcija.<sup>26</sup> Aktivacijske funkcije su važne jer unose nelinearnost u model neuronske mreže, čime omogućuju učenje složenijih uzoraka u podacima. Sigmoidna funkcija preslikava vrijednosti ulaza u raspon vrijednosti između 0 i 1, no kod njezine upotrebe može doći do pojave nestajanja gradijenta, kad gradijent funkcije postaje vrlo malen, što može usporiti učenje modela. Tangens-hiperbolna aktivacijska funkcija, s druge strane, preslikava odnosno normalizira vrijednosti izlaza neurona na raspon vrijednosti između  $-1$  i  $1$ . Na taj način kreira izlaze koji su centrirani oko nule, što može pomoći u sprječavanju problema s nestajanjem gradijenta te je mreži lakše naučiti temeljne obrasce u podacima. *ReLU* (ispravljena linearna jedinica) funkcija postavlja sve negativne ulazne vrijednosti na nulu, a ostavlja sve pozitivne vrijednosti nepromijenjene. Izbor aktivacijske funkcije ovisi o specifičnom problemu koji se rješava i o arhitekturi neuronske mreže. Uobičajena je praksa isprobati različite aktivacijske funkcije i usporediti njihovu izvedbu na skupu za validaciju modela.<sup>26</sup> Za skupove podataka D1 i D2 najprije je ispitan utjecaj broja neurona u samo jednome skrivenom sloju za pojedine funkcije. Zatim se pratio utjecaj vrste optimizatora na iznose *MSE* i  $R^2$  za odabrani broj neurona te za pojedinu funkciju. Za najbolju kombinaciju hiperparametara modela iz prethodna dva ispitivanja provjeren je utjecaj broja epoha te veličina grupe. Za najbolje modele iz skupova D1 i D2 ispitan je utjecaj dodavanja drugoga skrivenog sloja te broja neurona u oba sloja. Tako je dobiven skup D3 koji čine hiperparametri najboljeg modela skupa D1, ali s dva skrivena sloja, odnosno D4 koji čine hiperparametri najboljeg modela skupa D2 s dva skrivena sloja. Prilikom razvoja modela neuronskih mreža u *Pythonu* za sprječavanje pretreniranosti modela primijenila se metoda ranog zaustavljanja. Ta metoda omogućuje zaustavljanje učenja onda kad pogreška na skupu za učenje i dalje pada, a pogreška na skupu za provjeru dosegne minimum za zadani broj epoha. Je li razvijeni model pretreniran, pratilo se i usporedbom dobivenih kriterija srednje kvadratne pogreške i koeficijenta determinacije na skupu za razvoj i na skupu za validaciju modela.

### 3. Rezultati i rasprava

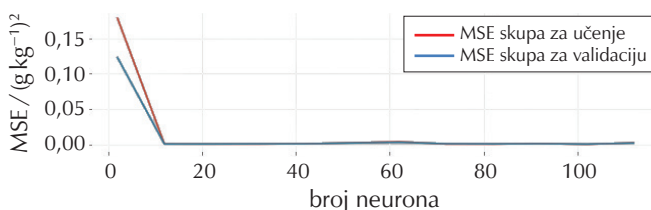
Zbog iznimno velikog broja podataka, u ovom su poglavlju grafički i tablično prikazani i protumačeni rezultati samo za D1 i D2 skup podataka uzimajući u obzir broj neurona u skrivenom sloju, vrstu aktivacijske funkcije, vrstu upotrijebljenog optimizatora te broj epoha i veličina grupe.

#### 3.1. Rezultati za skup podataka D1

Za skup podataka D1, veličine matrice  $354 \times 122$ , nije provedena predobrada podataka, već su se primijenili prikupljeni podatci s valnim brojem u rasponu  $840\text{--}1800\text{ cm}^{-1}$ .

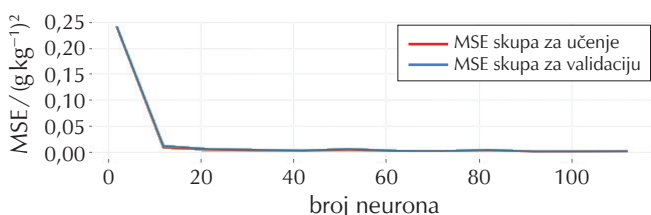
#### 3.1.1. Ispitivanje broja neurona za pojedine aktivacijske funkcije

U postupku razvoja i optimizacije modela neuronske mreže najprije se ispitaio utjecaj broja neurona u samo jednome skrivenom sloju počevši od 2 neurona te s obzirom na velik broj podataka do 112 neurona prateći pritom iznose *MSE* te  $R^2$  za pojedine aktivacijske funkcije. Rezultati za *MSE* grafički su prikazani na slikama 7, 8, i 9 te se iz njih može vidjeti da za aktivacijsku funkciju *ReLU* nema potrebe za brojem neurona u skrivenom sloju većim od 10, jer se *MSE* ne smanjuje s daljnjim povećanjem broja neurona u skrivenom sloju. Sličan je trend primijećen i kod sigmoidne aktivacijske funkcije za broj neurona veći od 20. Kod aktivacijske funkcije *tanh* vide se oscilacije u vrijednostima *MSE* kako se broj neurona u skrivenom sloju povećava. Za daljnja ispitivanja za *tanh* funkciju kao optimalan broj neurona uzet je broj 30, jer je za njega dobivena vrlo mala pogreška *MSE*, no raspon vrijednosti 10 – 20 također je zadovoljavajući.



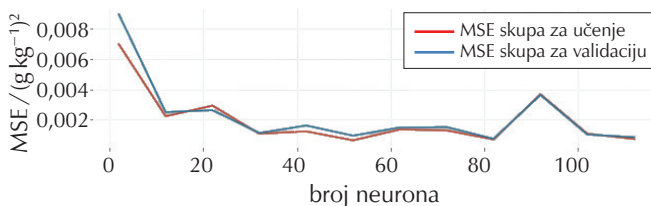
Slika 7 – *MSE* u ovisnosti o broju neurona u skrivenom sloju za *ReLU* aktivacijsku funkciju

Fig. 7 – *MSE* in dependence of number of neurons in hidden layer for *ReLU* activation function



Slika 8 – *MSE* u ovisnosti o broju neurona u skrivenom sloju za sigmoidnu aktivacijsku funkciju

Fig. 8 – *MSE* in dependence of number of neurons in hidden layer for sigmoid activation function



Slika 9 – *MSE* u ovisnosti o broju neurona u skrivenom sloju za *tanh* aktivacijsku funkciju

Fig. 9 – *MSE* in dependence of number of neurons in hidden layer for *tanh* activation function



### 3.1.2. Ispitivanje utjecaja vrste optimizatora

U tablici 2 prikazani su iznosi MSE i  $R^2$  na skupu za učenje i skupu za provjeru za ispitane optimizatore te aktivacijske funkcije. Iz tablice je vidljivo da su za sve tri aktivacijske funkcije najniže vrijednosti MSE skupa za učenje i provjeru, a analogno najviše vrijednosti  $R^2$  postižu se uporabom ADAM, NADAM te RMSprop optimizatorima. Za funkciju ReLU najbolje rezultate daje NADAM, za sigmoidnu RMSprop, a za tanh ADAM. U ovome radu, najbolje predikcije neuronskih mreža postignute su kombinacijom tanh aktivacijske funkcije s 30 neurona u skrivenom sloju i NADAM optimizatorom, što je rezultiralo najnižim vrijednostima srednje kvadratne pogreške i najvišim vrijednostima koeficijenta determinacije među razvijenim modelima.

Tablica 2 – MSE i  $R^2$  u ovisnosti o pojedinome optimizatoru i aktivacijskoj funkciji

Table 2 – MSE and  $R^2$  in dependence of the optimiser and activation function

Aktivacijska funkcija Activation function	Optimizatori Optimisers	Skup za učenje Training data		Skup za provjeru Validation data	
		MSE	$R^2$	MSE	$R^2$
ReLU	ADAM	0,0049	0,9952	0,0061	0,9932
	NADAM	0,0031	0,9970	0,0023	0,9974
	RMSprop	0,0102	0,9901	0,0137	0,9846
sigmoid	ADAM	0,0079	0,9923	0,0118	0,9868
	NADAM	0,0084	0,9918	0,0096	0,9893
	RMSprop	0,0058	0,9943	0,0057	0,9936
tanh	ADAM	0,0025	0,9975	0,0035	0,9961
	NADAM	0,0011	0,9989	0,0014	0,9983
	RMSprop	0,0166	0,9837	0,0199	0,9776

### 3.1.3. Ispitivanje utjecaja broja epoha i veličine grupe

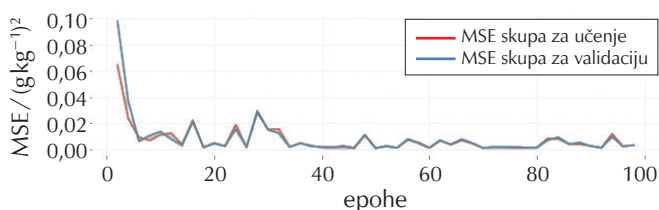
Nakon što je pronađena najbolja kombinacija aktivacijske funkcije, broja neurona u skrivenom sloju te vrsta optimizatora, ispita se utjecaj broja epoha i veličina grupe. Iz prethodnih ispitivanja, za skup podataka D1, to je model koji sadrži tangens hiperbolnu funkciju te 30 neurona u skrivenom sloju i NADAM optimizator. Na slikama 10 i 11 prikazan je utjecaj broja epoha i veličina grupe na iznos

Tablica 3 – Najbolji model neuronske mreže za skup podataka D1

Table 3 – Best neural network model for D1 data set

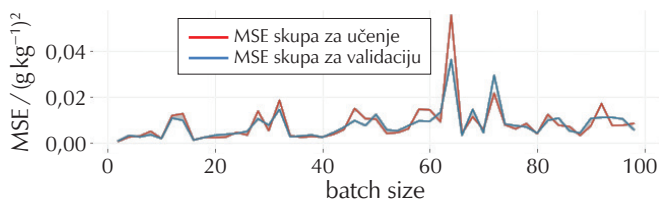
Broj neurona u skrivenom sloju Number of neurons in the hidden layer	Aktivacijska funkcija – skriveni sloj Activation function in the hidden layer	Aktivacijska funkcija – izlazni sloj Activation function in the output layer	Loss funkcija Loss function	Algoritam učenja Training algorithm	Broj epoha Number of epochs	Veličina grupe Batch size
30	tanh	LINEAR	MSE	NADAM	38	6

kriterija uspješnosti, MSE. Iznos srednje kvadratne pogreške najprije naglo pada, uz poneka odstupanja između 5. i 40. epoha. Daljnjim povećanjem broja epoha iznad 40 nema većeg utjecaja na kriterije uspješnosti na oba skupa podataka. Trend je sličan i za veličinu grupe, ali uz povećanje srednje kvadratne pogreške s povećanjem veličine grupe iznad 40. Za izradu modela odabrana je kombinacija od 38 epoha i veličine grupe 6.



Slika 10 – MSE u ovisnosti o broju epoha

Fig. 10 – MSE in dependence of the number of epochs



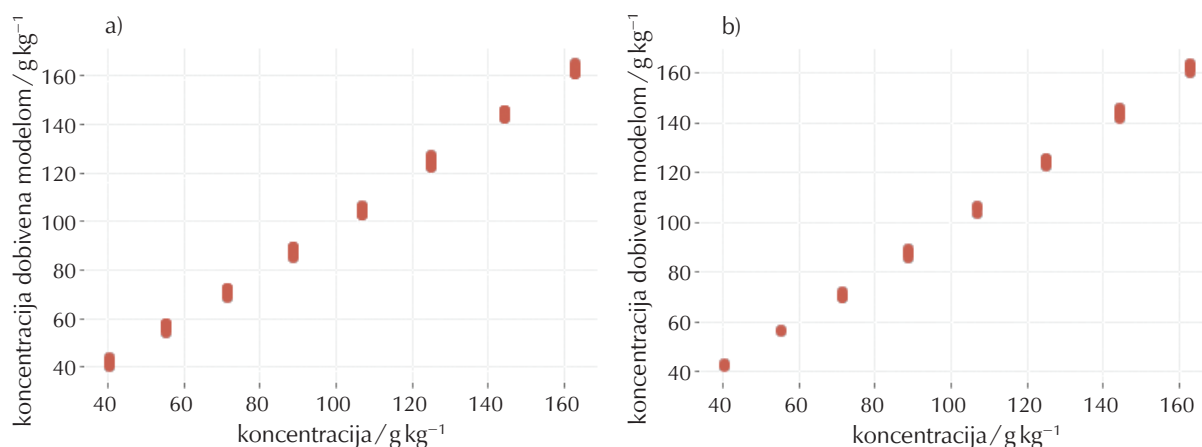
Slika 11 – MSE u ovisnosti o veličini grupe

Fig. 11 – MSE in dependence of the batch size

### 3.1.4. Najbolji model za skup podataka D1

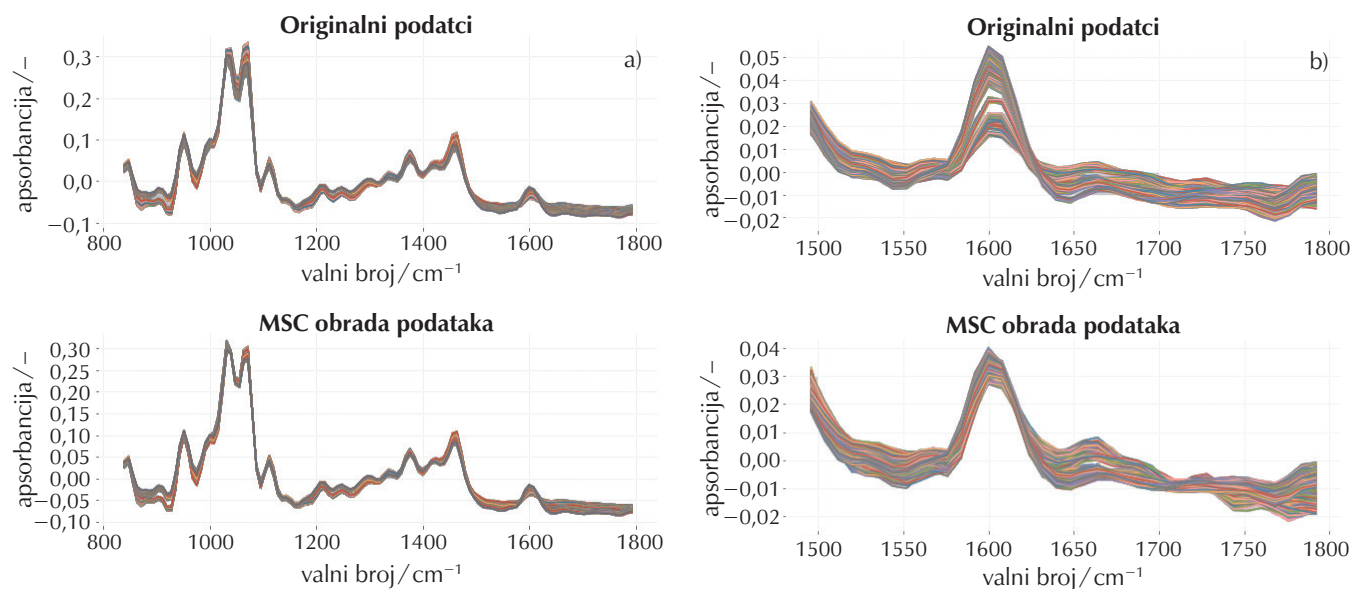
Na temelju svih prethodnih ispitivanja dobiven je najbolji model neuronske mreže, čiji su hiperparametri prikazani u tablici 3.

Na slici 12 nalaze se grafički prikazi usporedbe stvarnih koncentracija te onih dobivenih modelom neuronske mreže. Linearni trend i malo rasipanje točaka oko pravca potvrđuju prikladnost tog neuronskog modela za skup D1. Prikladnost modela također dokazuju iznimno visok i sličan iznosi koeficijenta determinacije, kako na skupu za razvoj modela tako i na skupu za vrednovanje modela, koji su vrlo blizu jedinici, kao i vrlo male vrijednosti srednje kvadratne pogreške koje su vrlo blizu nuli (tablica 6).



Slika 12 – Usporedba stvarnih koncentracija i onih dobivenih NN modelom za skup podataka D1 na a) skupu podataka za učenje i b) skupu podataka za provjeru

Fig. 12 – Comparison of real and predicted concentrations of neural network model for D1 data set on a) training, and b) validation data set



Slika 13 – Usporedba spektra s originalnim podatcima i podatcima obrađenim MSC tehnikom s rasponom valnih brojeva a) 840 – 1800 cm<sup>-1</sup> i b) uvećan prikaz 1500 – 1800 cm<sup>-1</sup>

Fig. 13 – Comparison of original spectra and the one preprocessed with MSC technique with the span of wave number a) 840–1800 cm<sup>-1</sup>, and b) zoom view 1500–1800 cm<sup>-1</sup>

### 3.2. Rezultati za skup podataka D2

Skup podataka D2, veličine matrice 354 × 122, predobrađen je tehnikom MSC da bi se smanjila raspršenja u spektralnim podatcima, čime bi se eventualno pospješila uspješnost modela neuronske mreže. Na slici 13 prikazana je usporedba spektra bez MSC obrade i s njom. Osim modela neuronske mreže, razvijen je i PLSR model koji se radi većeg broja modela razvijao samo za D2 skup (predobrađeni podatci).

#### 3.2.1. Ispitivanje broja neurona za pojedine aktivacijske funkcije

Za aktivacijske funkcije *ReLU* i *sigmoid* optimalni broj neurona u skrivenom sloju je 10, jer daljnjim povećanjem broja neurona u skrivenom sloju nema promjena u iznosima *MSE* i *R*<sup>2</sup>. Za funkciju *tanh* uzet je broj od 20 neurona u skrivenom sloju, jer daljnjim povećanjem njihova broja dolazi do povećanja iznosa *MSE*.



### 3.2.2. Ispitivanje utjecaja vrste optimizatora

U tablici 4 prikazani su dobiveni rezultati pojedinih aktivacijskih funkcija s određenim brojem neurona u skrivenom sloju s obzirom na vrstu primijenjenog optimizatora. Za aktivacijsku funkciju *ReLU* najbolji rezultat dobiven je s NADAM optimizatorom, što je vidljivo iz najnižih vrijednosti *MSE* te najviših vrijednosti koeficijenta determinacije  $R^2$ . Za sigmoidnu i *tanh* aktivacijsku funkciju dobiveni su vrlo slični rezultati uporabom ADAM optimizatora.

### 3.2.3. Ispitivanje utjecaja broja epoha i veličina grupe-a

Nakon što je pronađena najbolja kombinacija aktivacijske funkcije, *ReLU*, broja neurona u skrivenom sloju, 10, te vrste optimizatora, NADAM, ispitao se utjecaj broja epoha te veličina grupe. Povećanjem broja epoha od 0 do 20 naglo opada iznos *MSE*, no daljnjim povećanjem broja epoha dolazi samo do manjih promjena u iznosu *MSE*. Za izradu modela uzeto je 40 epoha, jer je za njega dobivena najmanja *MSE*. Povećanjem veličine grupe u rasponu od 0 do 20 pada i iznos *MSE*, a daljnjim povećanjem veličine grupe dolazi do oscilacija u iznosima *MSE* i  $R^2$ . Za izradu modela odabrana je veličina grupe 10 zbog najvećega dobivenog  $R^2$  te najmanje *MSE*.

### 3.2.4. Najbolji model za skup podataka D2

Nakon provedenih ispitivanja dobiven je model za skup podataka D2 s hiperparametrima prikazanim u tablici 5.

Tablica 5 – Najbolji model neuronske mreže za skup podataka D2

Table 5 – Best neural network model for D2 data set

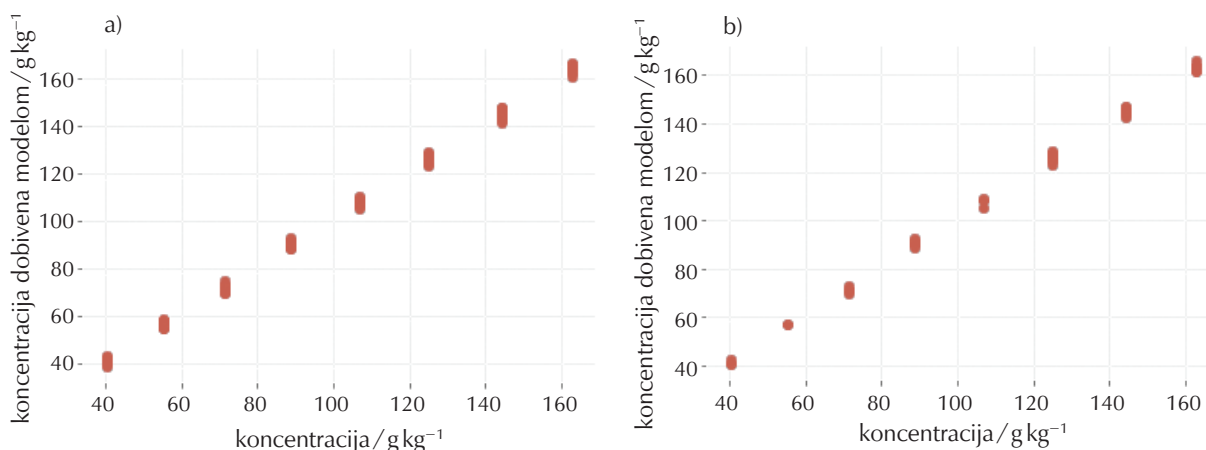
Broj neurona u skrivenom sloju Number of neurons in the hidden layer	Aktivacijska funkcija – skriveni sloj Activation function in the hidden layer	Aktivacijska funkcija – izlazni sloj Activation function in the output layer	Loss funkcija Loss function	Algoritam učenja Training algorithm	Broj epoha Number of epochs	Veličina grupe
10	<i>ReLU</i>	<i>LINEAR</i>	<i>MSE</i>	NADAM	40	10

Tablica 4 – *MSE* i  $R^2$  u ovisnosti o pojedinom optimizatoru i aktivacijskoj funkciji

Table 4 – *MSE* and  $R^2$  in dependence of the optimiser and activation function

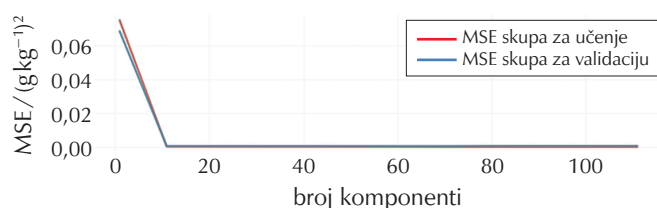
Aktivacijska funkcija Activation function	Optimizatori Optimisers	Skup za učenje Training data		Skup za provjeru Validation data	
		<i>MSE</i>	$R^2$	<i>MSE</i>	$R^2$
<i>ReLU</i>	ADAM	0,0062	0,9940	0,0079	0,9912
	NADAM	0,0011	0,9989	0,0012	0,9987
	RMSprop	0,0040	0,9961	0,0038	0,9958
<i>sigmoid</i>	ADAM	0,0064	0,9937	0,0078	0,9912
	NADAM	0,0305	0,9702	0,0421	0,9527
	RMSprop	0,0085	0,9917	0,0069	0,9922
<i>tanh</i>	ADAM	0,0013	0,9987	0,0017	0,9981
	NADAM	0,0033	0,9968	0,0037	0,9959
	RMSprop	0,0023	0,9978	0,0027	0,9970

Usporedbe stvarnih koncentracija i onih previđenih modelom prikazane su na slici 14 na skupu za učenje i skupu za provjeru. Ponovno se uočava izrazito linearni trend što, uz kriterije pogreške i visoke koeficijente determinacije, potvrđuje dobro slaganje modela i stvarnih koncentracija.



Slika 14 – Usporedba stvarnih koncentracija i onih dobivenih NN modelom za skup podataka D2 na a) skupu podataka za učenje i b) skupu podataka za provjeru

Fig. 14 – Comparison of real and predicted concentrations with the NN model for D2 data set on a) training, and b) validation data set



Slika 15 – Ovisnost MSE pogreške od PLSR modela u ovisnosti o broju komponenti

Fig. 15 – MSE of PLSR model in dependence of the number of PLSR components

Na slici 15 može se vidjeti ovisnost MSE pogreške kod PLSR modela o broju komponenata (PLS faktora), te je zaključeno da se za visoku točnost modela ne treba razmatrati više od 10 PLSR komponenti.

### 3.6. Usporedba najboljih modela

U tablici 6 prikazani su kriteriji uspješnosti najboljih modela neuronske mreže te PLSR modela za sva četiri skupa podataka na kojima su provedena ispitivanja. Koeficijent determinacije na svim je skupovima podataka dosta visok, u rasponu vrijednosti od 0,9979 do 0,9995, dok je srednja kvadratna pogreška od 0,0020 do 0,0005. Vrijednosti kriterija uspješnosti na skupu za učenje i na skupu za validaciju vrlo su slične, što potvrđuje dobre generalizacijske sposobnosti razvijenih modela. Na sva četiri skupa dobiveno je odlično slaganje stvarnih koncentracija te onih određenih modelom. Iz toga razloga nema potrebe za razvijanjem kompleksnijeg modela neuronske mreže od onoga s jednim skrivenim slojem. Također, predobrada podataka u ovoj studiji nije imala velik utjecaj na konačne rezultate. Uspoređujući rezultate za skup podataka D2 za model neuronske mreže te PLSR model, može se primijetiti da nema velike razlike u kriterijima uspješnosti.

Tablica 6 – Kriteriji uspješnosti najboljih modela neuronske mreže i PLSR modela za sve skupove podataka

Table 6 – Performance criteria for the best neural network models and PLSR model for all data sets

Skup podataka Data set	$R^2$ skupa za učenje	$R^2$ skupa za provjeru	$MSE$ skupa za učenje	$MSE$ skupa za provjeru
D1	0,9985	0,9981	0,0015	0,0017
D2	0,9980	0,9979	0,0020	0,0019
D3	0,9989	0,9985	0,0011	0,0014
D4	0,9984	0,9980	0,0016	0,0018
D2 PLSR	0,9995	0,9990	0,0005	0,0009

## 4. Zaključak

S ciljem pronalazjenja adekvatne metode za kontinuirano praćenje koncentracije djelatne tvari, ksilometazolin hidroklorida u *n*-butanolu, u ovom je radu razvijen i ispitivan linearan i nelinearan kemometrijski model primjenom PLSR i neuronskih mreža. Ispitivanja su provedena na četiri skupa podataka, D1, D2, D3 i D4, koji su se međusobno razlikovali po predobradi ulaznih podataka te broju skrivenih slojeva u neuronskoj mreži. Razvijeni modeli PLSR metodom i modeli neuronske mreže na sva četiri skupa podataka dali su iznimno dobre rezultate za ispitivani sustav. Iznosi koeficijenta determinacije i srednje kvadratne pogreške pokazali su da predobrada podataka u ovom slučaju nema velik utjecaj na konačne rezultate. To se može pripisati preciznom mjerenju i dobro odabranome rasponu valnih brojeva u kojemu je sadržana informacija o sustavu. Također, istraživanje je pokazalo da nema potrebe za dodavanjem drugoga skrivenog sloja neuronske mreže. Način praćenja i kontrole procesa primjenom procesne analitičke tehnologije rezultirao bi efikasnijom proizvodnjom, što u konačnici farmaceutskoj industriji omogućuje brže plasiranje proizvoda željene kvalitete na tržište te manju vjerojatnost za pojavom pogreške.

### ZAHVALA

Ovaj izvorni znanstveni rad izrađen je u sklopu projekta “CrystAPC – KK.01.1.1.07.0017 – Cryst APC – Crystallization Advanced Process Control (CrystAPC)” na Fakultetu kemijskog inženjerstva i tehnologije.

### Popis kratica i simbola

#### List of abbreviations and symbols

ADAM	– adaptivna procjena momenta – adaptive moment estimation
ATR	– prigušena ukupna reflektancija – attenuated total reflectance
$b_j$	– izlazni signal neurona – output signal from neuron
CPP	– ključne procesne varijable – critical process parameters
FDA	– Agencija za hranu i lijekove Sjedinjenih Američkih Država – Food and Drug Administration
FTIR	– Fourier transformacijska infracrvena spektroskopija – Fourier transform infrared spectroscopy
NADAM	– Nesterova metoda ubrzane adaptivne procjene momenta – Nesterov-accelerated Adaptive Moment Estimation
NN	– neuronska mreža – neural network
MLR	– viševeričinska linearna regresija – multiple linear regression

MSC	– korekcija višestrukog raspršenja – multiplicative scatter correction
MSE	– srednja kvadratna pogreška – mean square error
PAT	– procesna analitička tehnologija – process analytical technology
PCA	– analiza glavnih komponenata – principal component analysis
PLSR	– parcijalne regresije metodom najmanjih kvadrata – partial least squares regression
$R^2$	– koeficijent determinacije – coefficient of determination
RMSprop	– korijen srednje kvadratne propagacije – root mean squared propagation
$T_j$	– prag osjetljivosti neurona – threshold
$w_{ij}$	– vrijednost težinskog faktora neurona $i$ povezanog s neuronom $j$ – value of weight of neuron $i$ connected with neuron $j$

## Literatura

## References

1. A. S. Myerson, D. Erdemir, A. Y. Lee, Handbook of industrial crystallization, 2019., doi: <https://doi.org/10.1017/97811139026949>.
2. URL: <https://www.fda.gov/media/71012/download> (13. 2. 2022.).
3. N. Bolf, H. Dorić, Procesna analitička tehnologija, Kem. Ind. **67** (9-10) (2018) 466–469.
4. URL: [https://en.wikipedia.org/wiki/Fouriertransform\\_infrared\\_spectroscopy#Attenuated\\_total\\_reflectance](https://en.wikipedia.org/wiki/Fouriertransform_infrared_spectroscopy#Attenuated_total_reflectance) (20. 2. 2022.).
5. D. D. Dunuwila, L. B. Carroll, K. A. Berglund, An investigation of the applicability of attenuated total reflection infrared spectroscopy for measurement of solubility and supersaturation of aqueous citric acid solutions, J. Cryst. Growth **137** (1994) 561–568, doi: [https://doi.org/10.1016/0022-0248\(94\)90999-7](https://doi.org/10.1016/0022-0248(94)90999-7).
6. D. D. Dunuwila, K. A. Berglund, ATR FTIR spectroscopy for in situ measurement of supersaturation, J. Cryst. Growth **179** (1997) 185–193, doi: [https://doi.org/10.1016/S0022-0248\(97\)00119-X](https://doi.org/10.1016/S0022-0248(97)00119-X).
7. F. Zhang, T. Liu, X. Z. Wang, J. Liu, X. Jiang, Comparative study on ATR-FTIR calibration models for monitoring solution concentration in cooling crystallization, J. Cryst. Growth **459** (2017) 50–55, doi: <https://doi.org/10.1016/j.jcrysgro.2016.11.064>.
8. M. Barrett, M. McNamara, H. Hao, P. Barrett, B. Glennon, Supersaturation tracking for the development, optimization and control of crystallization processes, Chem. Eng. Res. Des. **88** (2010) 1108–1119, doi: <https://doi.org/10.1016/j.cherd.2010.02.010>.
9. Z. Q. Yu, P. S. Chow, R. B. H. Tan, W. H. Ang, Supersaturation control in cooling polymorphic co-crystallization of caffeine and glutaric acid, Cryst. Growth Des. **11** (2011) 4525–4532, doi: <https://doi.org/10.1021/cg200745q>.
10. L. C. Lee, C. Y. Liong, A. A. Jemain, A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum, Chemometr. Intell. Lab. Syst. **163** (2017) 64–75, doi: <https://doi.org/10.1016/j.chemolab.2017.02.008>.
11. F. Zhang, K. Du, L. Guo, Q. Xu, B. Shan, Comparative Study of Preprocessing on an ATR-FTIR Calibration Model for In Situ Monitoring of Solution Concentration in Cooling Crystallization, Chem. Eng. Technol. **44** (2021) 2279–2289, doi: <https://doi.org/10.1002/ceat.202100104>.
12. F. Lewiner, J. P. Klein, F. Puel, G. Févotte, On-line ATR FTIR measurement of supersaturation during solution crystallization processes. Calibration and applications on three solute/solvent systems, Chem. Eng. Sci. **56** (2001) 2069–2084, doi: [https://doi.org/10.1016/S0009-2509\(00\)00508-X](https://doi.org/10.1016/S0009-2509(00)00508-X).
13. URL: <https://wiki.anton-paar.com/en/attenuated-total-reflectance-atr/> (n.d.).
14. D. R. Baughman, Y. A. Liu, Neural Networks in Bioprocessing and Chemical Engineering, Academic Press, San Diego, SAD, 1995., str. 8–16; 21–28, doi: <https://doi.org/10.1016/B978-0-12-083030-5.50008-4>.
15. N. Bolf, Strojno učenje, Kem. Ind. **70** (9-10) (2021) 591–593.
16. S. Raschka, Python Machine Learning, Packt Publishing Ltd., Birmingham, 2015., str. 8–15.
17. Ž. Ujević Andrijić, N. Bolf, Metode umjetne inteligencije u kemijskom inženjerstvu, Umjetne neuronske mreže, Zagreb, 2021., str. 4–26.
18. S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. Syst. **58** (2001) 109–130, doi: [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1).
19. Y. Golander, W. J. DeWitte, Xylometazoline Hydrochloride, Anal. Profiles Drug Subst. Excipients. **14** (1985) 135–156, doi: [https://doi.org/10.1016/S0099-5428\(08\)60579-1](https://doi.org/10.1016/S0099-5428(08)60579-1).
20. Å. Rinnan, L. Nørgaard, F. van den Berg, J. Thygesen, R. Bro, S. Balling Engelsen, Infrared Spectroscopy for Food Quality Analysis and Control, Chapter 2 – Data Pre-processing, Elsevier, CA, SAD, 2009, str. 30–47, doi: <https://doi.org/10.1016/B978-0-12-374136-3.00002-X>.
21. R. Gautam, S. Vanga, F. Ariese, S. Umamathy, Review of multidimensional data processing approaches for Raman and infrared spectroscopy. EPJ Techn. Instrum. **2** (8) (2015) 8, doi: <https://doi.org/10.1140/epjti/s40485-015-0018-6>.
22. T. Dozat, Incorporating Nesterov Momentum into Adam, 4<sup>th</sup> International Conference on Learning Representations, ICLR, 2016.
23. Y. Tian, Y. Zhang, Z. Haibin, Recent Advances in Stochastic Gradient Descent in Deep Learning, Mathematics **11** (3) (2023) 682, doi: <https://doi.org/10.3390/math11030682>.
24. D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, 3<sup>rd</sup> International Conference on Learning Representations, ICLR, 2015.
25. S. Afaq, S. Rao, Significance of Epochs on Training a Neural Network, Int. J. Sci. Technol. Res. **9** (2020) 485–488.
26. C. Cui, T. Fearn, Modern practical convolutional neural networks for multivariate regression: Applications to NIR calibration, Chemometr. Intell. Lab. Syst. **182** (15) (2018) str. 9–20, doi: <https://doi.org/10.1016/j.chemolab.2018.07.008>.



## SUMMARY

### Application of Neural Networks for Estimating the Concentration of Active Ingredients Solution using *In-situ* ATR-FTIR Spectroscopy

Tesa Herceg,<sup>a</sup> Željka Ujević Andrijić,<sup>a\*</sup> Matea Gavran,<sup>a</sup> Josip Sacher,<sup>a</sup> Ivan Vrban,<sup>b</sup> and Nenad Bolf<sup>a</sup>

Process analytical technology (PAT) is increasingly applied in the crystallization process for continuous monitoring of some of the key process parameters and product quality features. Very important process variables for cooling crystallization are the temperature and concentration of the mother liquor. Continuous measurement of concentration is made possible by advanced *in situ* spectroscopic instruments. Attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR), used in this work belongs to such methods. The calibration model which estimates the concentration of the solution in real time can be developed using machine-learning methods. The aim of this work was to develop and analyse partial least squares regression (PLSR) and neural network models for modelling the dependence of the concentration of the active ingredients, xylometazoline hydrochloride in *n*-butanol, on temperature and spectral data obtained by measurements with an ATR-FTIR spectrometer. In this work, pre-processing of the collected data was performed with MSC technique (*multiplicative scatter correction*), *Min-Max* and *Z-score* normalization; the number of neurons in the first and second hidden layers, the number of hidden layers, the type of learning algorithm applied (ADAM, NADAM, RMSprop), and the influence of the type of transfer function (*ReLU*, *sigmoid*, *tanh*) on the quality of the developed neural networks were analysed. Considering values of coefficient of determination and mean square error, developed models gave very good results on all four datasets. The neural network model gave coefficients of determination in the range of values from 0.9979 to 0.9989, and the mean square error from 0.0020 to 0.0011. With the PLSR model, coefficients of determination from 0.9990 to 0.9995, and mean square errors from 0.0009 to 0.0005, were obtained.

Obtained results showed that the pre-processing of the data and the addition of a second hidden layer of the neural network in this study did not have a major impact on the final results. This type of monitoring and control of the process would lead to more efficient production with a lower probability of error, enabling the pharmaceutical industry to bring products to market faster.

#### Keywords

*Crystallisation, process analytical technology, ATR-FTIR spectroscopy, machine-learning, xylometazoline hydrochloride*

<sup>a</sup> University of Zagreb Faculty of Chemical Engineering and Technology, Department of Measurements and Process Control, Savska c. 16/5A, 10 000 Zagreb, Croatia

<sup>b</sup> PLIVA Croatia Ltd. (member of TEVA Group), Prilaz baruna Filipovića 25, 10 000 Zagreb, Croatia

Original scientific paper  
Received February 14, 2023  
Accepted April 21, 2023