

Transzmissziós Raman spektrometria alkalmazása a gyógyszeriparban

OTDK-dolgozat

Készítette:

Farkas István

gyógyszervegyész-mérnök M.Sc.

VI. évfolyam

Témavezető:

Dr. Marosi György

egyetemi tanár

Szerves Kémia és Technológia Tanszék

**Budapesti Műszaki és Gazdaságtudományi Egyetem
Vegyésszmérnöki és Biomérnöki Kar**

Budapest, 2012

A dolgozat bemutatását a TÁMOP - 4.2.2.B-10/1--2010-0009 "Új tehetséggondozó programok és kutatások a Műegyetem tudományos műhelyeiben" című program támogatta.

Tartalomjegyzék

1.	Bevezetés	4
2.	Szakirodalmi áttekintés.....	5
2.1.	A Raman-mikrospektrometria kialakulása.....	5
2.2.	A rezgési spektrometria elméletének rövid áttekintése a mennyiségi elemzés szempontjából....	5
2.3.	A transzmissziós Raman spektrometria bemutatása	8
2.4.	A bevonás műveletének nyomon követése spektrometriai módszerekkel	11
2.5.	Az alkalmazott sokváltozós adatelemzési (kemometriai) módszerek transzmissziós Raman spektrometriában történő felhasználásának áttekintése.....	12
3.	Felhasznált anyagok és módszerek	14
3.1.	Kísérletek során felhasznált anyagok	14
3.2.	A kísérletek során alkalmazott minta előkészítési módszerek és berendezéseik	16
3.2.1.	Minta előkészítési módszerek	16
3.2.2.	Analitikai vizsgálatok berendezései.....	17
3.3.	Alkalmazott vizsgálati körülmények.....	18
3.4.	A felhasznált előzetes adatkezelési eljárások, teljesítményjellemzők és sokváltozós regressziós módszerek ismertetése.....	19
3.4.1.	Alkalmazott adatelőkészítési módszerek	19
3.4.2.	Alkalmazott modell illeszkedésének mértékét bemutató jellemzők	20
3.4.3.	Alkalmazott regressziós módszerek	22
3.4.4.	Változó kiválasztási módszerek	25
3.4.5.	Rangszám különbségek abszolútérték-összegének módszere	26
4.	Kísérleti eredmények és értékelésük.....	27
4.1.	Aszimmetrikus legkisebb négyzetek módszerével (ALS) végzett alapvonal-korrektúra	27
4.2.	Panangin tabletták bevonásának nyomon követése.....	28
4.3.	Famotidine kvantitatív meghatározása.....	34
4.3.1.	Transzmissziós Raman spektrometriás elővizsgálatok eredményei.....	36
4.3.1.1.	Nagyobb töltöttségű kapszulák teljes spektrális tartományt felhasználó vizsgálatának eredményei.....	40
4.3.1.2.	Kisebb töltöttségű kapszulák redukált spektrális tartományt felhasználó vizsgálatának eredményei.....	42
4.3.1.3.	Kisebb töltöttségű kapszulák teljes spektrális tartományt felhasználó vizsgálatának eredményei.....	43
4.3.1.4.	Az összes mintasorozat eredményeinek összehasonlítása	46

4.3.1.5. A transzmissziós Raman vizsgálatok eredményeinek összegzése	53
4.3.2. FT-Raman spektrometriás vizsgálatok	54
4.3.3. A rangszám különbségek abszolútérték-összeg módszerének (SRD) eredményei	55
4.4. A transzmissziós és reflexiós Raman spektrometria összehasonlítása SRD módszerrel	61
5. Összefoglalás	63
6. Köszönetnyilvánítás	65
7. Függelék	66
8. Irodalomjegyzék	80

A dolgozatban előforduló rövidítések és betűszavak magyarázata:

ALS	aszimmetrikus legkisebb négyzetek módszere (<i>asymmetric least squares</i>) nem összetévesztendő az MCR-ALS-ből ismeretes ALS-sel (<i>alternate least squares</i>)
CCD	töltéscsatolt eszköz (<i>charge-coupled device</i> , Raman-spektrometriához használt detektor)
CLS	klasszikus legkisebb négyzetek módszere (<i>classical least squares</i> , keverékspektrumok modellezési eljárása a tiszta összetevők referenciaspektrumainak felhasználásával)
GA	genetikus algoritmus v. algoritmusok (<i>genetic algorithms</i> , változókiválasztási módszer)
IR	infravörös spektrometria (<i>infrared spectroscopy</i>)
MCR	többváltozós görbefulbontás (<i>multivariate curve-resolution</i>)
NIR	közeli infravörös spektrometria (<i>near-infrared spectroscopy</i>)
PAT	teljes folyamatszabályozás (<i>process analytical technology</i>) a nemzetköz gyógyszerhtóságok által preferált irányelv
PC	főkomponens (<i>principal component</i>)
PCA	főkomponens-elemzés (<i>principal component analysis</i>)
PCR	főkomponens-regresszió (<i>principal component regression</i>)
PLS	részleges legkisebb négyzetek módszere (<i>partial least squares</i> , a dolgozat ezzel jelöli az ezzel végzett regressziót is)
RMSE	közepes négyzetes hiba (<i>root mean square error</i>)
SORS	térben eltolt Raman spektrometria (<i>surface offset Raman spectroscopy</i>)
SRD	rangsámkülönbségek abszolútérték-összegének módszere (<i>sum of ranking differences</i> , matematikai eljárás módszerek, jellemzők stb. összehasonlítására)
SVM	támogatóvektor-módszer (<i>support vector machines</i> , nemlineáris gépi tanulási, ill. regressziós eljárás)
TRS	transzmissziós Raman spektrometria (<i>transmission Raman spectroscopy</i>)
XRPD	porröntgen diffrakció (<i>X-Ray powder diffraction</i>)

1. Bevezetés

A gyógyszeripari folyamatok minden lépése során kiemelt fontosságú az alapanyagok, köztitermékek és végtermékek minőségének ellenőrzése és folyamatos biztosítása. A gyógyszerkészítmények receptúráinak beállítása, illetve a gyógyszerkészítmények minőségellenőrzése a gyártás során sem kivétel ez alól. Eme folyamatok minőségellenőrzése az analitika más területein elért eredményeihez képest gyerekcipőben jár. Az utóbbi években egyre erősebb igény tapasztalható a formulálási technológiák és a termékek gyártásának valós idejű, illetve a gyártási fázisok közötti ellenőrzésére.

A rezgési spektrometriai módszerek elterjedése várható a közeljövőben ezen a területen. A gyógyszeriparban egyre nagyobb hangsúlyt kap a teljes folyamatszabályozás (angol nevén: „Process Analytical Technology”, rövidítve: PAT) igénye. Az irányelv szerint a készítménygyártás minden lépését vizsgálat kíséri, így a folyamatok paraméterei és a termékek jellemzői pedig minden eddiginél jobban kézben tarthatóvá válnak. A dinamikusan fejlődő rezgési spektrometriai módszerek alkalmasak erre a célra.

Ez a célkitűzés a spektrometriai módszerek fejlesztésének is új lendületet adott. Egy új technika, a transzmissziós Raman spektrometria a fenti elvárásoknak megfelelően illeszthető be a gyártási folyamatokba. A technika magában foglalja a hagyományos (reflexiós) Raman spektrometria előnyeit (a minták gyors, roncsolásmentes minőségi és mennyiségi elemzése, kis koncentrációban jelen lévő szennyezők mennyiségének pontos becslése), valamint további előnyei is vannak: lehetővé válik a minta teljes térfogatának elemzése, az automatizálhatóság és így felhasználható a készítménygyártás folyamán.

Dolgozatomban a transzmissziós Raman spektrometria alkalmazhatóságnak felderítését tűztam ki célul a készítménygyártás különböző lépései során. Feladatom egy bevonási technológia időbeli lefutásának nyomon követése a bevonó készülékből vett minták spektrometriai elemzésével, valamint egy hatóanyag két polimorf módosulatából készült különböző arányú keverékeinek mennyiségi elemzése volt. Az eredmények kiértékelésére a hagyományos módszereknél pontosabb elemzést lehetővé tévő többváltozós (kemometriai) módszereket választottam. A módszerek összehasonlítását empirikus és egy új, magyar fejlesztésű nem paraméteres statisztikai módszer is szolgálta.

2. Szakirodalmi áttekintés

2.1. A Raman-mikrospektrometria kialakulása

A fotonok molekulákon bekövetkező rugalmatlan szóródásának elméleti úton történő levezetése Brillouin [1] és Smekal [2] nevéhez köthető. Ezt követte néhány év elteltével a kísérleti úton történő kimutatás (1928, Raman) [3]. A kezdetben fellendülő kutatásoknak gátat szabott a Raman-szórás kis hatáskeresztmetszete (kb. minden 10^6 - 10^8 gerjesztő fotonra jut egy „Raman foton”) A lézerek felfedezésével [4], [5] új lendületet vett a műszerfejlesztés, és 1974-ben már az első mikrospektrométer megépítésére is sor került [6], [7], amely térképfelvételek készítését is lehetővé tette.

A kezdetben (fotoelektron-sokszorozó detektor mellett) több órás mérési idő a CCD detektorok kifejlesztésével másodperces nagyságrendűvé csökkent. A rugalmasan szórt, ún. Rayleigh-fotonokat elnyelő, notch filternek nevezett féligáteresztő tükrök alkalmazásával az átvitt jelintenzitás megnövekedett és így lehetővé vált nagyon gyenge, kis intenzitású sávok detektálása is. Az ún. konfokális rendszerek kialakításával a térbeli felbontás mikrométeres nagyságrendűvé és a vizsgált mintatérfogat „vastagsága”, azaz a mélységélesség szabályozhatóvá vált. (Diszperziós készülékeknél a mélységélesség növelésével gyakran jelentős fluoreszcencia csökkenés is elérhető.)

2.2. A rezgési spektrometria elméletének rövid áttekintése a mennyiségi elemzés szempontjából

A rezgési spektrometriának az anyagok rezgési energia-átmeneteit felhasználó szerkezetvizsgálati módszerek összességét nevezzük. Ha a fény frekvenciája megegyezik az atomok közötti rezgések frekvenciájával, az adott rezgések gerjesztése révén az anyag elnyelheti a ráeső fény energiájának egy részét. Az anyagok azonban csak egy részét nyelik el a rájuk eső megfelelő frekvenciájú fénynek, a többi részét visszaverik, vagy áteresztik. A három mennyiség aránya az anyagra jellemző tulajdonság, így beszélünk abszorpció, reflexió és transzmisszió színeiről. A molekulában az atomok állandó rezgőmozgásban vannak egyensúlyi helyzetük körül, ún. normálrezgéseket végeznek. A normálrezgés olyan mozgásforma, amelynek során a molekula minden atomja azonos fázisban és azonos frekvenciával, de eltérő amplitúdóval rezeg. A rezgési spektrumot alkotó frekvencia értékek a

molekulát felépítő atomok tömegétől, geometriai elrendeződésétől és a közöttük lévő kötések erősségétől függnék. A molekularezgések gerjesztésére alkalmas fény frekvenciája a közepes infravörös tartományba esik, általában $4000\text{--}200\text{ cm}^{-1}$ közötti.

Az *infravörös (IR)* spektrometria elméletét és gyakorlatát Griffiths és de Haseth részletesen tárgyalják könyvükben [8].

A *Raman spektrometria* elméleti és gyakorlati aspektusait szintén részletesen tárgyalja több összefoglaló munka [9], [10], [11]. A Raman-spektrometria a Raman-szórás jelenségén alapul, melynek során a besugárzó fény rugalmatlan szórást szenved, és a visszavert fény a besugárzó frekvenciánál kisebb, vagy nagyobb frekvenciájú fotonokat is tartalmaz (Stokes, ill. anti-Stokes szórás). Mivel ezek részaránya rendkívül kicsiny az eredeti frekvenciájú sugárzáshoz (ún. Rayleigh-szórás) képest, a vizsgálathoz nagy energiájú monokromatikus fény szükséges. Ezt a látható vagy a közeli IR tartományba eső fényt lézerrel állítják elő. Mivel a Raman-szórás intenzitása a besugárzó fény hullámhosszának negyedik hatványával arányos, a látható tartományba eső fény alkalmazásával sokkal jobb jel/zaj viszony érhető el. Ez esetben viszont problémát jelenthet a fluoreszcencia jelensége, ami nem jelentkezik közeli infravörös gerjesztő sugárzás alkalmazása során. A fluoreszcencia jelenségének kiküszöbölésére alkalmazhatunk az UV tartományban sugárzó lézereket, ez esetben azonban kiemelten fennáll a minta roncsolódásának veszélye. Látható tartományba eső lézer esetén segíthet a hosszabb ideig történő besugárzás is.

A Raman spektrometria kiválasztási szabályai eltérnek az infravörös spektrometria kiválasztási szabályaitól. Utóbbi esetben a molekula dipólus-momentumának kell megváltozni a rezgési átmenet során, míg Raman-szórás esetén a polarizálhatósági tenzornak. Így az IR és a Raman spektrum tehát egymást kölcsönösen kiegészítő információtartalmat hordoz [12].

A Raman spektrometria nagy előnye, hogy a spektrum felvételéhez gyakorlatilag semmiféle mintaelőkészítést nem igényel.

A Raman spektrometria kvantitatív alkalmazásának alapja a következő egyenlet [8].

$$I(\nu)_R = \frac{2^3 \pi^3}{45 \cdot 3^2 \cdot c^4} * \frac{h I_L N * (\nu_0 - \nu)^4}{\mu \nu (1 - e^{-h\nu/kT})} * [45(\alpha'_a)^2 + 7(\gamma'_a)^2] \quad (2.1)$$

ahol c a fénysebesség, h a Planck-féle állandó, I_L a besugárzás intenzitása, N a fotonnal ütköző molekulák száma, ν a molekula rezgési frekvenciája (Hz), ν_0 a besugárzó lézer frekvenciája (Hz), μ a rezgő atomok redukált tömege, k a Boltzmann-állandó, T az abszolút

hőmérséklet, α_a' a polarizálhatósági tenzor izotróp tagjának nyoma, γ_a' a polarizálhatósági tenzor anizotróp tagjának második invariánsa.

Az egyenlet szerint a Raman szórás intenzitása arányos azon molekulák számával, amelyek a besugárzás által érintett mintatérfogatban helyezkednek el. Ebből következik továbbá, hogy a Raman intenzitás arányos a forrás intenzitásával, tehát az érzékenység növelhető nagyobb teljesítményű lézer alkalmazásával. Ennek határt szab a minta hőtűrő képessége: nagy lézerteljesítmény alkalmazásakor a minta degradálódhat.

A kvantitatív Raman elemzés alapjai a lineáris szuperpozíció nyugszanak. Egy keverék Raman spektruma a keveréket felépítő komponensek Raman spektrumának súlyozott átlaga. A súlyfaktorokat a Raman kereszteffektusok és a komponensek mintabeli koncentrációja befolyásolja. Figyelembe kell venni, hogy egy alkotó spektruma nem feltétlenül egyezik meg tiszta, illetve keverékállapotban, mert kémiai interakciók a spektrumot szignifikánsan változtathatják.

A lineáris szuperpozíció elvének érvényesüléséhez járul hozzá az, hogy a Raman szórás önmagával való kölcsönhatása igen csekély. Gyakorlatilag zérus annak a valószínűsége, hogy egy Raman-szórt foton egy másik, szintén Raman-szórt fotonnal ütközzön, így hullámhossza megváltozzon. A Raman-fotonok és a gerjesztő lézer fotonjai közti interferencia esélye szintén elhanyagolhatóan csekély. Ellenben a gyógyszerformálásban előfordulhat – akár szándékosan, akár kiküszöbölendő körülményként – hogy a többkomponensű rendszerekben kémiai kölcsönhatás alakul ki. A Raman-spektrometriát – és természetesen minden más analitikai módszert – tehát megfelelő értékelési módszerekkel alkalmassá kell tenni arra, hogy ezekben az esetekben is pontos mennyiségi elemzést tudjunk megvalósítani.

A meghatározandó komponensre jellemző Raman-sávok gyakran jól elkülönülnek egymástól, de általában alapvonalon helyezkednek el, illetve egyéb mérési hibák is terhelhetik (fókuszálásból eredő intenzitás-ingadozás, sávarány-ingadozások). Így az elemzés előtt szükség van a spektrumok előkezelésére.

A rezgési spektroszkópiát felhasználó módszerekkel felvett spektrumok egy- ill. többváltozós modellt alkalmazva egyaránt kiértékelhetők. Amennyiben a spektrum tartalmaz a meghatározandó komponensre specifikus, annak koncentrációjával lineárisan arányos területű (magasságú), a többi sávtól elkülönülő sávot, valamint ha a spektrum nem zajos a vizsgált régióban, alkalmazhatunk egyváltozós (nulladrendű) kalibrációt, mely a gyakorlatban a Lambert-Beer törvény megvalósítását jelenti. Amennyiben a fenti feltételek nem, vagy csak részben teljesülnek, akkor többváltozós (elsőrendű) kalibrációs módszereket kell

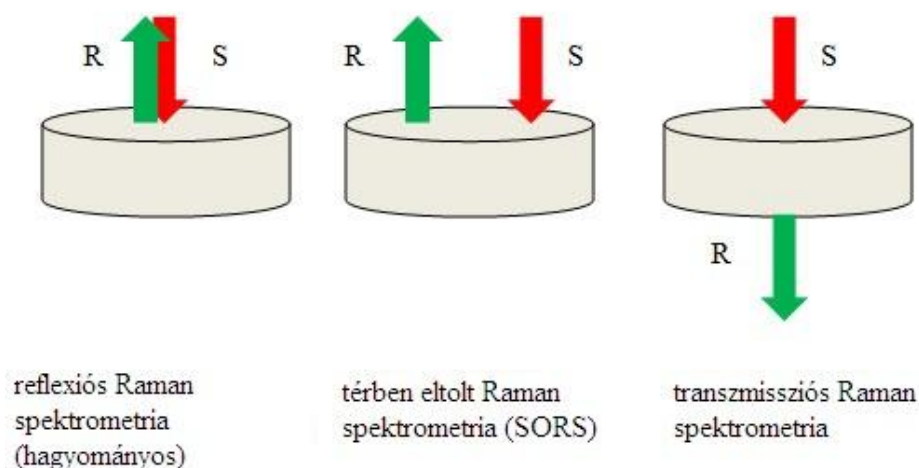
alkalmazzunk. Ilyen módszer például a klasszikus legkisebb négyzetek módszere [13] (Classical Least Squares, CLS), a főkomponens-regresszió (Principal Component Regression, PCR), a részleges legkisebb négyzetek módszere (Partial Least Squares, PLS) [16], vagy a maximális geometriai elhatárolódást támogató vektorokat kikereső módszer (a továbbiakban: támogató vektor módszer, Support Vector Machine, SVM).

2.3. A transzmissziós Raman spektrometria bemutatása

A gyógyszeriparban számos esetben szükséges a készítmények (tabletták, kapszulák) teljes térfogatának elemzése. Számos esetben (pl. kész termékek vizsgálata bliszterben, gyógyszerkészítmények roncsolásmentes minőségi és mennyiségi elemzése) követelmény a gyors, roncsolásmentes analitikai módszer. A Raman spektrometria alkalmas lehet erre a feladatra, mivel számos előnnyel rendelkezik: a kapott spektrumok könnyen értelmezhetőek, a mérést nem zavarja a víz jelenléte.

Az általánosan elterjedt reflexiós Raman spektrometria azonban nem tud a teljes mintatérfogatról pontos képet adni. A minták inhomogenitásának kiküszöbölésére a vizsgált mintatérfogat több pontjáról vesznek mintát. Ez a mérési eljárás azonban csak a minta felszínéről nyújt információt, a minta mélységi (vertikális) összetételéről pedig nem. A Raman spektrometriai módszerek fejlesztése az utóbbi években a teljes térfogat jellemzésére alkalmas irányába tolódott el [14]. Ez a törekvés vezetett el a transzmissziós Raman spektrometria megszületéséhez.

A transzmissziós Raman spektrometria egy másik újabb Raman spektrometriai mérési technikából, a térben eltolt Raman spektrometriából (spatially offset raman spectrometry, SORS) fejlődött ki. A transzmissziós Raman spektrometria elvét azonban már 1967-ben bemutatták [15]. A módszer lényege, hogy a szilárd fázisban a Raman fotonok a minta teljes térfogatában szóródnak, így a teljes mintatérfogatról reprezentatív mintavételezés történik már egy mérés alkalmával is. A technika kifejlesztését a Raman fotonok mintán belüli mozgását vizsgáló kutatások alapozták meg [16], [17]. A gyakorlatban ez lehetővé teszi azt, hogy a minta gerjesztése és a Raman fotonok detektálása egymástól térben elválasztható legyen, ezt mutatja be a 2.1. ábra.



2.1. ábra: A különféle Raman technikák működési elvét szemléltető ábra (piros: gerjesztő sugárforrás (S), zöld: Raman-fotonok (R))

A különféle Raman technikák működési elvét szemlélteti a 2.1. ábra. A legelterjedtebb a bal oldali, ún. visszaszórásos (reflexiós) Raman spektrometria. Itt a minta gerjesztése és a Raman fotonok begyűjtése a sugárforrás által gerjesztett területről történik. Ez az elrendezés egyszerűen megvalósítható, jelenleg a legelterjedtebb. Mellette látható a térben eltolt Raman spektrometriás technika elrendezése. Változás a hagyományos Raman spektrométerek elrendezéséhez képest, hogy itt a gerjesztés és a detektálás helye máshol történik a minta felületén. Ezt az elkülönülést használja ki a transzmissziós Raman technika, amely a térben eltolt Raman spektrometria extrém változatának tekinthető: a gerjesztés és a detektálás térbeli eltolása oly mértékű, hogy az a vizsgálandó minta ellentétes oldalán (!) történik. Ez megoldja a visszaszórásos Raman technika problémáját: a teljes mintatérfogatról képes (tehát nem csak a felületről, hanem vertikálisan is) Raman fotonokat begyűjteni. A nem teljeskörű mintavételezés problémáját érintőlegesen vizsgálták [18].

A technika alkalmazhatóságát vizsgálta meg Matousek és Parker [19]. A Raman fotonok mozgását szilárd fázison belül pedig Everall vizsgálta egyik cikkében [20]. Pikoszekundumos impulzusokkal bombázták a mintákat, és követték nyomon a lézer és a Raman jelek útját a szilárd fázisban. Monte Carlo modellt fejlesztettek a Raman impulzusok profiljának modellezésére. A minták különböző mélységi pontjaiból érkező Raman fotonok helyfüggését vizsgálta Matousek egy másik cikkében [21]. A minták közepéhez közletről érkező Raman jelek átlagolódását állapították meg, ezt az eredményt a szimulációk is alátámasztották.

A transzmissziós Raman spektrometria elterjedése főként a gyógyszeripar különböző területein várható, de figyelemreméltó próbálkozások történtek orvosi diagnosztikai célú felhasználására. Laboratóriumi modelleken próbálták meg a mellrák diagnosztizálását sikerrel [22], [23]. A diganosztikai gyakorlatban való felhasználás azonban még várat magára.

A transzmissziós Raman technika alkalmazásáról kevés cikk született napjainkig, azok túlnyomó többsége a készítménytechnológiai felhasználhatóságot emeli ki.

Az eddig publikált cikkek nagyobb része a készítmények hatóanyag-tartalmát vizsgálta. Matousek és Parker a kapszula mint gyógyszerforma megbontás nélküli mérésének lehetőségét mutatja be [24]. Propranolol hatóanyag-tartalmú készítmény kapszulákban történő mennyiségi elemzését mutatja be Sparén és Johansson cikke [25]. Egy meg nem nevezett hatóanyag három, a gyógyszeriparban gyakran használt segédanyag (laktóz, talkum, keményítő) melletti meghatározását mutatta be Hargreaves [26] kapszulákban. Különleges megoldást ismertetett cikkében Lee: különböző színű kapszulákba töltött ambroxol-laktóz keverékek ambroxol koncentrációját határozta meg üveg ampullákban mért kalibrációs sorozat felhasználásával [27]. Arra a következtetésre jutott, hogy a kapszulák színének a kék és a zöld színű kapszulák kivételével nincs nagy befolyásuk a hatóanyag-tartalom meghatározására. Nem szükséges tehát különböző színű kapszulák esetében külön-külön kalibráló sort készíteni. A hagyományos és transzmissziós Raman spektrometria összehasonlítását végzi el Johansson propranolol és acetil-szalicilsav mennyiségi elemzésével tablettákban és kapszulákban [28]. A két spektrometriai módszer összehasonlítását polimorfok mennyiségi elemzésében azonban a szakirodalmi adatok alapján még nem vizsgálták. Eliasson [29] és Fransson [30] pedig többféle kemometriai módszer alkalmazhatóságát hasonlította össze ezen a területen.

Hatóanyagok különböző polimorf formáinak mennyiségi elemzésével foglalkozó transzmissziós Raman spektrometriai cikk csak kettő található az irodalomban: Aina flufenénsav I-es formáját határozta meg [31], McGoverin [32] pedig ranitidine hatóanyag polimorf módosulatait határozta meg mikrokristályos cellulóz mellett.

Érdekesség, hogy a transzmissziós Raman spektrometriát alkalmazták kokristályok jellemzésére is [33], összevetve az eredményeket pordiffrakciós vizsgálatokkal, reflexiós Raman spektrometriával és differenciális pásztázó kalorimetriával.

A technika a műanyagiparban is alkalmazható: becsomagolt polietilén pelleték sűrűségét mérte meg Kim transzmissziós Raman spektrometriai úton [34].

Az eddigi eredményeket eddig mindössze két összefoglaló cikk gyűjtötte össze [35], [36]. A szakirodalomban nem találtam példát kemometriai módszerek összehasonlítására különböző polimorf formák mennyiségi elemzésében transzmissziós Raman spektrometriai mérések során. Dolgozatomban tehát megvizsgáltam, hogy a különböző kemometriai módszerek hogyan befolyásolják a polimorfok mennyiségének meghatározhatóságát transzmissziós Raman vizsgálatok esetében. Szintén nem találtam példát a szakirodalomban arra, hogy vizsgálták-e a transzmissziós és reflexiós Raman spektrometria összehasonlíthatóságát a polimorfok mennyiségének meghatározásában.

2.4. A bevonás műveletének nyomon követése spektrometriai módszerekkel

A bevonás műveletének valós időbentörténő spektrometriai nyomon követéséről igen kevés cikk számolt be. A vizsgálatok hangsúlya a bevonat vastagságának meghatározásán volt. Palou és Cruz NIR térképezéssel vizsgálta a bevonat vastagságát [37]. Torres fluoreszcens bevonat vastagságának meghatározását végezte el Raman spektrometriai úton [38]. Filmbevonat-vastagságot mért Sovány is [39]. A gyógyszeriparban elterjedten használt módszer bevonat nyomon követésére tömegnövekedésen alapul, e mérési módszert egészítette ki a bevonatok vastagságának meghatározásával a különböző tömegnövekedési értékeknél Kauffmann és társai [40].

A bevonat vastagságának szórását mérte Torres és kutatócsoportja, ők a bevonás időbeli lefolyását is vizsgálták Raman-spektrometriai úton [41]. A kioldódás és a bevonat vastagsága közötti összefüggést is vizsgálták az elmúlt időszakban [42], s a gyógyszeripari folyamatok NIR és Raman spektrometriai úton történő nyomon követését vizsgáló összefoglaló cikk erre vonatkozó cikkeket is tartalmaz [43].

A bevonás műveletének transzmissziós Raman spektrometriával történő nyomon követésére a szakirodalomban nem találtam példát. Dolgozatomban tehát megvizsgáltam, hogy a transzmissziós Raman spektrometria alkalmas-e a bevonási folyamat időbeli nyomon követésére.

2.5. Az alkalmazott sokváltozós adatelemzési (kemometriai) módszerek transzmissziós Raman spektrometriában történő felhasználásának áttekintése

A mennyiségi elemzésben sok lehetőség adódik a minta pontos összetételének meghatározására. A legegyszerűbb módszer egy összetevő meghatározására a rá jellemző rezgési sáv intenzitásának ábrázolása a koncentráció függvényében, amely még napjainkban is igen elterjedt a Raman-spektrometria és az analitika más területein is. Ezek az eljárások azonban nem minden esetben megbízhatóak. Ezen módszer alkalmazásánál alapvető követelmény, hogy a meghatározni kívánt komponensnek legyen csak rá jellemző, éles rezgési sávja.

Sok esetben ez a megközelítés nem elégséges, ezért olyan módszerekre van szükség, amelyek képesek az előbb ismertetett módszernél pontosabban (vagy ismert pontossággal) jelezni a minták koncentrációját a mintában. Ehhez ismert összetételű mintasorozatok elkészítésére, mérésére van szükség. Az így kapott adatok feldolgozásához pedig regressziós módszerek szükségesek.

A transzmissziós Raman spektrometriai vizsgálatok során a kemometriai módszereket elterjedten alkalmazzák. Szinte minden publikáció, amely nem csak egy adott komponens kimutatását tűzte ki célul, kemometriai módszereket használ fel az eredmények értékelése során. A bevonatvastagságok meghatározásakor is alkalmaznak kemometriai módszereket. Ezeket a publikációkat az előző fejezetekben már bemutattam, s a 2.1. táblázatban az általuk felhasznált regressziós módszerek szerint csoportosítottam azokat.

2.1. táblázat: A dolgozatban tárgyalt módszerekhez kapcsolódó kemometriai alkalmazások

	műanyagipar	polimorfia vizsgálatok	mennyiségi elemzés	bevonás nyomonkövetése
PLS	[34]	[31]	[25], [26], [27], [28], [29], [30], [32]	[37], [38], [40], [41], [42]
CLS			[29], [30]	[40]
MCR			[30]	
PCA			[26], [29]	

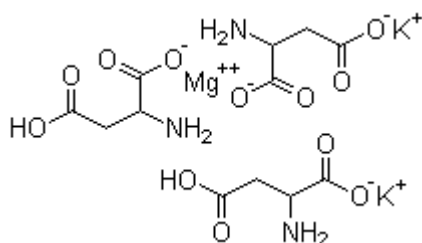
A transzmissziós Raman spektrometria szakirodalmában jelenleg a legelterjedtebb kemometriai módszer a PLS. A 2.1. táblázatban is látható, hogy a PLS módszert használták a legtöbb alkalommal. Nem találtam példát azonban a PLS-nél modernebb kemometriai módszerek alkalmazására: senki nem használt még fel SVM, illetve genetikus változó kiválasztási algoritmusokkal támogatott PLS regressziót transzmissziós Raman spektrumok kiértékelésére. Utóbbi módszerek reflexiós Raman spektrometriai vizsgálatokban történő felhasználása is ritka.

A különböző regressziós módszerek összehasonlítása a szakirodalmi példák alapján a legtöbbször egy modell-jóság jellemző alapján történik, holott a regressziós modellek összes paraméterét megvizsgálva ez sokszor nem a legjobb modell kiválasztását eredményezi. Dolgozatomban ezért a modellek összehasonlításakor az összes rendelkezésre álló modell-jóság jellemző értékelésével választottam ki a legjobb regressziós modellt. Ez a módszer azonban nem mindig adja ugyanazt az eredményt, ezért egy statisztikai alapokon nyugvó, objektív eljárást kerestem [44]. Ez a módszer a rangszám különbségek abszolútérték összegének módszere (sum of ranking differences, SRD). A szakirodalomban nem találtam példát arra vonatkozóan, hogy SRD módszert alkalmaztak volna transzmissziós Raman spektrometriai vizsgálatokban, a módszer azonban regressziós modellek összehasonlítására is alkalmas.

3. Felhasznált anyagok és módszerek

3.1. Kísérletek során felhasznált anyagok

Panangin:



Összegképlet: $C_8H_{11}KMgN_2O_8 \times C_4H_6KNO_4$
Molekulatömeg: 497,78 g/mol
Megjelenés: fehér por

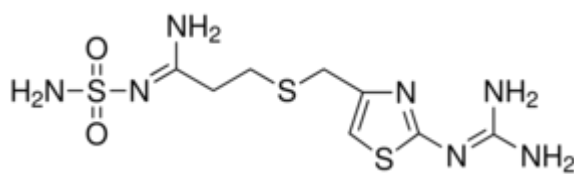
A tabletta magok összetétele:

Hatóanyag (kálium-magnézium-aszpartát), kolloid szilícium-dioxid, polividon, burgonyakeményítő, magnézium-sztearát, talkum, kukoricakeményítő, makrogol 6000, Eudragit E.

Bevonatok:

	Advantia red	Opadry red
Gyártó	Advantia	Colorcon
Összetétel	hipromelloz, poliszorbát 80, titándioxid, vörös vas-oxid, polietilén-glikol, talkum	hipromellóz, sárga vas-oxid, titán-dioxid, fekete vas-oxid, makrogol 400, vörös vas-oxid

Famotidine:



Összegképlet: $C_8H_{15}N_7O_2S_3$
Molekulatömeg: 337,45 g/mol
Megjelenés: fehér por
Sűrűség: 1,838 g/cm³
Olvadáspont: 161-163 °C

A kapszulák összetétele:

	Fehér kapszula	Átlátszó kapszula
Összetétel	Zselatin: 98%, TiO ₂ : 2%	Zselatin: 100%

A mért mintasorozatok bemutatása:

3.1. táblázat: A mért Famotidine polimorf keverék mintasorozatok jellemzői

Mintasorozat neve	Kapszula típusa	Bemért anyagmennyiség tartomány (gramm)	Mérési idő (s)	Akkumulációk száma	Mintasorozat jelölése
AtlatszoMore15sec100scan	Átlátszó zselatinkapszula	0,342-0,351	15	100	AM15s100
AtlatszoMore15sec100scan_ism	Átlátszó zselatinkapszula	0,342-0,351	15	100	AM15s100_ism
AtlatszoLess15sec100scan	Átlátszó zselatinkapszula	0,078-0,079	15	100	AL15s100
AtlatszoLess15sec100scan_ism	Átlátszó zselatinkapszula	0,078-0,079	15	100	AL15s100_ism
AtlatszoLess5sec50scan	Átlátszó zselatinkapszula	0,078-0,079	5	50	AL5s50
AtlatszoLess5sec50scan_ism	Átlátszó zselatinkapszula	0,078-0,079	5	50	AL5s50_ism
TiO2Less1sec1scan	Fehér, TiO2 tartalmú zselatinkapszula	0,078-0,079	1	1	TL1s1
TiO2Less1sec1scan_ism	Fehér, TiO2 tartalmú zselatinkapszula	0,078-0,079	1	1	TL1s1_ism
TiO2Less15sec100scan	Fehér, TiO2 tartalmú zselatinkapszula	0,078-0,079	15	100	TL15s100
TiO2Less15sec100scan_ism	Fehér, TiO2 tartalmú zselatinkapszula	0,078-0,079	15	100	TL15s100_ism
TiO2More1sec1scan	Fehér, TiO2 tartalmú zselatinkapszula	0,359-0,350	1	1	TM1s1
TiO2More1sec1scan_ism	Fehér, TiO2 tartalmú zselatinkapszula	0,349-0,350	1	1	TM1s1_ism
TiO2More15sec100scan	Fehér, TiO2 tartalmú zselatinkapszula	0,349-0,350	15	100	TM15s100
TiO2More15sec100scan_ism	Fehér, TiO2 tartalmú zselatinkapszula	0,349-0,350	15	100	TM15s100_ism
ParhuzamosBemeres	Átlátszó zselatinkapszula	0,078-0,079	1	1	PB1s1
ParhuzamosBemeres	Átlátszó zselatinkapszula	0,078-0,079	5	50	PB5s50
ParhuzamosBemeres_ism	Átlátszó zselatinkapszula	0,078-0,079	1	1	PB1s1_ism
ParhuzamosBemeres_ism	Átlátszó zselatinkapszula	0,078-0,079	5	50	PB5s50_ism

A 3.1. táblázatban szürkével kiemelt mérési sorozatokat nem értékeltem. Az AL15s100 és AL15s100_ism mintasorozatok mérése során a detektor „kiütött”, azaz a detektorba érkező fotonok száma meghaladta a maximális értéket, ezért a spektrumok további elemzésre alkalmatlanná váltak. A PB1s1 és PB1s1_ism mintasorozatok pedig nem voltak beilleszthetőek a kiértékelési rendszerbe: nem készült ilyen paraméterekkel kalibráló sorozat, amelynek robusztusságát ezekkel a mintasorozatokkal tesztelhettem volna.

3.2. A kísérletek során alkalmazott minta előkészítési módszerek és berendezéseik

3.2.1. Minta előkészítési módszerek

A tablettamagok bevonásának bemutatása:

A bevonást mindkét bevonó keverék esetében Glatt GC 250 típusú bevonó berendezésben végeztem. A berendezés felépítése a következő volt: 0,8 mm porlasztó nyílással szerelt Schlick porlasztó fej, Watson Marlow perisztaltikus pumpa és analóg szabályzó rendszer. A bevonószert (40 g) minden esetben szobahőmérsékletű vízben (260 g) állandó keverés mellett szuszpendáltam, majd 45 percig kevertettem. A berendezésbe betöltött tablettamagokat (800 g) 55 °C-os levegő befúvatásával, az üst 3 fordulat/perc forgatása mellett 45 °C kilépő levegő hőmérséklet erejéig felmelegítettem. A bevonó diszperziók porlasztását 6 g/min adagolási sebesség és 3 bar porlasztó levegőnyomás alkalmazása mellett végeztem, az üst fordulatszámát 5 fordulat/perc értékre emeltem. A bevonást 4% tömegnövekedés eléréséig végeztem.

A famotidine polimorf keverékek előállítás:

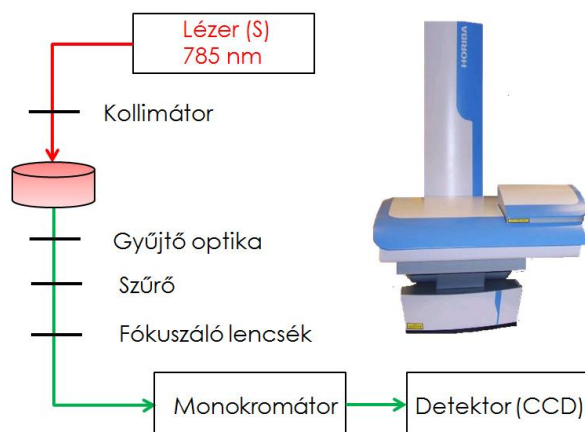
A kalibráló és validáló mintasorozat tagjait az irodalomban fellelhető eljárásához hasonlóan [45] állítottam elő: a pontosan kimért polimorf módosulatokat hexánban szuszpendáltam, majd 40 percig kevertettem, szűrtem, majd exikkátorban szárítottam.

3.2.2. Analitikai vizsgálatok berendezései

Transzmissziós Raman spektrometria:

A berendezés rövid bemutatása:

A transzmissziós Raman spektrometria széleskörben még nem terjedt el, ezért röviden bemutatom a berendezés működési elvét. A hagyományos Raman spektrometria a reflexió elvén alapul. Ennek során a mintára eső, az esetek túlnyomó többségében fókuszált lézersugár a minta felületére érkezte Raman szóródást szenved, majd visszaverődik a minta felületéről. Ezek a visszaszóródó fotonok az objektíven keresztül a detektorba jutnak. A transzmissziós Raman spektrometria ezzel ellentétben nem fókuszált, de nagyobb teljesítményű lézersugarat bocsát a mintára. A lézersugár a minta teljes belső térfogatán a szemcsék közötti pórusokon áthaladva, azokkal rugalmatlanul ütközve (Raman szórás jelensége) átjut a mintán, és a spektrális információkat tartalmazó Raman-fotonok egy gyűjtő lencserendszeren egy szűrőn majd egy újabb lencserendszeren keresztülhaladva az optikai szálba gyűjtődnek össze. Az optikai szál az optikai rácshoz vezeti a fotonokat, amely hullámszám szerint felbontja őket, ezek után a monokromátorba kerülnek a Raman-fotonok, ahonnan a detektorba jutnak. A berendezést mutatja be a 3.1. ábra.



3.1. ábra: Transzmissziós Raman spektrométer

A transzmissziós Raman spektrometriai vizsgálatokat egy Horiba Jobin Yvon AccuRA típusú transzmissziós Raman spektrométeren végeztem el. A lézersugár átmérője a minta felszínén 6 mm, míg a gerjesztő lézer hullámhossza 785 nm volt. A Raman-fotonok

detektálását egy Peltier elemmel hűtött Ge detektor végezte. A berendezés felbontása 4 cm^{-1} /pixel volt.

FT-Raman spektrometria:

Az FT-Raman vizsgálatokat egy Thermo Nicolet NXR-9650 típusú reflexiós FT-Raman spektrométeren végeztem, amely 1064 nm-es Nd-YAG lézerrel és folyékony nitrogénnel hűtött Ge detektorral volt felszerelve. A lézersugár átmérője a minta felületén $50\text{ }\mu\text{m}$ volt.

A kiértékelésben használt szoftverek:

Thermo Nicolet Omnic 7.2, Horiba Jobin Yvon LabSpec 5.76, Matlab 7.6, PLS Toolbox 6.5.2, MIA Toolbox 2.5.1

3.3. Alkalmazott vizsgálati körülmények

Panangin tabletták bevonásának nyomon követése:

Transzmissziós Raman spektrometria:

A mintákat a mintatartóba helyeztem, majd a mérési programot elindítva a berendezés automatikusan elvégezte a beállított mérési paraméterekkel az összes mérést. A spektrumok 30 másodperces mérési idővel és két akkumulációval készültek.

Famotidine hatóanyag polimorf módosulatainak mennyiségi elemzése

Transzmissziós Raman spektrometria:

A mintákat kapszulákba töltöttem, majd a mintatartóba helyezve, a mérési programot elindítva lemértem őket. A bemért anyagmennyiségek, mérési idők és akkumulációk a 3.1. táblázatban találhatóak meg.

FT-Raman spektrometria:

Minden összetételről 25 különböző mérési pontban spektrumokat vettem fel 500 mW megvilágító lézerteljesítménnyel és 4 cm^{-1} /pixel felbontással.

3.4. A felhasznált előzetes adatkezelési eljárások, teljesítményjellemzők és sokváltozós regressziós módszerek ismertetés

3.4.1. Alkalmazott adatelőkészítési módszerek

Alapvonal-korrekción:

Az alapvonal korrekciót egy Raman spektrometriában még ismeretlen, ám nagyon hatékony módszerrel: az aszimmetrikus legkisebb négyzetek módszerével végeztem (asymmetric least squares, ALS). A módszer egyszerű, de nagyon hatásos az alapvonal becslésére, egy automatikus algoritmus segítségével könnyen számolható. Az algoritmus a (3.1.) képlet alapján működik:

$$Q = \sum_i v_i (y_i - f_i)^2 + \lambda \sum_i (\Delta^2 f_i)^2 \quad (3.1)$$

ahol y a korrigálandó spektrum, f a simasági trend és v a fő súly. A v értékei mindenhol egynek adódnak azokon a helyeken ahol az y megfigyeltlen vagy megengedetten befolyásolhatja f értékét. Minden más esetben a v értéke 0, ott az f értékére simára interpolált értékeket számol ki az algoritmus. A pozitív λ paraméter állítja be az érdességi „büntetést”, minél nagyobb a λ értéke, annál simább lesz az f .

Normálás

A normálás többféleképpen lehetséges: egységnyi abszolútérték-összegre (vagyis görbe alatti területre) és egységnyi négyzetösszegre. Dolgozatomban egy harmadik változatot használtam, a Minkowski-távolságot, ennek matematikai megvalósítása látható a (3.2) képletben

$$n_x = \left(\sum_{i=1}^N |x_i^p| \right)^{1/p} \quad (3.2)$$

ahol N a mátrix sorainak száma, p a normálás típusa (jelen esetben 2), n_x a normáló függvény. A normált x_n vektort az $x_n = x/n_x$ művelet állítja elő.

Differenciálás

A differenciálás a Savitzky-Golay féle polinom illesztési módszer segítségével történt. Az algoritmus a mátrix soraira történő polinom illesztése után differenciálja a kapott függvényt. A polinom illesztésekor több paraméter is módosítható: az adott pont környezetében lévő pontok száma, amelyek részt vesznek a polinom illesztésében, illetve az illesztett polinom fokszáma. Az alkalmazott paraméterek minden spektrum esetében azonosak voltak: az adott

mérési pont környezetében lévő pontok száma 15, az illesztett polinom optimális fokszáma előzetes kísérletek alapján 2 volt. Minden spektrumot elsőrendű differenciálásnak vettem alá.

Skálázás:

A módszer során léptékváltás történik nulla középére és egységnyi szórásra. A formulát mutatja be a (3.3) képlet.

$$x_{ij}^{scal.} = \frac{(x_{ij} - \bar{x}_j)}{s_j}; s_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}} \quad (3.3)$$

ahol X_{ij} az adatmátrix, $\mathbf{x}_j$ pedig az adatmátrix egy oszlopa, a tulajdonságvektor, az s_j pedig az adott tulajdonságvektor elemeinek tapasztalati szórása. A standardizálás konstans eltolást és zsugorítást vagy nyújtást jelent az eredeti skálán oly módon, hogy a tulajdonságvektorok elemeinek számtani közepe 0, szórása pedig 1 lesz. A standardizált változó dimenziómentessé válik. Ebben az esetben is szükséges mind az $\mathbf{X}$, mind az $\mathbf{Y}$ mátrix előzetes adatkezelése.

3.4.2. Alkalmazott modell illeszkedésének mértékét bemutató jellemzők

Négyzetes középhibák számítása

A kalibráció négyzetes középhibája (Root-mean-square error of calibration, RMSEC)

Az RMSEC értéket a következő egyenlet definiálja:

$$RMSEC = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3.4)$$

ahol n a kalibrációban résztvevő minták száma, $\hat{y}_i$ az i -edik minta meghatározandó jellemzőjének (pl. koncentrációjának) valódi értéke, y_i pedig a jellemző (pl. koncentráció) regressziós módszerrel becsült értéke az i -edik minta esetében. Ebben az esetben a modell felépítéséhez minden egyes mérési pontot felhasználunk. Az RMSEC értéke megmutatja, hogy a modell mennyire illeszkedik a mért adatokhoz. Minél kisebb ez a szám, annál jobban illeszkedik a modell. Természetesen a képlet nemcsak koncentrációkra érvényes, hanem bármilyen meghatározandó fizikai mennyiségre.

A keresztellenőrzés egyelem-kihagyásos négyzetes középhibája (Root-mean-square error of cross validation, RMSECV)

Az RMSECV értéke hasonló módon, a (3.4) egyenlettel számítható, mint az RMSEC értéke, de ekkor minden egyes y_i változót úgy becslünk meg, hogy a hozzá tartozó valódi $\hat{y}_i$ értéket kihagyjuk a regresszióból a keresztellenőrzés során.

Az előbecslés négyzetes középhibája (Root-mean-square error of prediction)

Az egyenlet megegyezik az RMSEC egyenletével (3.4), a különbség itt annyi, hogy az y_i becsült értékek (pl. koncentrációk) számítása egy, a kalibráló mátrixtól különböző ellenőrző (validáló) mátrixban szereplő mérési adatok alapján történik, amelyek nem vettek részt sem a modell felépítésében, sem a keresztellenőrzésben.

Tengelymetszet

A számított mennyiségeket a valódi értékeik függvényében ábrázolva ideális esetben egy 0 tengelymetszetű 45°-os egyenest kapnánk. Ennek megfelelően az adott modell esetében tapasztalt meredekség illetve tengelymetszet ideálistól való eltérésének nagysága is információval szolgál a modell jóságáról. A két tényező közül a dolgozatomban főként a tengelymetszetet kísérttem figyelemmel, mert érzékenyebben mutatja az ideálistól való eltérést.

Determinációs együttható (R^2)

A determinációs együtthatót a (3.5) egyenlet definiálja:

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (3.5)$$

ahol $\hat{y}_i$ a mért érték, y_i a regresszióval becsült értéke, $\bar{y}$ pedig a mérendő változó valódi értékeinek átlaga. Ennek megfelelően az R^2 érték megmutatja, hogy a mért y_i értékek varianciájából mekkora részt magyaráz meg a regressziós becslések varianciája, azaz a regressziós becslés milyen mértékben írja le a mérendő változó viselkedését. Értéke 0 és 1 közötti, minél nagyobb az értéke, annál jobb a regresszió. Ez a modell-jósági paraméter is számítható a kalibráció (R_C^2), a keresztellenőrzés (R_{cv}^2) és az előrebecslés (predikció, R_{pred}^2) esetére egyaránt.

3.4.3. Alkalmazott regressziós módszerek

Egyváltozós lineáris regresszió [46] (egyváltozós legkisebb négyzetek módszere)

Az egyváltozós lineáris regresszió a leggyakrabban alkalmazott regressziós módszer a mennyiségi elemzésben. Az összefüggés alapja:

$$y = \alpha + \beta(x - \bar{x}) + \varepsilon,$$

$$\text{ahol } \bar{x} = \frac{\sum x_i}{n} \quad (3.6)$$

ahol α és β a becsülendő paraméterek, α az egyenes tengelymetszete, β a meredeksége, ε pedig a hiba. A módszer leírása megtalálható a szakirodalomban, ezért részletesebb ismertetésétől eltekintek.

Klasszikus legkisebb négyzetek módszerével végzett regresszió [49]

A módszert legjobban úgy lehet szemléltetni, ha a Lambert-Beer törvény egyenletét vesszük, amely a mért értékeket mátrixokba foglalva tartalmazza a (3.7) egyenlet:

$$\mathbf{X} = \mathbf{K} \mathbf{C} + \mathbf{E} \quad (3.7)$$

ahol az $\mathbf{X}$ ($\lambda \times n$) a mért intenzitásokat tartalmazó oszlopvektor (a mért spektrum), $\mathbf{C}$ ($p \times n$) a keveréket alkotó komponensek koncentrációit tartalmazó mátrix, a $\mathbf{K}$ ($\lambda \times p$) mátrix pedig oszlopaiban az egyes tiszta komponensek intenzitás koefficienseit (azaz spektrumait) tartalmazza. Az $\mathbf{E}$ ($\lambda \times n$) mátrix a reziduális hibákat mutatja. Ez utóbbi a számítások során elhanyagolásra kerül.

Vesszük a koncentrációkat tartalmazó mátrixot $\mathbf{C}$ ($p \times n$) és egy a mért adatokat tartalmazó $\mathbf{X}$ ($\lambda \times n$) mátrixot, majd felírjuk a fenti egyenletet (3.7), és megoldjuk a tiszta komponensek egyikének spektrumát tartalmazó mátrixra. A célunk egy olyan $\mathbf{K}$ ($\lambda \times p$) megtalálása, amely esetében a hibák aszimmetrikus négyzetének összege minimális. A módszer részletesebb ismertetése megtalálható a szakirodalomban [49].

Főkomponens regresszió (Principal Component Regression, PCR)

A PCR a főkomponens analízis módszerén alapul, ezért bemutatását ennek ismertetésével kezdem. Az adatokat az eredeti változótér helyett egy olyan ortogonális térre szűkíti le, ahol az új változók (főkomponensek, principal component, PC) a lehető legjobban leírják az eredeti adathalmaz varianciáját. Ez a gyakorlatban azt jelenti, hogy az első néhány változó írja

le az adathalmaz varianciájának a legnagyobb részét, tehát ez hordozza a legtöbb információt az eredeti adathalmazról. Ennek következtében ezek a változók elválaszthatóvá válnak a többi (újonnan létrehozott) változótól. Erre matematikailag többféle lehetőség is adódik, az egyik legegyszerűbb megoldás az adatmátrix szingulárisérték-felbontása. Ezt mutatja be az alábbi egyenlet:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T + \mathbf{E} \quad (3.8)$$

Ahol $\mathbf{U}$ ($r \times r$) az ún. bal oldali, $\mathbf{V}$ ($k \times k$) az ún. jobboldali sajátvektor-mátrix, $\mathbf{\Sigma}$ ($r \times k$) pedig a szinguláris értékeket a főátlójában tartalmazó mátrix. Ezek a sajátvektorok ortonormáltak: egységnyi hosszúak, és merőlegesek egymásra. A felhasználásra kerülő főkomponensek számának meghatározása történhet a szinguláris értékek, vagy a főkomponensek által leírt variancia elemzésével. Utóbbi vizuálisan is megvalósítható, a fenti értékek ábrázolásával a főkomponensek sorszáma függvényében. Ezt az ábrázolásmódot hegymelés-ábrának [47] (scree-plot) nevezik. Egyéb statisztikai módszerek is ismeretesek a főkomponensek ideális számának meghatározására [48], [49].

A főkomponens-regresszió lényege, hogy az adathalmazt először a fent tárgyalt PCA eljárással a főkomponens-térbe transzformáljuk, majd ennek a transzformált adathalmaznak a dimenzióját lecsökkentjük [50]. A regressziót az így létrejött redukált adathalmazon végezzük. Ezzel jelentősen csökken a minták matematikailag megkövetelt száma, de nagymértékben nő a kapott modell stabilitása.

Parciális legkisebb négyzetek módszere (Partial Least Squares, PLS)

A részleges legkisebb négyzetek módszere a legelterjedtebb többváltozós regressziós módszer, amely a PCR módszeren alapul. Az előbb bemutatott PCR egyik legelőnyösebb tulajdonsága, hogy a látens változók úgy vannak megválasztva, hogy az eredeti adathalmaz varianciáját a lehető legjobban leírják. A PLS módszer ezt azzal fejleszti tovább, hogy figyelembe veszi a függő változó értékeit is. Tehát a létrehozandó főkomponensek már nem az adatrendszer varianciájának maximalizálására, hanem a függő változóval vett kovariancia maximalizálására törekednek [51]. Ennek eredményeképpen a kapott főkomponensek jobban korreálnak a függő változóval, a meghatározási hiba így tovább csökken a PCR módszerhez képest. A PLS módszer elméletét Geladi és Kowalski mutatja be összefoglaló közleményükben [52].

Támogató vektor módszerrel végzett regresszió (Support Vector Machine, SVM)

A támogató vektor regresszió a gépi tanulási módszerek közé sorolható, alapvetően osztályozási feladatokra lett kifejlesztve [53], de regressziós feladatok megoldásában is sikerrel alkalmazható [54]. Az SVM módszer egyik fontos eleme az ún. magfüggvény (kernelfüggvény) alkalmazása, amellyel az adatrendszer dimenziója megnövelhető. A célunk olyan magfüggvény kiválasztása, mellyel az adatpontok elrendeződése a megnövelt dimenziójú tulajdonságtérben (feature space) kedvezőbb az eredeti tulajdonságtérnél (input space). Ez regressziós esetben azt jelenti, hogy az eredetileg nemlineáris görbe mentén elhelyezkedő pontok a megnövelt dimenziójú térben elvileg egy egyenesre esnek, ami a gyakorlatban csak a hiba miatt nem teljesül.

Az SVM másik fontos jellemzője a regressziós függvény illesztésének módja. Az illesztés célfüggvénye itt egy sajátos összetett függvény. Az eljárás sajátossága, hogy a célfüggvény egy előzetesen megadott ε tűréshatárnál kisebb hibákat egyáltalán nem bünteti, az ennél nagyobb hibákat azonban igen, a tűréshatárral vett távolsággal egyenesen arányos mértékben. A módszer elnevezését az adja, hogy a fentiekből következően a csak a tűréshatáron fekvő, illetve az azon kívül eső pontok vesznek részt a függvény illesztésében, az ezen belüli (ε sugarú „hipercsőbe” eső pontok egyáltalán nem. E tolerancia sávon kívülre, illetve a határára eső pontokba mutató vektorokat nevezik támogató vektoroknak. Az SVM egyik legnagyobb hátránya a túlillesztésre való erős hajlama, ezért a paraméterek optimálásakor kellő óvatossággal kell eljárni, és fokozottan ügyelni a megfelelő validálásra.

A túlillesztés fogalma alatt a dolgozatban a következőket értem:

1. a modell a kalibráló sorozat tagjaira igen jól illeszkedő modellt készít kis RMSEC érték), de a keresztellenőrzés jósági paramétere (RMSECV) ennél sokkal rosszabb eredményt ad, illetve a validáló halmaz jósági paramétere is nagyon eltér a kalibrálóétól (RMSEP), emellett a modell különféle R értékei között is nagy az eltérés.
2. a modell felépítéséhez „túl sok” komponens szükséges, ami lehet például az, hogy a rendszer sajátosságainál sokkal több főkomponenst használ fel, a szakirodalomban közmegegyezés szerint optimálisnál több komponenst használ fel a modell felépítéséhez.

3.4.4. Változó kiválasztási módszerek

Genetikus algoritmusok (Genetic Algorithms, GA)

A főkomponens-elemzésen alapuló módszerek előnye, hogy a regresszió során felhasznált főkomponensekben azok a hullámszám-tartományok vesznek részt leginkább, amelyek valós kémiai információt hordoznak. A zajt és más zavaró tényezőket leíró hullámszám-csatornák a regresszióban csak kis súllyal vesznek részt. Amennyiben számuk az adathalmazon belül nagy, hatásuk csak akkor nem fog érvényesülni, ha eltávolítjuk őket az adatsorból. Ez megtehető manuálisan, ahol a felhasználó jelöli ki a használni kívánt hullámszám-tartományokat.

A felhasználó által végzett változó-kiválasztás szubjektív és nehezen optimalizálható. Ellenőrzött körülmények között megvalósítható, illetve optimalizálható változó-kiválasztási algoritmusra számos példa akad [55]. A rezgési spektrometriában főleg az intervallum-PLS, és a genetikus algoritmusokkal segített PLS regresszió (PLS-GA) alkalmazása jellemző. Munkám során az utóbbit alkalmaztam. A genetikus algoritmusok a véletlen által is befolyásolt lépések sorozatával végzik az optimalizálást [56].

A módszer neve az öröklődés mechanizmusával való erős analógiából ered. A genetikus algoritmusok a globális optimumkeresési eljárások közé sorolhatók [57], [58], a változó kiválasztás csak egy a sokféle felhasználási lehetősége közül. A változók kiválasztásának mechanizmusa az alábbiakban foglalható össze: egy több különböző „génnel” (esetünkben a kiválasztott változóintervallumok halmazával) rendelkező regressziós modellt („egyedet”) hasonlítunk össze a meghatározási hiba („életrevalóság”) alapján, melyek egy „populációt” alkotnak. Az egyedek közül „kiszelektáljuk” a rosszakat, a nagy meghatározási hibával rendelkezőket (célszerűen pl. a rosszabbik 50%-ot). Ezek után valamelyest véletlenszerű „keresztetést” hajtunk végre a fennmaradt „túlélők” között, és ezzel az eredeti létszámmra töltjük fel a populációt. A létrejött, illetve fennmaradt egyedekben ezek után mutáció is bekövetkezhet. A fenti lépések után a következő generációban ugyanezek a lépések történnek meg. A módszerrel az evolúciós tapasztalatoknak megfelelően nő az életrevalóság, azaz csökken a létrejövő modellek meghatározási hibája.

3.4.5. Rangszám különbségek abszolútérték-összegének módszere

A regressziós modellek összehasonlítására legelterjedtebb módszer a 3.4.2. fejezetben alkalmazott teljesítmény-jellemzőik összehasonlítása. Ez az esetek nagy részében szubjektív vizsgálatot jelent. A modellek objektív összehasonlíthatóságára azonban létezik egy magyar fejlesztésű, nemparaméteres statisztikai módszer, a rangszám különbségek abszolútérték-összegének módszere (sum of ranking differences, SRD) [59].

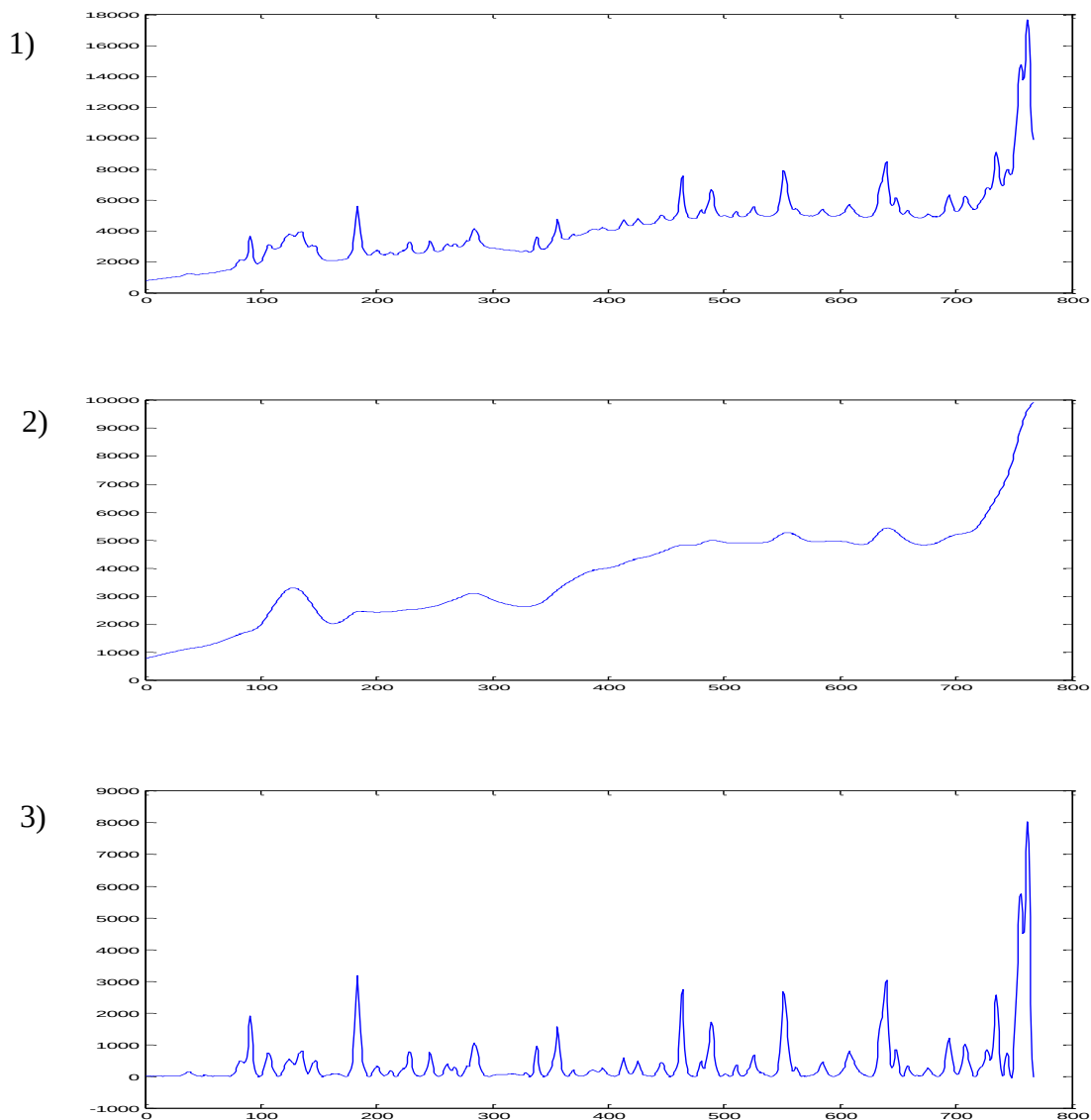
A módszer lényege, hogy a különböző minták sorba rendezhetők valamely vizsgált tulajdonságuk alapján. A mintákat többféle módszerrel vizsgálva minden módszer esetében kialakul egy sorrend. A mintákra szokás definiálni egy referencia-sorrendet, ezek után a különféle módszerekkel kialakult sorrendeket ehhez a referencia-sorrendekhez hasonlítjuk. Ez a következőképpen történik: minden egyes minta esetében vesszük az adott módszerrel kapott és a referencia sorrendben szereplő rangszám különbségét, majd ezeket összegezzük minden objektumra.

Az SRD eljárás szerint a dolgozatomban összehasonlított regressziós módszerek esetében egy módszer annál jobbnak minősül, minél kisebb a rangszám különbségek összege a referenciához képest. Abban az esetben, amikor léteznek az összehasonlítás alapját képező tulajdonság valódi értékei (esetemben ez jó közelítéssel teljesül), akkor egy (közel) egzakt referencia-sorrend is felállítható.

4. Kísérleti eredmények és értékelésük

4.1. Aszimmetrikus legkisebb négyzetek módszerével (ALS) végzett alapvonal-korrektúra

Dolgozatomban rutinszerűen alkalmaztam egy, a kromatográfiában kifejlesztett alapvonalat illesztő módszert, az aszimmetrikus legkisebb négyzetek módszerét. A módszert Raman spektrometriában még nem alkalmazták. Az ALS módszer működését mutatja be a 4.1.1. ábra.



1) a nyers spektrum 2) az ALS által illesztett alapvonal 3) A nyers spektrumból kivont alapvonal

4.1.1. ábra: Az ALS módszerrel történő alapvonal-illesztés lépései

A 4.1.1. 1) ábrán látható, hogy a berendezés által felvett spektrumok alapvonala jelentős mértékű. Erre illeszt az algoritmus egy, a 2) számmal jelölt részen látható alapvonalat, majd ezt kivonva az eredeti spektrumból megkapható a 3) alapvonal-korrigált spektrum. Bár a módszer Raman-spektrumok korrigálására is kitűnően alkalmazható, a szakirodalomban nem találtam példát a módszer spektrometriai alkalmazására.

4.2. Panangin tabletták bevonásának nyomon követése

A készítménytechnológiai folyamatok műszerezettsége az ezen a területen általánosan megkövetelt analitikai szinthez képest nem eléggé fejlett. A spektrometriai módszerek elterjedésével ez a helyzet javul, a készítménytechnológia egyre több területén kerülnek felhasználásra a rezgési spektrometriai módszerek. A bevonás műveletét többen is vizsgálták különféle módszerekkel [37], [40], [41].

A bevonás műveletének reflexiós Raman spektrometriai úton történő vizsgálata körülményes, mivel a felület több pontjáról szükséges mintát venni minden tabletta esetében, hogy a vizsgálat reprezentatívan tükrözze az adott mintán a bevonatréteg vastagságát. A reflexiós Raman vizsgálatot tovább nehezíti az, hogy a bevonás kezdetén még nem alakul ki egységes bevonatréteg a tablettamagok (a továbbiakban: korpuszok) felületén, így könnyen előfordulhat, hogy olyan pont mérése történik, ahol sok bevonó anyag van, míg a tabletta egy másik pontján készült mérés egyáltalán nem detektálja a bevonó szuszpenziót. A bevonatréteg inhomogenitásából adódó probléma térképfelvételek készítésével kiküszöbölhető, ez azonban időigényes folyamat. Így a reflexiós Raman spektrometria nehezen illeszthető be a gyártási folyamatba.

Dolgozatomban a transzmissziós Raman spektrometria alkalmazhatóságát vizsgáltam a bevonás folyamatának nyomon követésére. Előzetes várakozásaim alapján erre a technika alkalmasnak tűnik, mivel az a tabletták egész térfogatáról szolgáltat információt, így a bevonat átlagos vastagságáról is. Feltételeztem, hogy a tabletta korpusz valamely alkotórészének (célszerűen a hatóanyag) jelintenzitása a bevonat vastagságának növekedésével arányosan csökken, ebből pedig következtetni lehet a bevonás idejére, illetve a folyamat végpontjának megállapítására. A gyógyszeriparban jelenleg a bevonás folyamatának nyomon követését meghatározott számú minta átlagos tömegnövekedési értékeiből állapítják

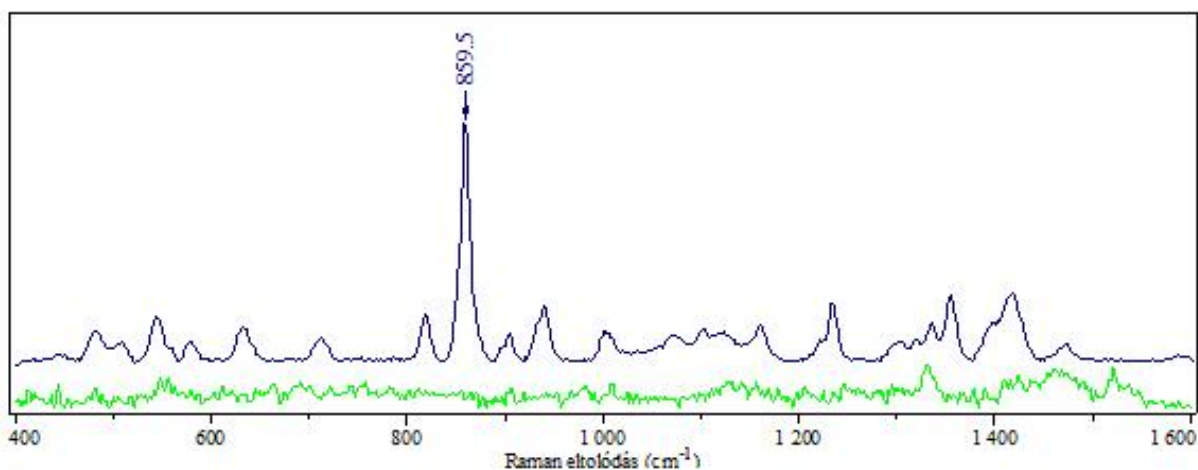
meg. A bevonás akkor ér véget, amennyiben a tabletták tömegnövekedése elér egy meghatározott szintet.

Vizsgálataim során Panangin tabletták korpuszok bevonását követtem nyomon transzmissziós Raman spektrometriai úton. Kétféle, az iparban elterjedt vasoxid-tartalmú bevonó keverékkel (AdvantiaRED, illetve OpadryRED bevonatokkal) vontam be a tablettákat, majd a 4.2.1. táblázatban megjelölt időpontokban 7 darab mintát vettem a bevonó üstből. A mintákat 30 másodperces mérési idővel és két akkumulációval mértem meg.

4.2.1. táblázat: A mintavételek időpontja a Panangin korpuszok különböző bevonatokkal történő bevonása során

AdvantiaRED	0 min	11 min	20 min	30 min	40 min	50 min	60 min	62 min
OpadryRED	0 min	15 min	25 min	35 min	45 min	55 min		

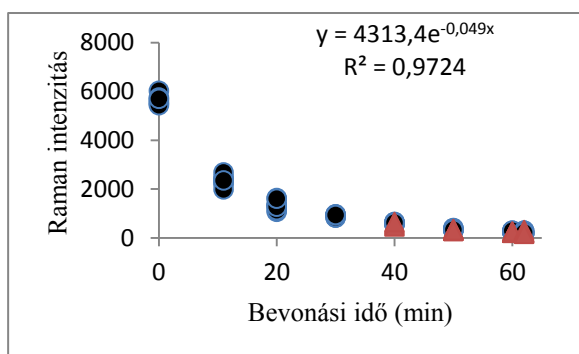
A kapott spektrumokat ezek után ALS módszer alkalmazásával alapvonal-korrigáltam, majd a hatóanyag (aszparaginsav kálium és magnézium sója) jellemző rezgési sávjának intenzitását ábrázoltam a mintavételi idők függvényében. A bevonatlan Panangin korpusz és a bevonó anyag rezgési spektrumát mutatja be a 4.2.1. ábra. Az eredmények értékelését a bevonatok szerint csoportosítva végeztem el. A bevonás során nem tapasztaltam a 859 cm^{-1} -es sáv eltolódását.



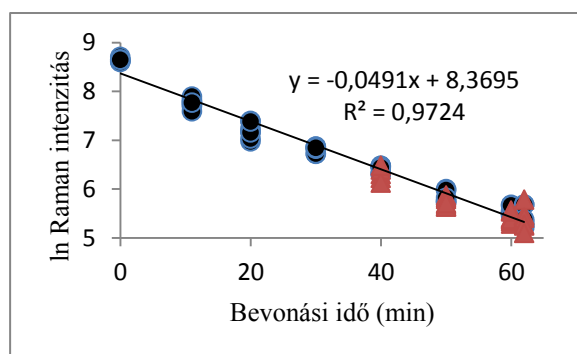
4.2.1. ábra: A Panangin tabletták korpusz (kék) és a bevonó anyag (zöld) spektruma, a kálium-magnézium aszparaginat jellemző rezgési sávja megjelölve (859 cm^{-1})

AdvantiaRED bevonat

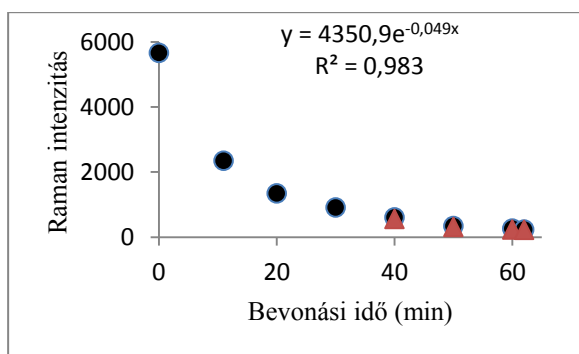
Az AdvantiaRED bevonattal bevont minták esetében a bevonás megkezdése előtt, hozzávetőlegesen 10 percenként, majd a bevonás végén (62 percnél) vettem hét mintát. A fent ismertetett előzetes adatkezelések után az aszparaginsav jellemző csúcsának (859 cm^{-1}) intenzitását ábrázoltam a mintavétel időpontjának függvényében (4.2.2. ábra). Ez nem a hatóanyag mennyiségének csökkenését jelenti, hanem azt, hogy a bevonat vastagságának növekedésével az aszparaginsav jellemző csúcsának intenzitása lecsökken, egyre nehezebben detektálható.



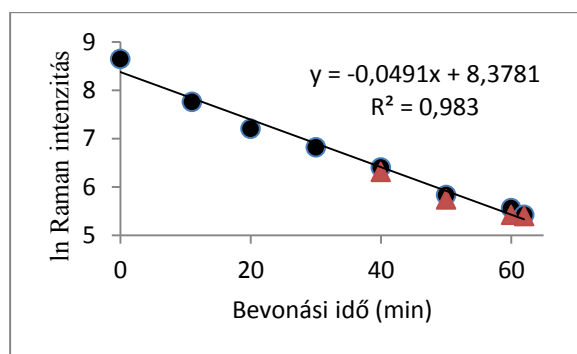
a) Az összes mért mintát tartalmazó görbe



b) Az a) görbe linearizált alakja



c) A minták átlagértékeire illesztett görbe



d) A c) görbe linearizált alakja

4.2.2. ábra: Az AdvantiaRED bevonattal bevont Panangin korpuszok jellemző csúcsának intenzitásának változása a bevonási idő függvényében (fekete pontok: a bevonás során vett minták, piros háromszög: a minták újbóli mérésének eredményei)

A 4.2.2. ábrán látható, hogy az aszparaginsav jellemző csúcsának intenzitása a bevonási idő növekedésével közelítőleg exponenciálisan csökken. A tabletták jellemző csúcsintenzitásainak szórása a bevonás kezdetén nagy, a szórás azonban jelentősen lecsökken a bevonat vastagságának növekedésével, amint az a 4.2.2. a) ábrán látható.

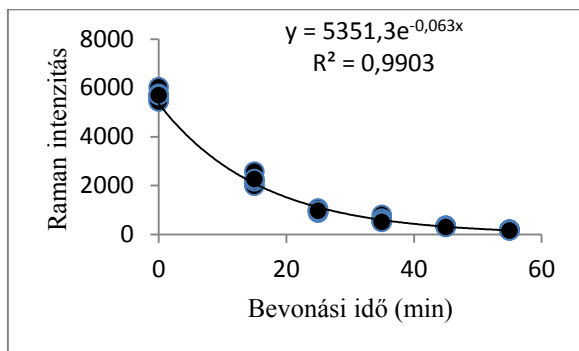
A jelenség a bevonási folyamat mechanizmusával magyarázható: a bevonás kezdeti szakaszában a bevonó oldat cseppjei csupán foltokban vannak jelen a korpuszon, ekkor a tabletták spektrumok intenzitásának szórása mellett ez az inhomogenitás is növeli a minták intenzitás értékeinek szórását. A homogén bevonat kialakulásával azonban ez a jelenség megszűnik, a fokozatosan és egyenletesen vastagodó bevonat pedig maszkírozza a tablettamagok hatóanyag-tartalma közötti eltérést. Az összes tabletták-korpusz azonos mennyiségű hatóanyagot tartalmazott, de a gyártás során ez meghatározott hibahatáron belül változhat. A tabletták-korpuszok hatóanyag-tartalmának pontos meghatározása után a Raman intenzitás értéke és a pontos hatóanyag-tartalom között kapcsolat teremthető. Így a módszer alkalmas lehet a gyártott tételek dózis-egyenletességének meghatározására is. Látható, hogy a bevonás követésére az összes mintát felhasználó görbe nem alkalmas: nem határozható meg egyértelműen az intenzitásértékek alapján, hogy az adott minta melyik mintavételi időponthoz tartozik.

A minták átlagolásával azonban ez a probléma kiküszöbölhető, amint az a 4.2.2. c) ábrán látható. Az exponenciális összefüggések alkalmazása a mindennapokban sokszor körülményes, ezért az intenzitás-értékek természetes alapú logaritmusát véve a görbét linearizáltam. A modellparaméterek értékeit tanulmányozva látható, hogy mindez nem okoz változást a modellparaméterek értékeiben, az összefüggés pedig sokkal egyszerűbben értelmezhető és alkalmazható lesz.

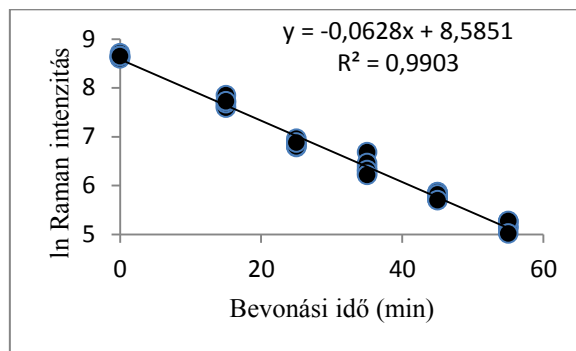
Az AdvantiaRED sorozat esetében látható, hogy a bevonás során vett minták szórása a fent részletezett jelenségen kívül is nagy. Eme mintasorozat esetében az utolsó négy mintavételi pontnál mért mintákat a mintatartóban megfordítva ismét lemértem, hogy megállapítsam, a módszer ismételhető-e. Ezeket a mérési pontokat a diagramokon piros háromszöggel jelöltem. Az eredmények ebben az esetben sem tértek el számottevően az eredeti vizsgálati eredményektől.

OpadryRED bevonat

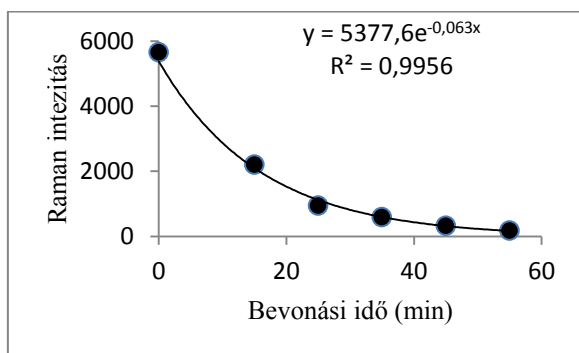
Az OpadryRED bevonatos mérési sorozat esetében hat alkalommal vettem mintát. Ebben az esetben is az aszparaginsav jellemző csúcsának intenzitását ábrázoltam bevonási idő függvényében. Az eredményeket mutatja be a 4.2.3. ábra.



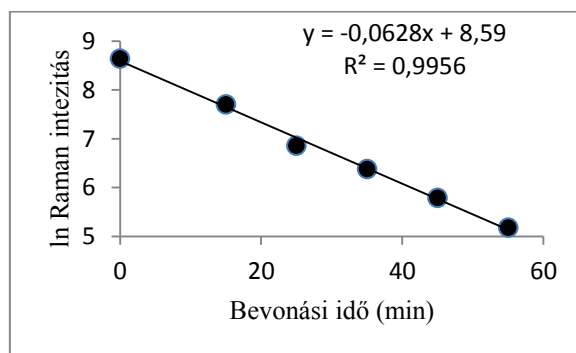
a) Az összes mért mintát tartalmazó görbe



b) Az előző a) linearizált alakja



c) A minták átlagértékeire illesztett görbe



d) A c) görbe linearizált alakja

4.2.3. ábra: Az OpadryRED bevonattal bevont Panangin tabletták korpuszok jellemző csúcshintenzitásának változása a bevonási idő függvényében

A 4.2.3. ábrán látható, hogy a hatóanyag csúcshintenzitása ebben az esetben is exponenciális csökkenést mutat a bevonási idő növekedésével. A tabletták szórása ebben az esetben nem olyan jelentős, mint az AdvantiaRED bevonat esetében (4.2.3. a) ábra). A minták szórásának csökkenése a bevonási idő előrehaladtával látványosabb, mint az AdvantiaRED bevonószer alkalmazásakor (4.2.3. a) ábra). Az összefüggés linearizálása után kapott egyenes modellparaméterei meggyőzőek, átlagolás nélkül is jó eredményt adnak. Az adott mintavételi pontok lineáris átlagolása tovább javítja ezeket a paramétereket.

Összefoglalva a két bevonattal készült bevonási folyamatokat megállapítható, hogy az intenzitás-bevonási idő összefüggés exponenciális, amely könnyen linearizálható. Egyszerű

adatkezelési módszerekkel és gyors mérésekkel pontos információhoz juthatunk a bevonás időbeli lefolyásáról és a dózis-egyenletességről is egy mérés során. A bevonás nyomonkövetése a kezdeti szakaszban pontatlan a nagy szórás miatt, de a transzmissziós Raman spektrometria kiválóan alkalmazható gyors és roncsolásmentes vizsgálati módszer a bevonás végpontjelzésére. A módszer a későbbiekben alkalmas lehet az eddig szinte egyeduralkodó tömegnövekedés-mérési bevonás nyomon követés leváltására, illetve alkalmazásával eleget tehetünk a dolgozat bevezetésében említett – nemzetközi gyógyszerhatóságok által preferált – PAT irányelveknek.

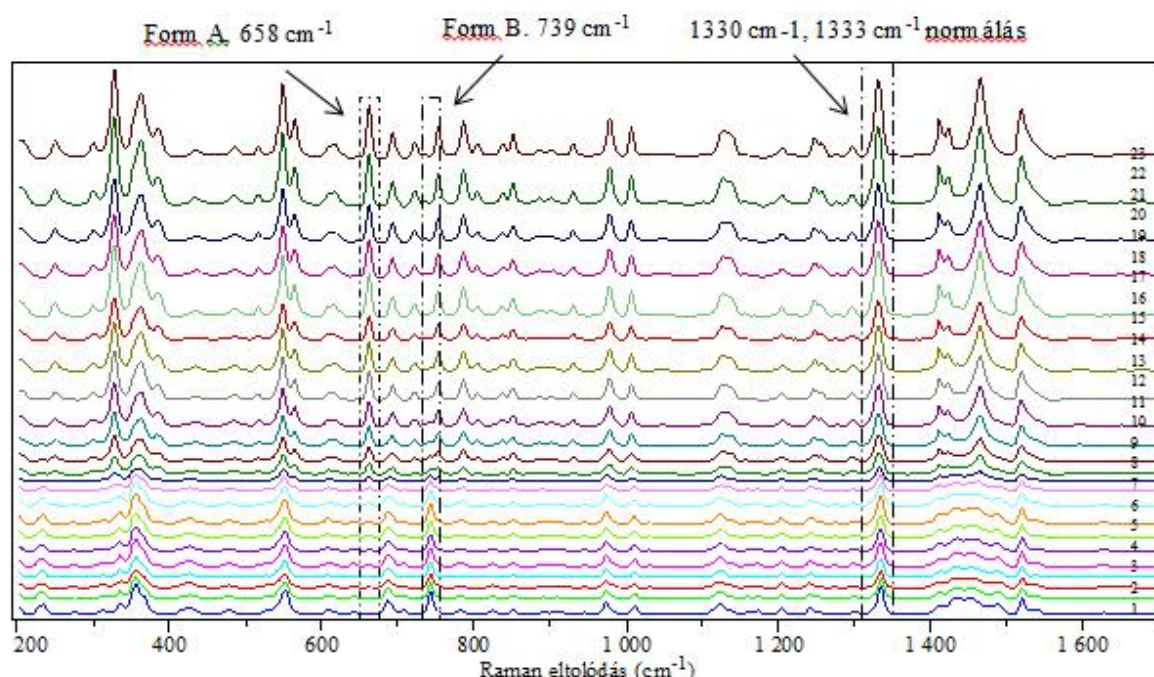
4.3. Famotidine kvantitatív meghatározása

A gyógyszerhatóanyagok polimorf összetételének pontos meghatározása a szilárdfázisú analitika fontos kihívása. A jelenleg legelterjedtebb módszer, az XRPD azonban nem alkalmas a kis koncentrációban jelen lévő polimorf szennyezések kimutatására. A mikro-Raman spektrometria érzékenysége korábbi tapasztalatok szerint meghaladja az XRPD technikáét, így kis mennyiségben jelen lévő polimorf forma kimutatását is lehetővé teszi. A meghatározás pontossága tovább növelhető kemometriai módszerek alkalmazásával, melyekkel a meghatározás hibájának számszerűsítése is megoldható. Felmerült a kérdés, hogy a különböző gyógyszerformák roncsolásmentes vizsgálatát a gyártási folyamatba illesztve is lehetővé tévő transzmissziós Raman spektrometria megközelíti-e a mikro-Raman spektrometria érzékenységét. A Raman-fotonok begyűjtése ez utóbbi esetben már egy mérés esetén is a készítmény teljes területéről történik. Dolgozatomban a famotidine hatóanyag két polimorf módosulatának mennyiségi elemzésén keresztül mutatom be a transzmissziós Raman technika alkalmazhatóságát a polimorf módosulatok mennyiségi meghatározásában. Az eredményeket összehasonlítottam a jelenleg alkalmazott reflexiós Raman-technikával. Az analízis célja a B polimorf módosulat, mint szennyező meghatározása volt.

A mennyiségi azonosítást a famotidine két polimorf módosulatának 23 különböző összetételével végeztem el. A minták spektrumait mutatja be a 4.3.1. ábra.

A 4.3.1. ábrán a különböző összetételű polimorf keverékek spektruma látható. A 658 cm^{-1} -nél fokozatosan megjelenő sáv a famotidine A módosulatára jellemző csúcs, amelynek intenzitása az A polimorf koncentrációjával párhuzamosan nő. Hasonló jelenség figyelhető meg a B módosulatra jellemző rezgési sáv (739 cm^{-1}) esetében is. A B forma koncentrációjának csökkenésével párhuzamosan a rá jellemző csúcs intenzitása is csökken, majd eltűnik.

A mennyiségi meghatározáshoz a dolgozatban bemutatott összes spektrum esetében azonos előzetes adatkezelési eljárásokat alkalmaztam. Minden spektrumot ALS módszerrel alapuló alapvonal-korrekciónak vettem alá, majd normáltam. A normálást egységnyi hosszra végeztem: a famotidine A módosulata esetében az 1330 cm^{-1} -es csúcsra, a B formánál pedig az 1333 cm^{-1} -es csúcsra. A két csúcs nagyon közel helyezkedik el egymáshoz, ezért választottam ezt a két rezgési sávot.



4.3.1. ábra: A 23 különböző összetételű famotidine polimorf keverék transzmissziós Raman spektruma (Form A tartalom: 1: 2%, 2: 3%, 3: 5%, 4: 7%, 5: 8%, 6: 10%, 7: 11%, 8: 15%, 9: 20%, 10: 30%, 11: 50%, 12: 70%, 13: 80%, 14: 85%, 15: 89%, 16: 90%, 17: 92%, 18: 93%, 19: 95%, 20: 97%, 21: 98%, 22: 99%, 23: 100%)

Az így előkészített spektrumokat ezek után minden egyes mérési sorozatban két halmazra bontottam szét: egy 17 elemű kalibráló és egy 6 elemű validáló halmazra. A kalibráló és validáló sorozat elemeit mutatja be a 4.3.1. táblázat. A kalibráló halmaz elemeiből kalibrációs modelleket készítettem, amely modellek pontosságát a validáló halmaz elemeivel ellenőriztem. Minden mérési sorozat esetében a fent leírt módon jártam el. Így vizsgálhatóvá válik a transzmissziós Raman spektrometria meghatározási pontossága, s egyéb módszerekkel (FT-Raman, porröntgen-diffrakció) is összehasonlíthatóvá válik. Famotidine hatóanyagra ezek a vizsgálatok a 4.3.1. táblázat szerinti felosztással már elvégezték [60]. Dolgozatomban én is ezt a felosztást követtem, hogy az összehasonlítás teljesen korrekt lehessen.

A 4.3.1. táblázatban feketével jelöltem a kalibráló halmaz, pirossal pedig a validáló halmaz elemeit. A 4.3.1. táblázatban látható, hogy a validáló halmaz tagjait a kalibrációs egyenes két végpontjához közel választottam ki. Azért ezeket a pontokat választottam, mert a

gyógyszeripari gyakorlatban az egyik tiszta polimorf módosulat (famotidine Form A) mellett a másik polimorf (Form B), mint szennyező minél pontosabb meghatározása a cél.

4.3.1. táblázat: A famotidine polimorf keverék minták kalibráló ill. validáló halmazainak elemei

Minta	Form A tartalom	Halmaz	Minta	Form A tartalom	Halmaz	Minta	Form A tartalom	Halmaz
FA298	2%	kalibráló	FA2080	20%	kalibráló	FA928	92%	kalibráló
FA397	3%	validáló	FA3070	30%	kalibráló	FA937	93%	validáló
FA595	5%	kalibráló	FA5050	50%	kalibráló	FA955	95%	kalibráló
FA793	7%	validáló	FA7030	70%	kalibráló	FA973	97%	validáló
FA892	8%	kalibráló	FA8020	80%	kalibráló	FA982	98%	kalibráló
FA1090	10%	kalibráló	FA8515	85%	kalibráló	FA991	99%	kalibráló
FA1189	11%	validáló	FA8911	89%	validáló	FA100	100%	kalibráló
FA1585	15%	kalibráló	FA9010	90%	kalibráló			

4.3.1. Transzmissziós Raman spektrometriás elővizsgálatok eredményei

Transzmissziós Raman spektrometriás méréseket először annak érdekében végeztem, hogy megismerkedjek a módszer pontosságával, megbízhatóságával és ismételhetőségével. E célból két eltérő kapszulatípusban, két különböző bemért mintamennyiséggel, három féle mérési idővel és akkumulációval végeztem méréseket.

Az összes bemutatott mintasorozat esetén a fent tárgyalt előzetes adatelőkészítési módszereket alkalmaztam, és a fent bemutatott halmaz-felosztással készítettem el a modelleket. Minden mintasorozat esetében az összes mintáról készült még egy mérés, ami a dolgozatban ismételt mérés néven szerepel. Ezeket a méréseket az adott mintasorozatból készült kalibrációs modellek jószágának ellenőrzésére használtam fel. Ez a gyakorlatban azt jelenti, hogy a 17 elemű kalibráló sorozatokat nem csak a 4.3.1. táblázatban pirossal jelölt 6 elemű halmaz elemeivel validáltam, hanem az ismételt mérés 23 mintájával is. Így minden kalibrációs egyenes kétféleképpen is validálva lett. Ezzel a méréssel a transzmissziós Raman spektrometria méréseinek ismételhetőségét vizsgáltam.

Minden mintasorozat elemeinek felhasználásával egyváltozós regressziós, CLS, PCR, PLS, SVM és PLS-GA modelleket készítettem. A modelleket a 4.3.1. ábrán bemutatott tartományon (65 cm^{-1} és 1700 cm^{-1}), illetve egy, csak a polimorfokra jellemző sávon ($732\text{--}763$

cm^{-1} és $648\text{--}668\text{ cm}^{-1}$), ún. redukált tartományon is elkészítettem. Az utóbbi két részből álló hullámszám tartományt irodalmi adatok alapján választottam ki [60].

Mindkét polimorf módosulatra (Form A és B) elkészítettem a modelleket az említett regressziós módszerek felhasználásával az előző bekezdésben ismertetett hullámszám sávokban, ebben a fejezetben azonban csak az A módosulatra jellemző értékeket mutatom be. A B módosulat eredményeit a Függelék tartalmazza.

A modellek paramétereit táblázatokba gyűjtöttem. Ezen táblázatok leglényegesebb modellparamétereket tartalmazó formáin mutatom be a különböző modellek eredményeit. A teljes, minden modell paramétert bemutató táblázatokat a Függelék tartalmazza.

Az ismételt mérések paramétereinek ismertetése során csak a modellek előrebecsléssel kapcsolatos modellparamétereit mutatom be, mivel a többi érték (RMSEC, RMSECV, stb.) megegyezik az eredeti, kalibráló modell paramétereivel, amelyek a táblázatokban megtalálhatóak.

A PLS-GA modellek egy optimalizált tartományokat kiválasztó GA lefutás eredményei, ugyanezeket a kiválasztott tartományokat alkalmaztam az összes modell elkészítése során a jobb összehasonlíthatóság végett.

A különböző regressziós modelleket minden mintasorozat esetén összehasonlítottam, és kiválasztottam a legjobbat közülük, melynél minden esetben az alábbiak szerint jártam el:

- I. A hegyomlás ábra (scree-plot) alapján kiválasztottam a komponensek optimális számát (az egyszerűség kedvéért a főkomponenseket, látens változókat, támogató vektorokat összefoglaló néven komponenseknek hívom). A modellek kiválasztásánál törekedtem a lehető legkevesebb főkomponens felhasználására. Ahol lehetett, előnyben részesítettem az egy komponens felhasználásával készült modelleket, mivel az összes tárgyalt mintasorozat minden mintája csak a két polimorf arányában különbözik. Ezért azt feltételeztem, hogy a minták csak a koncentrációjukban különböznek egymástól, tehát a halmazban lévő variancia legnagyobb részét a koncentráció változása teszi ki. Ebből következően egy látens változó is elegendő a modell felépítéséhez.

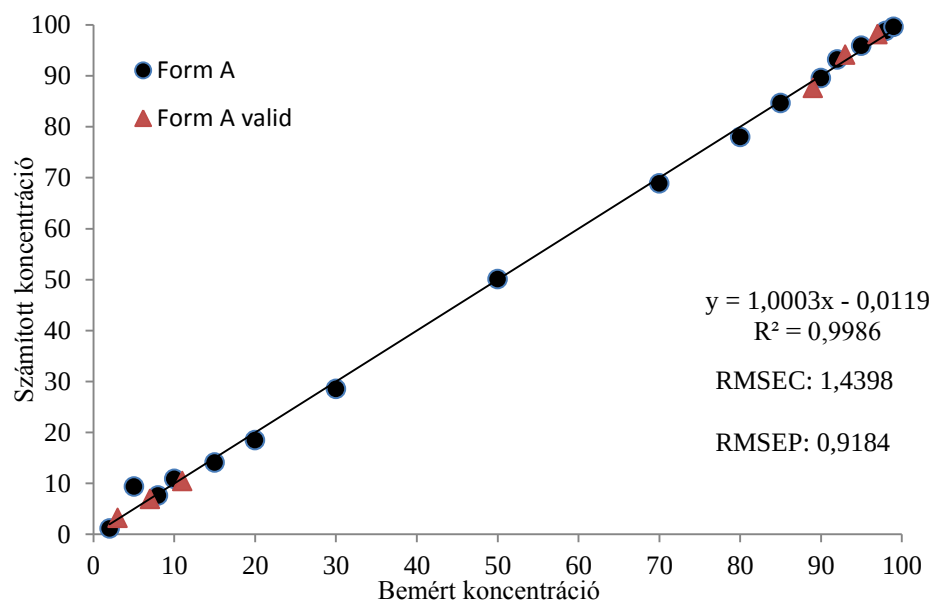
II. Minden modell esetében különös hangsúlyt fektettem a túlillesztés elkerülésére, ezért a jó teljesítményjellemzőkkel rendelkező, de a komponenseik száma, illetve a többi modellparaméter alapján túlillesztettnek tűnő modellek felhasználását a későbbi mérésekben nem javasoltam.

III. Ezek után összehasonlítottam a modelleket azok paramétereit alapján, és kiválasztottam a legeredményesebbet.

Az összehasonlítást nehezítette az, hogy a különböző módszerekkel készült regressziós módszerek között sok esetben a különbség szinte jelentéktelen, a modellparaméterek értékei között tizedekben, századokban mérhető a különbség. Ezekben az esetekben szubjektívan választottam ki a legjobb modelleket a szinte teljesen ugyanannyira jó modellek közül.

Egyváltozós regresszió:

Minden mintasorozat esetében egyváltozós regressziót is készítettem. Ennek során az alapvonal-korrigált, majd normált spektrumok esetében a polimorfokra jellemző rezgési sávok (A módosulat: 658 cm^{-1} , B módosulat: 739 cm^{-1}) csúcsainak maximum értékét feljegyeztem, és a kapott értékekre egyenest illesztettem. Ez az egyszerű módszer bizonyos esetekben meglepően jó eredményt szolgáltatott, ahogy azt a 4.3.2. ábra is mutatja. Az így elkészített egyváltozós regresszió eredményei azonban igen erősen függenek a spektrumok zajszintjétől, ezért a mérések többségében az így kapott eredmények már messze nem olyan jók, mint az itt bemutatott ábrán. Így a mintasorozatok bemutatásánál eltekintek ezen eredmények részletes ismertetésétől. Ezt a feltételezést támasztja alá a 4.3.3. fejezetben alkalmazott SRD módszer eredménye is. A fejezetben bemutatott táblázatok nem tartalmazzák az egyváltozós regresszió eredményeit, de a Függelékben megtalálhatók.



4.3.2. ábra: Az egyváltozós regresszió eredményei az AM15s100 minta esetében

A 4.3.2. ábrán látható, hogy eme egyszerű módszer ilyen hosszú mérési idő és magas akkumuláció-szám (100 darab 15 másodperces mérés készült minden mintáról) esetén kiváló eredményt szolgáltat. Kisebb mérési idő, illetve akkumuláció-szám azonban a modell paramétereinek nagymértékű romlását okozza.

Az egyváltozós regresszió eredményei után a különböző mintasorozatok eredményeit mutatom be. Elsőnek a teljes spektrális tartományt felhasználó modelleket ismertetem, ezt követi a redukált spektrális tartomány felhasználásával készült modellek értékelése.

A többváltozós regressziós módszereknél minden esetben azonos előzetes adatkezelés alkalmazására törekedtem. Ezeket szakirodalmi adatok [60], illetve a dolgozatban nem ismertetett előzetes kísérletek alapján választottam ki. Ennek eredményeként minden regressziós modellt a korábban már bemutatott alapvonal-korrekciónak és normálás után elsőrendű deriválásnak, majd skálázásnak vettem alá. Az azonos előzetes adatkezelések alkalmazására azért volt szükség, hogy a később bemutatandó SRD módszerrel jobban összehasonlítható modellekhez jussak, hogy a modellek közötti különbség ne az előzetes adatkezelések különbözőségének következménye legyen. Az általam legjobbnak ítélt modellt minden esetben szürkével emeltem ki a táblázatokban.

4.3.1.1. *Nagyobb töltöttségű kapszulák teljes spektrális tartományt felhasználó vizsgálatának eredményei*

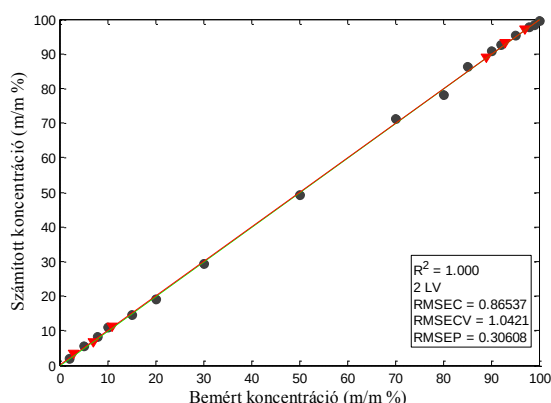
Átlátszó kapszulákba töltött nagyobb mennyiségű polimorf keverék mintasorozat (a továbbiakban: AM15s100T) tagjai 15 másodperc mérési idővel, és 100 akkumulációval készültek. A hosszú mérési idő és a sok akkumuláció miatt ezen spektrumok minősége kiváló, zajt csak igen kis mértékben tartalmaznak (ezeket a spektrumokat mutatja be a 4.3.1. ábra). A mintasorozatot felhasználva regressziós modelleket készítettem, amelyeknek általam kiválasztott eredményeit tartalmazza a 4.3.2. táblázat. Az összes modellparamétert a Függelék tartalmazza.

4.3.2. táblázat: AM15s100T mintasorozat különböző regressziós módszereinek eredményei

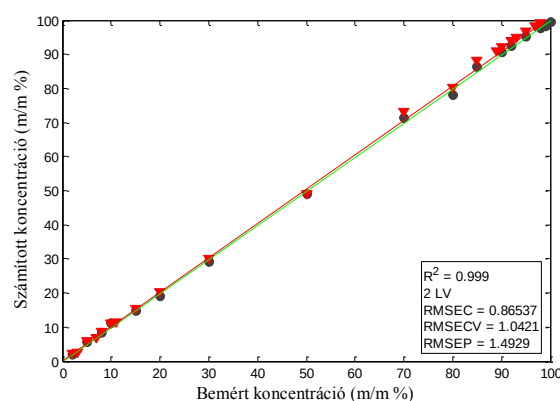
	PLS Form A	PCR Form A	CLS Form A	SVM Form A	PLS-GA Form A
Komponens	2	2	-	12	2
RMSEC	0,8654	0,9037	1,7995	0,3506	0,7982
RMSECV	1,0421	1,0417	2,0676	1,5085	1,0252
RMSEP	0,3061	0,2954	0,9282	0,9056	0,7484
R ² Kal	0,9995	0,9994	0,9978	0,9999	0,9996
R ² CV	0,9993	0,9993	0,9971	0,9986	0,9993
R ² Pred	1	1	0,9999	0,9998	0,9997
1.komp.%	81,67	81,68	81,65	-	73,77
Ismétlés					
RMSEP	1,4929	1,4562	2,3229	2,7456	3,5733
R ² Pred	0,9996	0,9996	0,9984	0,9998	0,9995

A 4.3.1. fejezet I. II. III. pontjaiban megadott kiválasztási folyamat szerint a legjobb modellt a PLS regresszió adta. A 4.3.2. táblázatban látható, hogy a faktoranalízisen alapuló modellek (PLS, PCR, PLS-GA) szolgáltatják a legjobb eredményt adó modelleket. Szinte minden modellparaméterük majdnem azonos. A PLS-GA modell sok esetben a legjobb eredményt mutatta, azonban az ismételt mérés esetében az RMSEP értéke elmaradt a többi modellétől. A PLS és PCR modell összes paramétere nagyon közel esik egymáshoz, ami a modellek majdnem azonos alapok szerinti működésének tudható be. Az irodalomban leggyakrabban alkalmazott PLS modellt választottam a legjobbnak, mert a PLS explicit módon figyelembe veszi a függő változókat is, ami stabilabb modellt eredményez. Az SVM

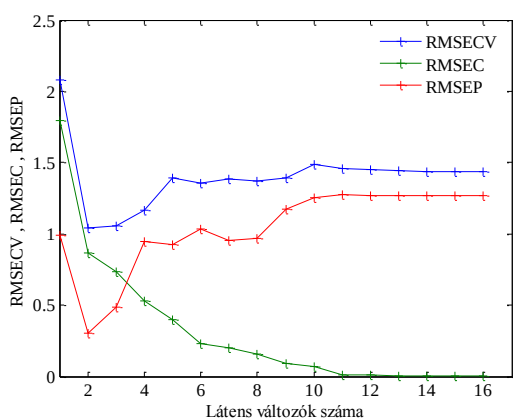
modell jelentősen túlilleszt, amit a támogató vektorok nagy száma, és az ismétlés nagy RMSEP értéke mutat be legjobban. A legrosszabb eredményt a CLS modellezés adta: minden modellparamétere elmarad a többi modelltől. A PLS modell által illesztett egyenest, hegyomlás-ábrát (scree-plot) és a reziduumjainak eloszlását mutatja be a 4.3.3. ábra.



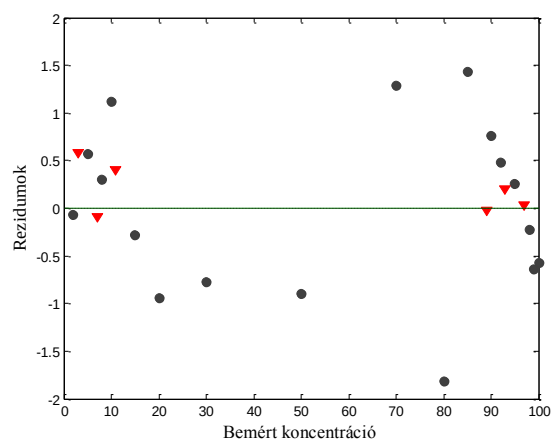
a) PLS modell



b) PLS modell ismételt mérés



c) PLS modell hegyomlás-ábra (scree-plot)



d) PLS modell reziduumjainak ábrázolása

4.3.3.ábra: A PLS modell kalibrációs egyenese, kalibrációs egyenes ismételt mérés esetén, hegyomlás ábrája és a reziduumok eloszlása

A 4.3.3. ábrán látható, hogy a PLS módszerrel készült kalibráció nem csak a modellparaméterek alapján nagyon jó modell, hanem a kalibrációs egyenese alapján is. A validáló minták is a kalibrációs egyenesen találhatók, alig van eltérés az ideális egyenestől.

Majdnem ugyanez mondható el az ismételt mérés esetében is, igaz, itt már nagyobb eltérés tapasztalható. A hegyomlás-ábrán látható, hogy a látens változók optimális száma kettő. A reziduumok eloszlása nem véletlenszerű, az ismétlés is hasonló ábrát eredményez. Ezeket a vizsgálatokat az összes többi modell esetében elvégeztem, azonban helyhiány miatt a dolgozatban részletesen nem közlöm őket.

4.3.1.2. *Nagyobb töltöttségű kapszulák redukált spektrális tartományt felhasználó vizsgálatainak eredményei*

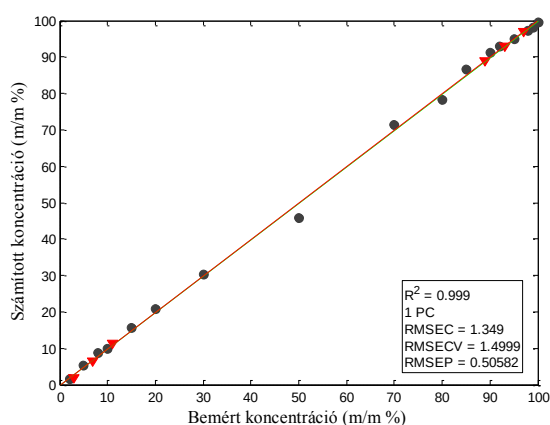
Az átlátszó kapszulákba töltött nagyobb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó (a továbbiakban: AM15s100R) mintasorozata az előző fejezetben bemutatott mintasorozat felvett spektrumainak megismétlése szűkített tartományon, amely csak a két polimorf módosulatra jellemző csúcsokat tartalmazza. Előzetesen ettől a lépéstől a modellek javulását vártam. Az eredményeket mutatja be a 4.3.3. táblázat.

4.3.3.táblázat: Az AM15s100R mintasorozat különböző regressziós módszereinek eredményei

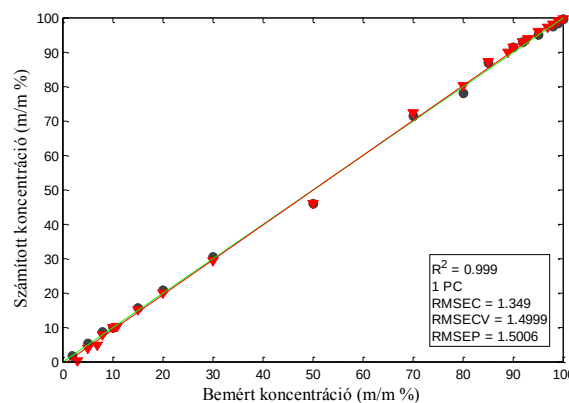
	PLS Form A	PCR Form A	CLS Form A	SVM Form A	PLS-GA Form A
Komponens	1	1	-	10	2
RMSEC	1,328	1,349	1,3288	0,3471	0,9354
RMSECV	1,4954	1,4999	1,4913	1,3532	1,1284
RMSEP	0,5109	0,5058	0,5439	0,5131	0,4228
R ² Kal	0,9988	0,9988	0,9988	0,9999	0,9994
R ² CV	0,9985	0,9985	0,9985	0,9989	0,9991
R ² Pred	0,9999	0,9999	0,9999	0,9999	0,9999
1.komp%	92,36	92,36	92,36	-	95,24
Ismétlés					
RMSEP	1,4834	1,5006	1,5139	1,1013	1,6186
R ² Pred	0,9991	0,9991	0,9991	0,9998	0,9993

A 4.3.3. táblázatban láthatóak a módszerek eredményei kivonatolva. A többi modellparaméter a Függelékben található meg. Ebben az esetben is a faktoranalízisen alapuló modellek szolgáltatták a legjobb rendszereket. Csak a modellparamétereket vizsgálva a legjobb eredményt a PLS-GA regresszió adta, habár a regresszió ismételt mérésbeli RMSEP

értéke az összes vizsgált modell közül a legrosszabb. A PLS és PCR regresszió esetében a hegyomlás-ábrát tanulmányozva (a dolgozat helyhiány miatt nem tartalmazza) azonban észrevehető, hogy a legkedvezőbb egy komponens alkalmazása. Ezek a módszerek ugyan elmaradnak a modellparaméterek értékeiben a genetikus algoritmusokkal támogatott PLS regresszió eredményeitől, de egyetlen komponens felhasználásával is jól ismételhető modellt adnak. A két modell szinte teljesen megegyezik minden paraméterében, különbség nem tehető köztük. Az SVM modellt a támogató vektorok optimálisnál nagyobb száma miatt túllílesztettnek ítélem, amit a nagy RMSECV értéke is bizonyít. A CLS kiértékelés ebben az esetben szinte teljes mértékben azonos eredményt ad a PLS és PCR modellekkel. A 4.3.4. ábrán a PCR modell kalibrációs egyeneseit mutatom be.



a) PCR modell



b) PCR modell ismételt mérés

4.3.4. ábra: A PCR regresszió kalibrációs egyenese és az ismételt mérés kalibrációs egyenese

A 4.3.4. ábrán látható, hogy az egy főkomponens felhasználásával készült PCR modell által illesztett egyenes is megközelíti az ideális egyenest, ahogyan az ismételt mérés eredményei sem térnek el az eredeti mérés eredményeitől.

4.3.1.3. *Kisebb töltöttségű kapszulák teljes spektrális tartományt felhasználó vizsgálatainak eredményei*

Az átlátszó kapszulákba töltött kevesebb mennyiségű polimorf keveréket tartalmazó mintasorozatban (a továbbiakban: AL5s50T) a bemért anyagmennyiség már jóval kevesebb

volt: 0,1 gramm alatti. Ebben az esetben is a fejezet elején ismertetett halmazokkal végeztem a kalibráló és validáló, illetve ismételt mérések eredményeit bemutató modellek felépítését és értékelését. Ennél a mintasorozathoz és mérési beállításához azonban készült egy újabb validálásra felhasznált mintasorozat. Az ide tartozó polimorf keverékek összetételét a 4.3.4. táblázat mutatja be.

4.3.4. táblázat: A kevesebb mennyiségű polimorf keveréket tartalmazó mintákról a 3.1. táblázatban ismertetett beállítások mellett készült független bemérés mintáinak Form A polimorf tartalmát bemutató táblázat

Minta	Form A tartalom
FA0892	8%
FA2080	20%
FA5050	50%
FA8020	80%
FA9208	92%
FA1000	100%

Ezek a minták az eddigi mintasorozatoktól függetlenül, a 4.3.1. táblázatban bemutatott polimorf keverékekből lettek bemérve, az előző mintasorozattal azonos mérési idővel és ismétléssel. A korábbi (amiből a kalibráló, validáló modell, és az ismételt mérés származik) és az új bemérés tömegre megegyezik, csak a bemérés ismételhetőségének hibájában különböznek egymástól. Minden, a táblázatban bemutatott összetételből 4 darab minta készült. Az így kimért mintákat ezek után a régi beméréssel azonos körülmények között lemértem. Erről a mintasorozatról is készült ismételt mérés, amelyet az „új bemérés 2” névvel láttam el. A két mintasorozat így válik összehasonlíthatóvá. Ezzel a kísérlettel a modellek stabilitását, prediktív erejét demonstrálom.

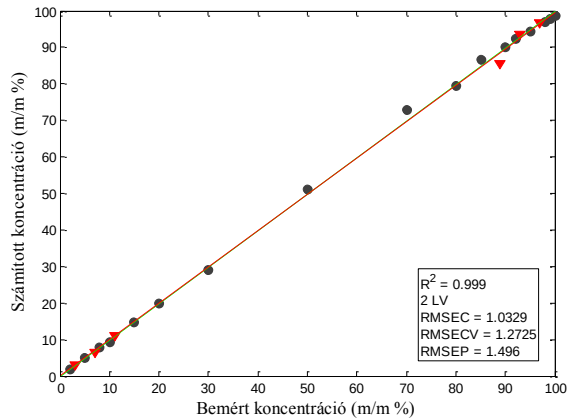
A 4.3.5. táblázatban látható, hogy a legjobb eredményt a PLS modell adta, nem sokban marad el tőle a PCR módszer. A PLS-GA modell kalibrációs és validáló sorozatot felhasználó eredményei jobbak, de az ismétlés, és az újabb bemérések esetében ez a módszer adja a legrosszabb eredményt. Az SVM regresszió itt is túlilleszt, amit a támogató vektorok nagy száma és az új bemérés eredményei is igazolnak. A CLS módszer most is elmarad a többi módszertől, de a PLS-GA módszernél jobban előrejelzi az új bemérések koncentrációit. A táblázatban látható, hogy a legrobustusabb modellt a PLS módszer adja, mivel az ismételt

mérés, és az új bemérés, és annak ismétlése ezeknél a módszereknél szolgáltatja a legjobb RMSEP értékeket.

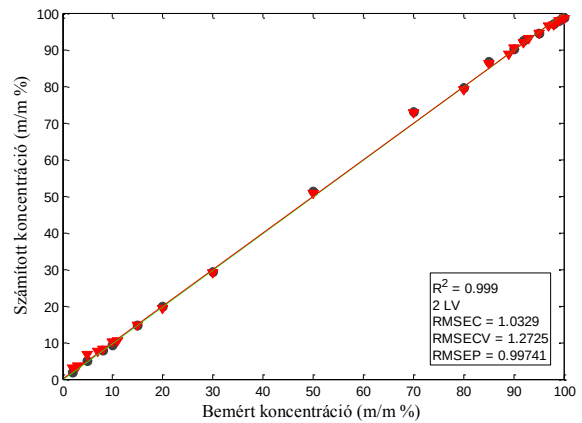
4.3.5. táblázat: Az AL5s50T mintasorozat különböző regressziós módszerek eredményeinek összehasonlítása

	PLS Form A	PCR Form A	CLS Form A	SVM Form A	PLS-GA Form A
Komponens	2	2	-	12	2
RMSEC	1,0329	1,2057	1,4695	0,5886	0,8749
RMSECV	1,2725	1,4635	1,5713	1,1475	1,3079
RMSEP	1,496	1,4476	1,5434	1,8924	2,3343
R ² Kal	0,9993	0,999	0,9985	0,9998	0,9995
R ² CV	0,9989	0,9985	0,9983	0,9991	0,9988
R ² Pred	0,9991	0,9992	0,9991	0,9985	0,9984
1. komp%	87,94	87,94	87,93	-	84,32
Ismétlés					
RMSEP	0,9974	1,0341	1,2353	1,1116	1,5385
R ² Pred	0,9994	0,9994	0,9991	0,9994	0,9987
Új bemérés 1					
RMSEP	0,8987	0,9866	1,1209	1,484	1,2203
R ² Pred	0,9997	0,9996	0,9997	0,9995	0,9995
Új bemérés 2					
RMSEP	0,8693	0,9673	1,0967	1,4663	1,688
R ² Pred	0,9997	0,9996	0,9997	0,9996	0,9988

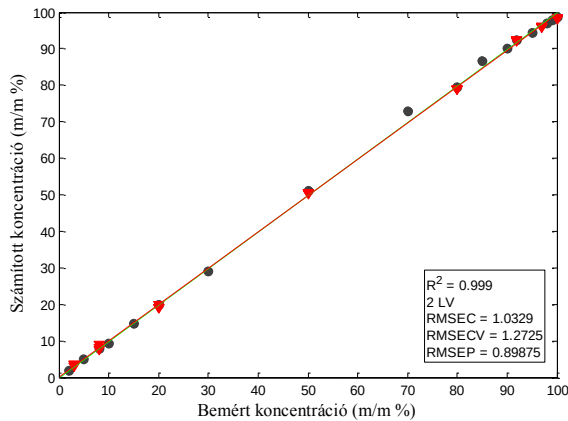
A 4.3.5. ábrán látható, hogy ebben az esetben is az ideálist nagyban megközelítő kalibrációs egyenest adott a PLS modell. A validáló sorozatban megfigyelhető, hogy a 89% Form A polimorfot tartalmazó minta nem illeszkedik az egyenesre, ám az eltérés nem nagymértékű. Az ismételt mérés esetében ez a különbség már nem tapasztalható, és az eddigiektől eltérően az ismételt mérés RMSEP értéke jobb, mint a validáló sorozaté. Az új bemérés, és annak ismétlésével további javulást értem el, ennek a validáló sorozatnak az elemei jobb RMSEP értékeket adnak az ismétlésnél. Látható, hogy a PLS modell robusztussága miatt a mérés új, független minták esetében is megbízható, és jó eredményt szolgáltat.



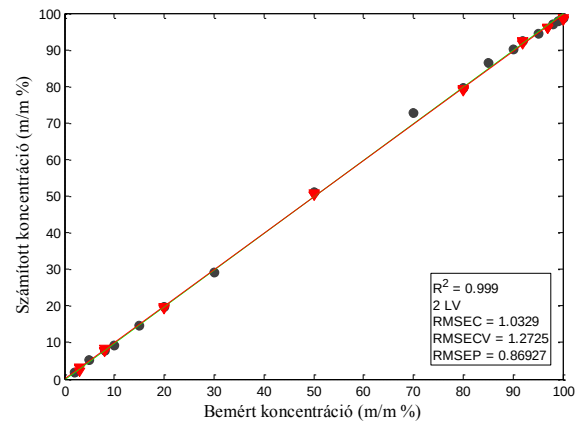
a) PLS modell



b) PLS modell ismételt mérés



c) PLS modell új bemérés



d) PLS modell új bemérés ismétlés

4.3.5. ábra: A PLS modell kalibrációs egyenesei a kalibráló, validáló, ismételt mérés, és az új bemérés, és annak ismétlése esetében

4.3.1.4. Az összes mintasorozat eredményeinek összehasonlítása

Dolgozatomban hat különböző mintasorozatot vizsgáltam meg. Ebből az előző fejezetekben két mintasorozat kiértékelését mutattam be részletesen. A többi modell ilyen részletesen történő kiértékelésétől dolgozatomban eltekintek. Az összes itt bemutatott modell valamennyi eredménye a Függelékben található meg. Az eddigi fejezetekben ismertettem azt a módszert, amely alapján minden mintasorozat esetében kiválasztottam a legjobb modellt. Ebben a fejezetben is ezt a módszert alkalmaztam. Ehelyett a különböző mintasorozatok összehasonlítására fektettem nagyobb hangsúlyt. Így lehetőség nyílik az eltérő paraméterek (a modellek alkalmazhatósága az adott problémában, robusztusságuk jellemzése) hatásának tanulmányozására. Az összefoglaló táblázatban szerepelnek az eddig kiértékelt modellek az

összehasonlítás szempontjából legfontosabb paramétereikkel. A mintasorozatok nevei a táblázatban rövidített formában szerepelnek, ezért a rövidítések feloldását tartalmazza a 4.3.6. táblázat.

4.3.6. táblázat: A mintasorozatok nevei, és a táblázatban alkalmazott rövidítésük

A mintasorozat elnevezése	A mintasorozat rövidítése
Átlátszó kapszulába töltött nagyobb mennyiségű polimorf keverék teljes spektrális tartományt felhasználó mintasorozat (mérési idő 15 sec, akkumulációk száma 100)	AM15s100T
Átlátszó kapszulába töltött nagyobb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó mintasorozat (mérési idő 15 sec, akkumulációk száma 100)	AM15s100R
Átlátszó kapszulába töltött kisebb mennyiségű polimorf keverék teljes spektrális tartományt felhasználó mintasorozat (mérési idő 5 sec, akkumulációk száma 50)	AL5s50T
Átlátszó kapszulába töltött kisebb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó mintasorozat (mérési idő 5 sec, akkumulációk száma 50)	AL5s50R
Titán-dioxid tartalmú kapszulába töltött kisebb mennyiségű polimorf keverék teljes spektrális tartományt felhasználó mintasorozat (mérési idő 1 sec, akkumulációk száma 1)	TL1s1T
Titán-dioxid tartalmú kapszulába töltött kisebb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó mintasorozat (mérési idő 1 sec, akkumulációk száma 1)	TL1s1R
Titán-dioxid tartalmú kapszulába töltött kisebb mennyiségű polimorf keverék teljes spektrális tartományt felhasználó mintasorozat (mérési idő 15 sec, akkumulációk száma 100)	TL15s100T
Titán-dioxid tartalmú kapszulába töltött kisebb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó mintasorozat (mérési idő 15 sec, akkumulációk száma 100)	TL15s100R
Titán-dioxid tartalmú kapszulába töltött nagyobb mennyiségű polimorf keverék teljes spektrális tartományt felhasználó mintasorozat (mérési idő 1 sec, akkumulációk száma 1)	TM1s1T
Titán-dioxid tartalmú kapszulába töltött nagyobb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó mintasorozat (mérési idő 1 sec, akkumulációk száma 1)	TM1s1R
Titán-dioxid tartalmú kapszulába töltött nagyobb mennyiségű polimorf keverék teljes spektrális tartományt felhasználó mintasorozat (mérési idő 15 sec, akkumulációk száma 100)	TM15s100T
Titán-dioxid tartalmú kapszulába töltött nagyobb mennyiségű polimorf keverék redukált spektrális tartományt felhasználó mintasorozat (mérési idő 15 sec, akkumulációk száma 100)	TM15s100R

A 4.3.7. táblázatban láthatóak az összes teljes spektrális tartomány felhasználásával készült modellek paramétereinek értékei. Szürkével jelöltem a kiszámolt modellparaméterek vizsgálatával általam legjobbnak ítélt modelleket minden mintasorozat esetében. A táblázatok eredményeit megvizsgálva látható, hogy sok esetben a regressziós modellek közötti különbség

nem jelentős, a jósági paraméterek értékei tizedekben, századokban térnek el csak egymástól. Ez megnehezítette a modellek közül a legjobb kiválasztását.

AM15s100: Az AM15s100 mintasorozat esetében látható, hogy a faktoranalízises modellek adták a legjobb eredményeket. A PLS modellt választottam legjobbnak, mert általában a PLS regresszió a függő változók explicit felhasználásával a PCR kiértékelésnél robusztusabb modellt szolgáltat. Az SVM regresszió hiába ad jó RMSEC értéket, azonban a módszer túlilleszt. A PLS-GA modell nem sokban marad el a PLS és PCR modell eredményeitől.

AL5s50: Az AL5s50 mintasorozat eredményeit értékelve megállapítható, hogy a legjobb kiértékelést ebben az esetben is a PLS modellezés adta. A modell paramétereinek értékei jobbak a PCR, CLS, PLS-GA modellnél is. Az SVM modell túlilleszt.

TL1s1: A PLS regresszió adta a legjobb eredményeket a TL1s1 sorozat esetében. Teljesítménye azonban nem kiegyensúlyozott: látszik, hogy az RMSEC, RMSECV és RMSEP értékei között nagy különbségek vannak, nagyobbak, mint a PCR, CLS, PLS-GA modelleknél. Az SVM regresszió itt is túlilleszt.

TL15s100: A megnövelt mérési idejű TL15s100 sorozat mennyiségi összetételének meghatározásában a PLS és az SVM modell is egyaránt jól teljesített. Az SVM modell eredményei jobbak, ugyanakkor tizenhat támogató vektor felhasználásával készült a modell. Az ismételt mérés azonban nem mutatott túlillesztésre utaló jelet.

TM1s1: A nagyobb bemért anyagmennyiséget tartalmazó TM1s1 sorozat rövid expozíciós idővel (1 másodperc) és 1 akkumulációval került lemérésre. Látható, hogy a modelleknek ezért gyenge paraméterei és előrebecslő képességük van. Legjobb eredménnyel a PLS modell jelzi előre a koncentrációkat, de az RMSEC és a többi négyzetes középhiba paraméter értéke között nagy a különbség. A többi modell eredményei elmaradnak a PLS modellétől, a túlillesztés az SVM esetében nagyon feltűnő.

TM15s100: Ugyanezzel a bemért mintatömeggel készült el a TM15s100 mintasorozat is, azonban a mérésidő és az akkumulációk száma itt már sokkal több volt (15 másodperc illetve 100 akkumuláció). Ez a modellek jóságán és előrebecslő képességén is megfigyelhető. Látható, hogy ebben az esetben a PLS és a PLS-GA adta a legjobb eredményt.

4.3.7. táblázat: A teljes spektrális tartomány felhasználásával készült modellek paramétereinek értékei kivonatolt formában

	PLS				PCR				CLS			
	RMSEC	RMSEP	Komp.	Ism RMSEP	RMSEC	RMSEP	Komp.	Ism RMSEP	RMSEC	RMSEP	Komp.	Ism RMSEP
AM15s100	0,8654	0,3061	2	1,4929	0,9037	0,2954	2	1,4562	1,7995	0,9282	-	2,3229
AL5s50	1,0329	1,496	2	0,9974	1,2057	1,4476	2	1,0341	1,4695	1,5434	-	1,2353
TL1s1	1,1679	4,779	2	2,7486	3,3685	3,5902	2	3,0909	3,1855	3,7476	-	2,9143
TL15s100	1,0653	1,2733	2	1,2393	1,266	1,3	2	1,3682	1,4393	1,4085	-	1,5389
TM1s1	1,8117	3,0344	2	4,5076	2,9763	3,03	2	5,9271	3,3086	2,8133	-	6,6582
TM15s100	0,9356	0,4243	2	1,0939	0,9962	0,398	2	1,1604	1,2245	0,2481	-	1,5156
	SVM				PLS-GA							
	RMSEC	RMSEP	Komp.	Ism RMSEP	RMSEC	RMSEP	Komp.	Ism RMSEP				
AM15s100	0,3506	0,9056	12	2,7456	0,7982	0,7484	2	3,5733				
AL5s50	0,5886	1,8924	12	1,1116	0,8749	2,3343	2	1,5385				
TL1s1	0,0402	5,0837	17	2,3809	3,3332	4,5382	2	4,7013				
TL15s100	0,8613	1,1311	16	1,158	1,0833	0,9461	2	1,47				
TM1s1	0,0089	3,3158	17	12,4333	3,2185	2,7507	2	6,2694				
TM15s100	0,6798	0,4637	12	0,9504	0,9442	0,6439	2	1,0525				

Az összes mintasorozat eredményeit összevetve a felhasznált regressziós módszerek összehasonlíthatóvá tették a különböző mérési beállítások eredményeit. A legjobb eredmény minden esetben a PLS modellezéssel állt elő. Ezek az eredmények is a PLS regresszió robusztusságát bizonyítják. Emellett két esetben (TL15s100 és TM15s100) az SVM regresszió és a PLS-GA módszer is a PLS regresszióval összemérhetően jó modellt adott. A PCR módszer eredményeiben nem sokkal maradt el a PLS módszertől. A CLS kiértékelés elmarad ezektől a módszerektől, de nem sokban. Az SVM regressziós modell az esetek nagy részében túlilleszt. A PLS-GA itt nem javította nagymértékben a modell jóságát, sok esetben az eredeti PLS modellhez képest a teljesítményjellemzők romlása figyelhető meg. Megjegyzendő azonban, hogy sok esetben a modellek közötti különbségek minimálisak voltak, ez megnehezítette a legjobb modell kiválasztását.

A mintasorozatokat összehasonlítva látható, hogy az átlátszó zselatinkapszulában mért mintasorozatok igen csekély mértékben, de jobb eredményt adtak a titándioxid tartalmú kapszulákba töltött mintákban elért értékeknél. Hosszabb mérési idő és több ismétlés alkalmazása nagyobb mértékű javulást eredményez a modellparaméterek értékeiben. Ez a javulás legfőképpen az ismételt mérések esetében érhető tetten. A mérési idő harmadára és az akkumulációk számának felére csökkenése nem okozza a modellparaméterek nagymértékű romlását, amit az az AL5s50 mintasorozat esetében látható. Összességében elmondható, hogy a kapszulák összetétele kevésbé, a mérési idő és az akkumulációk száma jobban befolyásolja a modellek jóságát a teljes spektrumtartományt felhasználó modellek esetében.

A 4.3.8. táblázat mutatja be a redukált spektrális tartomány felhasználásával készült modellek eredményeit. Ebben az esetben is a szürke színnel kiemelt sorok jelzik az általam legjobbnak ítélt modelleket.

AM15s100: Az AM15s100 mintasorozat esetében a faktoranalízisen alapuló módszerek (PLS, PCR, PLS-GA) között nem fedezhető fel nagy különbség, eredményeik hasonlóan jók. Külön kiemelendő, hogy a PLS és PCR modellel ugyanez egy látens változó felhasználásával érhető el. Ezekről az eredményekről a CLS modellezés sem ad szignifikánsan rosszabb kiértékelést. Az SVM regresszió paraméterei a legjobbak az összes vizsgált modell közül, azonban a modell túlilleszt, ami a támogató vektorok számán és az ismételt mérések eredményein is látszik.

4.3.8. táblázat: A redukált spektrális tartomány felhasználásával készült modellek paramétereinek értékei kivonatolt formában

	PLS				PCR				CLS			
	RMSEC	RMSEP	Komp.	Ism RMSEP	RMSEC	RMSEP	Komp.	Ism RMSEP	RMSEC	RMSEP	Komp.	Ism RMSEP
AM15s100	1,328	0,5109		1 1,4834	1,349	0,5058		1 1,5006	1,3288	0,5439	-	1,5139
AL5s50	0,7508	1,6814		2 1,5417	0,7549	1,6756		2 1,5009	1,7378	1,5837	-	1,5618
TL1s1	3,2936	6,5944		2 6,6226	4,8003	7,4036		2 6,7809	6,9265	8,3807	-	7,0904
TL15s100	1,1087	1,2952		2 1,3796	1,3204	1,1911		2 1,5256	1,3673	1,1856	-	1,4892
TM1s1	4,1983	3,3666		2 5,6157	5,0362	4,2331		2 5,9103	5,0807	4,0734	-	6,5759
TM15s100	1,23	1,2498		1 1,109	1,2318	1,2454		1 1,108	1,2306	1,2733	-	1,11
	SVM				PLS-GA							
	RMSEC	RMSEP	Komp.	Ism RMSEP	RMSEC	RMSEP	Komp.	Ism RMSEP				
AM15s100	0,3471	0,5131		10 1,1013	0,9354	0,4228		2 1,6186				
AL5s50	0,7777	1,6266		12 0,9293	1,3142	1,2817		2 1,3025				
TL1s1	0,0133	7,0624		17 6,5524	6,3364	4,1001		2 6,7569				
TL15s100	0,9788	1,3692		13 1,182	0,9617	1,1559		2 1,4154				
TM1s1	3,0825	3,0212		8 5,3569	5,1126	9,4475		1 10,8679				
TM15s100	0,0416	1,4757		16 0,7622	1,2535	1,1158		1 1,2897				

AM5s50: A PLS és a genetikus algoritmusok segítségével kiválasztott változókat felhasználó PLS modell (PLS-GA) adta a legjobb modelleket az AM5s50 mintasorozat esetében. A PLS modellnél megfigyelhető, hogy a kalibráció lényegesen jobb a validáló pontok, illetve az ismételt mérés modellparamétereinél. Eme modellektől nem marad el sokban a PCR kiértékelés sem. A CLS modellezés eredményei sem különböznek jelentős mértékben a többi modell paramétereitől. Az SVM modellezés ebben az esetben is túlilleszt.

TL1s1: A redukált tartomány felhasználásával készült modellek zajra való érzékenysége látszik a TL1s1 mintasorozat eredményeit tanulmányozva. Minden modell esetében a kalibráló sorozathoz képest a validáló sorozat teljesítményjellemzői kétszer rosszabbak a kalibráló sorozat megfelelő értékeinél. Ennél sokkal nagyobb mértékű a romlás a teljesítményjellemzők értékeiben az SVM regresszió esetében. A legrosszabb eredményt a CLS regresszió adta, míg a legjobb modellnek a PLS modellt választottam.

TL15s100: A mérési időt és az akkumulációk számát megnövelve a modellek eredményei sokkal jobbak lettek a TL15s100 mintasorozat esetében. A PLS és a PLS-GA modellezés közel egyenértékű teljesítményjellemzőket adott a kalibráló adatsor feldolgozásában. A PCR és CLS modellezés is megközelítőleg azonos eredményekkel jelezte előre a minták koncentrációit, az SVM regresszió viszont túlillesztett. A PLS és PLS-GA modelleket választottam legjobbnak.

TM1s1: A zajra való érzékenység a másik egy másodperces mérés esetében is tetten érhető, ezt mutatja be a TM1s1 mintasorozat kiértékelése. Itt nem olyan drámai mértékű az ismételt mérések eredményeinek romlása, de a kalibráló sorozatok RMSEC értékei magasak. Ebben az esetben meglepően jó adatokat szolgáltatott az SVM regresszió, egy nem túlillesztett modellben a többi modellnél jobb teljesítmény jellemzőket adó modellt épített fel. Következésképpen ennél a mintasorozatnál az SVM regresszió adta a legjobb modellt.

TM15s100: A TM15s100 mintasorozat eleminek felhasználásával már egy látens változó felhasználása is elegendő volt a megfelelő modelljóság paraméterekkel felruházott modellek felépítéséhez. A táblázatban látható, hogy minden modell esetében egy látens változót használtam fel a modellek felépítéséhez, kivéve az SVM regressziót, amely ebben az esetben is túlilleszt. A CLS regresszió eredményei közelítenek a faktoranalízises regresszió eredményeihez. A legjobb modellt itt a faktoranalízisen alapuló modellek adták (PLS, PCR).

Ebben az esetben is elvégeztem a regressziós módszerek összehasonlítását a mérési beállítások és a mért rendszerek szempontjából. A legjobb eredményeket jelen mintasorozatonál is a PLS regresszió szolgáltatta, több esetben egy látens változó felhasználásával értem el jó modellparamétereket. PCR regresszióval is ugyanilyen jó eredmények érthetők el nagy mérési idővel mért és nagyobb anyagmennyiséget tartalmazó mintasorozatok felhasználásával. A CLS regresszió kicsivel ugyan, de elmarad a faktoranalízises módszerek hatékonyságától. Az SVM regresszió a redukált spektrális tartományon is túlilleszt az esetek nagy részében, kivéve a TM1s1 mintasorozatot, ahol a legjobb modellt szolgáltatta. A genetikus algoritmusokkal segített PLS modell (PLS-GA) sokszor a modellparaméterek javulásához vezetett. A TL1s1 mintasorozat esetében azonban a legrosszabb modellt ez a módszer szolgáltatta. A legjobb modell kiválasztását itt is megnehezítették a regressziós módszerek közötti kis különbségek.

4.3.1.5. A transzmissziós Raman vizsgálatok eredményeinek összegzése

A mintasorozatokat összehasonlítva megállapítható, hogy a legjobb modelleket az előzőekhez hasonlóan a nagy mérési idejű, nagyobb bemért anyagmennyiséget tartalmazó kapszulák alkalmazásával értem el. A kevesebb bemért anyagmennyiséget tartalmazó mintasorozatok eredményei kis mérési időnél messze elmaradnak a nagyobb mérési idővel mért társaiktól. A kapszulák anyaga itt sem befolyásolja a modellek eredményeit jelentősen, a mérési idő és az akkumulációk hatása jelen esetben sokkal nagyobb.

A két spektrális tartomány eredményeit összevetve látható, hogy az általam kiválasztott hullámszám tartományok alkalmazása nem mindig okozza a modellek eredményeinek javulását. A modellek érzékenysége kis mérési időre sokkal inkább megfigyelhető, mint a teljes tartományt felhasználó modelleknél. Ilyen mértékű romlást a teljes spektrális tartomány felhasználásával készült elemzés zajra vonatkozó érzékenysége a modellparaméterek értékeiben nem okozott. Teljes tartományt alkalmazva valószínűleg lecsökken a zajra való érzékenység is. Nagyobb mérési idővel és több akkumulációval történő mérésnél azonban célszerűbb a redukált spektrális tartomány használata. A PLS modell (vagy a faktoranalízises modellek: PCR, PLS-GA) alkalmazása mindkét esetben jó választás.

4.3.2. FT-Raman spektrometriás vizsgálatok

A famotidine kétkomponensű keverékeket FT-Raman spektrometriai módszerrel is megvizsgáltam, hogy összehasonlíthassam a két módszert. Az FT-Raman eredmények értékelésénél is a 4.3. fejezet elején bemutatott összetételeket vizsgáltam, s az ott bemutatott kalibráló és validáló halmaz-felosztás alkalmazásával. A regressziós modelleket az ott ismertetett előzetes adatkezelési eljárásoknak vettem alá, továbbá a legjobb modellt is ugyanazon alapelvek szerint választottam ki. Így a két eltérő spektrometriai módszer összehasonlíthatóvá válik, s a különbségek várhatóan legnagyobb részben a kétféle mérési elrendezés (reflexiós illetve transzmissziós) közötti különbségeket fogják tükrözni. Ebben az esetben is azonban csak a fontosabb mérési paramétereket, illetve csak a famotidine Form A polimorf módosulatának eredményeit mutatom be a dolgozat fő részében. A teljes táblázatokat a Függelék tartalmazza.

Teljes spektrális tartományt felhasználó regressziós modellek eredményei

A teljes spektrális tartományt felhasználó regressziós modellek eredményeit mutatja be a 4.3.9. táblázat.

4.3.9. táblázat: A teljes spektrális tartományt felhasználó regressziós modellek eredményei

	PLS Form A	PCR Form A	CLS Form A	SVM Form A	PLS-GA Form A
Komponens	1	1	-	15	1
RMSEC	1,1579	1,1640	1,1585	0,3765	0,8558
RMSECV	1,3185	1,3143	1,3164	1,7621	0,9627
RMSEP	2,1523	2,148	2,1706	2,5471	3,4049
R ² Kal	0,9999	0,9990	0,9990	0,9999	0,9994
R ² CV	0,9988	0,9988	0,9988	0,9978	0,9993
R ² Pred	0,9992	0,9992	0,9992	0,9993	0,9996
1. komp%	89,53	89,53	89,53	-	90,84

A 4.3.9. táblázatban látható, hogy a legjobb regressziós modellek a faktoranalízisen alapuló modellek lettek. A PCR modellezés pár esetben jobb paraméterekkel rendelkezik, de a szakirodalomban elterjedtebb PLS robusztusabb modelleket szolgáltat, ezért ezt választottam legjobb modellnek. A CLS kiértékelés eredményei nem sokban maradnak el az előző két módszertől. Az SVM regresszió túlilleszt: látható, hogy a kalibráló egyenes RMSEC értéke a

legjobb az összes modell közül, azonban az RMSEP érték elmarad az előző modellekétől. A genetikus algoritmusokkal támogatott PLS regresszió szolgáltatja a legjobb RMSEC és RMSECV értékeket, azonban a legrosszabb RMSEP paraméter is ehhez a modellhez tartozik.

Redukált tartományt felhasználó modellek eredményei

A redukált spektrális tartomány, ahogy azt már a 4.3. fejezet elején bemutattam, nem tartalmazza csak a famotidine polimorf módosulataira egyedileg jellemző két kiválasztott Raman-sávot. Az ezeket tartalmazó spektrális tartományok felhasználásával készült modellek eredményeit mutatja be a 4.3.10. táblázat.

4.3.10. táblázat: A redukált spektrális tartományt tartalmazó modellek eredményei

	PLS Form A	PCR Form A	CLS Form A	SVM Form A	PLS-GA Form A
Komponens	2	2	-	13	2
RMSEC	1,07054	1,07836	1,28045	1,07626	1,11807
RMSECV	1,3381	1,30628	1,4756	1,43207	1,3794
RMSEP	1,2078	1,22086	1,3662	1,1672	1,5643
R ² Kal	0,9992	0,9992	0,9988	0,9992	0,9991
R ² CV	0,9987	0,9988	0,9985	0,9985	0,9986
R ² p	0,9996	0,9996	0,9996	0,9997	0,9997
1. komp%	91,24	91,24	91,24	-	97,7

A 4.3.10. táblázatban látható, hogy a legjobb modellparamétereket az SVM modellezés szolgáltatta, igaz ezt 13 támogató vektor felhasználásával érte el, ami túlságosan sok, azaz a modell túlillesztett. Alig marad el tőle a PLS regresszió, ezért ezt a két módszert választottam a legjobbnak. Kicsit rosszabb a PCR regresszió, ám ennek eredményei sem különböznek jelentősen az előbbi két módszertől. A CLS és PLS-GA módszerek kis mértékben, de elmaradnak az előbb bemutatott három modell hatékonyságától.

4.3.3. A rangszám különbségek abszolútérték-összeg módszerének (SRD) eredményei

Az előző fejezetben a regressziós módszerek összehasonlítását a teljesítményjellemzőik alapján végeztem el. Ennek hátránya, hogy szubjektív, a kiválasztott modell a készítő személy szaktudásától, tapasztalatától függ: a modellparaméterek értékei nem adnak egyhangú képet arról, hogy melyik a legjobb modell. Ezért dolgozatomban azokat a modelleket választottam

elsőrangúnak, amelyek a legtöbbször tartalmazták a legjobb modellparamétereket. Igyekeztem az összes paraméter figyelembe vételével döntést hozni.

A modellek összehasonlítása elvégezhető matematikai úton is, egy új, magyar fejlesztésű nemparaméteres statisztikai eljárással. A módszer neve rangszám különbségek abszolútérték-összegének (SRD) módszere. Az eljárás a bemért és számított koncentrációk közvetlen összehasonlításával, nem pedig a teljesítményjellemzők összevetésével állít fel sorrendet a modellek között. Jelen esetben ez a következő feltevést jelenti: minél jobb egy modell előrebecslési képessége, annál kisebb lesz az adott módszerre vonatkozó és a referencia sorszámok különbsége.

A referencia ebben az esetben a bemért Form A polimorf koncentrációja volt. Dolgozatomban a modelleket ezzel a módszerrel is összehasonlítottam. Az általam legjobbnak talált modelleket és az SRD eredményeit összevetettem. Az SRD módszert a kalibráló és a validáló, illetve az ismételt mérés eredményeit tartalmazó halmazon végeztem el minden esetben. Itt is csak a famotidine Form A polimorfra számolt eredményeket közlöm. Az előző fejezetekkel ellentétben szürkével jelöltem az SRD módszer szerint legjobb modelleket, az előző fejezetekben általam legjobbnak ítélt eredményeket pedig félkövér betűtípussal jelöltem. Így a kiválasztott rendszerek összehasonlíthatóvá válnak az SRD módszer eredményeivel.

Transzmissziós Raman spektrometria

Előzetes összehasonlító vizsgálatot végeztem a teljes spektrális tartományt felhasználó modellek eredményeiből, hogy képet kapjak a módszer működéséről, illetve a regressziós kiértékelésekről. Az eredményeket mutatja be a 4.3.11. táblázat.

4.3.11. táblázat: Előzetes összehasonlító vizsgálat a teljes spektrális tartományt felhasználó modellek esetében

	AM15s100	AL5s50	TL1s1	TL15s100	TM1s1	TM15s100
PLS	18	30	92	42	74	24
PCR	18	30	118	40	88	24
CLS	32	42	116	46	94	24
SVM	30	38	82	38	104	16
PLS-GA	44	44	152	48	124	38
Egyvált.	126	42	218	62	212	38

A 4.3.11. táblázatban látható az összehasonlítás eredménye. A referenciától való legnagyobb eltérés hatból öt alkalommal az egyváltozós regressziónál volt tapasztalható. Sötétszürkével jeleztem azokat a rangszámokat, amelynél valamely regressziós modell ugyanolyan, vagy rosszabb rangszámot adott, mint az egyváltozós regresszió. A 4.3.2. fejezetben leírtak és az itt tapasztalt értékek miatt döntöttem úgy, hogy az egyváltozós regresszió eredményeit csak a Függelék tartalmazza részletesen.

A teljes spektrumtartományt felhasználó modellek eredményeinek SRD módszerrel való összehasonlítását mutatja be a 4.3.12. táblázat.

4.3.12. táblázat: A teljes spektrumtartományt felhasználó modellek rangszámai mintasorozatok szerint

	AM15s100	AL5s50	TL1s1	TL15s100	TM1s1	TM15s100
PLS	18	30	92	42	74	24
PCR	18	30	118	40	88	24
CLS	32	42	116	46	94	24
SVM	30	38	82	38	104	16
PLS-GA	44	44	152	48	124	38

A 4.3.12. táblázatban látható, hogy az SRD a PLS és PCR módszert három esetben ítélte a legjobbnak, együttesen az AM15s100 és AL5s50 kapszuláknál, a PLS módszert egyénileg pedig a TM1s1 mintáknál. A fennmaradó három sorozatnál (TL1s1, TL15s100 és TM15s100) pedig az SVM regresszió eredményei állnak legközelebb a bemért koncentrációkhoz. Az eredmények a vizsgált rendszerek kétharmadában megegyeznek az általam legjobbként kiválasztott modellekkel. Ugyanakkor a táblázatban is megfigyelhető, hogy a sokszor általam túlillesztettnek gondolt modelleket jónak ítéli meg a módszer.

A mintasorozatok között sorrendet állítottam fel, hogy kiválasszam az optimális mérési időt, illetve akkumuláció számot, illetve összehasonlítsam, hogy a fehér, titándioxidot tartalmazó, átlátszatlan kapszulák mennyiben befolyásolják a kapott regressziós modellek eredményeit. Minden mintasorozatnál az általam és az SRD által legjobbnak ítélt modelleket együttesen figyelve a csökkenő rangszámok alapján állítottam fel a rangsort.

A sorrend: 1. AM15s100, 2. TM15s100, 3. AL5s50, 4. TL15s100, 5. TM1s1, 6. TL1s1

Látható, hogy a nagy mérésidőjű és több bemért anyagmennyiséget tartalmazó kapszulák felhasználásával készült modellek (AM15s100 és TM15s100) kapták a legkisebb rangszám

abszolútérték-összeget. Őket követte a másik átlátszó kapszulába töltött mintasorozat (AL5s50), ezek után pedig a titándioxid-tartalmú kapszulák következtek. A kis mérésidővel készült modelleket (TM1s1, TL1s1) jellemzik a legnagyobb rangszámok. Elmondható, hogy a hosszabb mérési idő jobb modelleket eredményez, a kapszula anyagának nincs jelentős hatása a pontos koncentráció meghatározására. Nem meglepő módon tehát a nehezen mérhető, kevés bemért anyagmennyiséget tartalmazó kapszulák felhasználásával készült modell lett a legrosszabb.

A redukált spektrumtartományt felhasználó modellek eredményeinek SRD módszerrel való összehasonlítását mutatja be a 4.3.13. táblázat.

4.3.13. táblázat: A redukált spektrumtartományt felhasználó modellek rangszámai

	AM15s100	AL5s50	TL1s1	TL15s100	TM1s1	TM15s100
PLS	28	36	150	40	156	38
PCR	28	34	158	32	154	38
CLS	28	42	184	34	128	38
SVM	28	22	122	52	108	44
PLS-GA	32	26	160	64	198	36

A 4.3.13. táblázat adatait tanulmányozva látható, hogy a redukált spektrális tartományt felhasználó modelleknél a rangszámok is nagyobb szórást mutatnak. Az AM15s100 mintasorozatnál ugyanolyan jónak ítélte a modellt a PLS, PCR, CLS, SVM regressziót. Ez nem meglepő, hiszen a modellparaméterek is nagyon hasonlóak. A legjobbnak az SVM regresszió bizonyult: az AM15s100, az AL5s50, a TL1s1 és a TM1s1 mintasorozatok esetében is első helyen szerepelt. Látható, hogy a kisebb bemérési anyagmennyiségű, rövid mérési idővel és kevés akkumulációval vizsgált kapszuláknál a rangszámok kiugróan magasak, a felhasznált modellek SRD módszer szerint sem adnak megbízható eredményeket.

Itt is sorrendet állítottam fel a mintasorozatok között az előbb bemutatott módon.

A sorrend: 1. AM15s100, 2. AL5s50, 3. TM15s100, 4. TL15s100, 5. TM1s1, 6. TL1s1

Az eredmény az előzőekben tárgyaltakhoz hasonlóan alakult. A nagyobb mérési idő és több akkumuláció használata itt is jobb modelleket eredményez. A kapszulák anyaga jelen esetben sem számít igazán, az AL5s50 és a TM15s100 rangszámai között nincs nagy

különbség. Sokkal szembetűnőbb a két legnagyobb rangszámot kapott mintasorozat leszakadása a többitől.

A két spektrális tartományt összehasonlítva megállapítható, hogy a teljes spektrális tartományban készült modellek rangszámai közelebb esnek egymáshoz, nem mutatkozik akkora szórás a mintasorozatok között, mint a redukált tartományokat felhasználók esetében. A teljes tartományban készült modellek stabilabban közelebb állnak a bemért koncentrációkhoz, hiszen rangszámaik kisebbek. A modellek pedig robusztusabbak szintén az alacsonyabb rangszámaik miatt.

FT-Raman spektrometria

A modelleket az FT-Raman spektrometria esetében is összehasonlítottam az SRD módszerrel, hogy ennek segítségével is kiválaszthassam a legjobb regressziós modellt, és azt összevethessem a saját tapasztalataimmal. A teljes tartomány eredményeit mutatja be a 4.3.14. táblázat.

4.3.14. táblázat: A teljes spektrális tartományt felhasználó modellek összehasonlítása SRD módszerrel

	FT-Raman
PLS	6
PCR	6
CLS	6
SVM	6
PLS-GA	8

A 4.3.14. táblázatban jól látható, hogy a hatból öt módszer azonos rangszámot kapott, és amelyik nem (PLS-GA), az sem marad el jelentősen a többitől, tehát az SRD módszer ez esetben nem tesz különbséget a módszerek között. Ezt a megerősíti a modellparaméterek vizsgálata is, ahol csak igen kis különbségek tapasztalhatóak a modellek között.

A redukált tartomány eredményeit mutatja be a 4.3.15. táblázat.

4.3.15. táblázat: A redukált spektrális tartományt felhasználó modellek eredményei

	FT-Raman
PLS	6
PCR	6
CLS	4
SVM	4
PLS-GA	6

A 4.3.15. táblázat alapján a redukált spektrális tartományt felhasználva sem nagyobbak a különbségek a modellek között, ám nagyobb differenciálódás figyelhető meg. A legjobb modellt a CLS és az SVM regresszió adta, őket nem sokkal elmaradva követik a faktoranalízisen alapuló módszerek (PLS, PCR, PLS-GA).

A dolgozatban ismertetett és az SRD módszer összehasonlítása

A 4.3.3. fejezet táblázatait tanulmányozva látható, hogy az általam legjobbnak talált és az SRD módszer által legkisebb rangszámúnak számolt modellek eredményei nem minden esetben egyeznek meg. Ennek több oka van:

1. A modellek közötti különbségek sok esetben nem számottevőek. Az általam kiválasztott modell nem nagyságrendekkel jobb a többinél, azonban, ha kis mértékben is, de a legjobb modellparaméterekkel rendelkezik. Így több esetben is előfordult, hogy a majdnem ugyanolyan jól teljesítő modellek közül az SRD módszer egy másik modellt választott ki. Az SRD eredményeit tanulmányozva ugyanerre a következtetésre jutottam: a rangszámok közötti különbség kevés esetben jelentős (százas nagyságrendű). Az SRD módszer esetében így esetleges, hogy melyik módszer kerül első helyre. Ezen a problémán segíthet az irodalomban megtalálható keresztellenőrzéses SRD. Ez a módszer képes a rangszámok intervallumának becslésére, és képes előrejelezni, hogy az általa adott eredmények statisztikailag szignifikánsak-e. Sajnos csak az alap (keresztellenőrzés nélküli) változat állt rendelkezésemre, de a későbbiekben tervezem a keresztellenőrzéses SRD módszerrel való összehasonlítást is.
2. Az SRD módszer szerint a legjobbnak ítélt modell nem feltétlenül jelenti azt, hogy az a modell a legjobb. A legkisebb rangszám azt jelenti, hogy az adott módszer áll legközelebb a referenciához. Előfordulhat azonban, hogy megjelenik egy olyan

effektus, ami miatt a mérési eredmények közel esnek a referenciához (például a modell túlilleszt), abban az esetben is kis rangszámot kap a modell. A módszer eredménye így nagyban függ a referencia megválasztásától.

Összességében tehát elmondható, hogy az SRD módszer alkalmas a különféle regressziós módszerek közül a legjobb kiválasztására, ám eredményeit kritikával kell fogadni. Erre megoldást jelenthet a későbbiekben a keresztellenőrzéses SRD módszer felhasználása.

4.4. A transzmissziós és reflexiós Raman spektrometria összehasonlítása SRD módszerrel

Az eddigi fejezetekben a célom a legjobb modellek kiválasztása volt. Ezt a modellek egyenkénti értékelésével, illetve az SRD módszer felhasználásával valósítottam meg. Ennek keretében kiválasztottam kétféle spektrális tartomány felhasználásával készült modellek, különböző mérési körülmények, illetve különböző kapszulákba töltött anyagmennyiségek bemérésével készült modellek közül a legjobbakat. Ugyanezt a kiválasztási folyamatot alkalmaztam a reflexiós Raman spektrometria esetében is. Az így kiválasztott modelleket használtam fel arra, hogy a kétféle Raman berendezés-elrendezést összehasonlíthassam. Ezért kiválasztottam a transzmissziós Raman spektrometria legjobb eredményt elért átlátszó, illetve fehér kapszulába töltött modelleket mind a két spektrális tartomány esetében, majd összehasonlítottam a reflexiós Raman spektrometria eredményeivel. A dolgozat fő célkitűzése, hogy eldönthessem, a transzmissziós Raman spektrometria gyorsabb, megbízhatóbb alternatívája-e a reflexiós Raman spektrometriának.

A reflexiós Raman spektrometria csak a minta felületéről egy adott – a minta teljes térfogatához képest elenyészően kis – térfogatról szolgáltat információt. Ezért ahhoz, hogy reprezentatív becslhető legyen a minta koncentrációja, ahhoz a minta különböző helyein kell méréseket végezni, ami jelentősen megnöveli a mérési időt. A mintavételi pontok optimális számát pedig sok esetben nem lehet egyértelműen meghatározni. A reflexiós Raman spektrometriai úton történő mennyiségi elemzés így körülményessé és időigényessé válik, ami alkalmatlanná teszi a gyártásközi ellenőrzésben való felhasználására.

Ezzel szemben a transzmissziós Raman spektrometria már egy mérés esetében is reprezentatív képet ad a teljes vizsgált mintatérfogatról, ami lerövidíti a szükséges mérési

időt, és csak egy pontban szükséges minden mintát lemérni. Így a transzmissziós Raman spektrometria sokkal gyorsabb a reflexiósnál, és ezért beilleszthető a gyártási folyamatba.

Dolgozatomban az SRD módszer segítségével megvizsgáltam, hogy a transzmissziós Raman spektrometria lényegesen kevesebb idő és energia ráfordítással képes-e olyan eredményeket elérni, mint a reflexiós Raman. Ezeket az eredményeket mutatja be a 4.4.1. táblázat.

4.4.1. táblázat: A transzmissziós és reflexiós Raman spektrometria eredményeinek összehasonlítása

FT-Raman	Teljes tart.	Redukált tart.
PLS	6	6
CLS	-	4
SVM	-	4
Transzmissziós Raman		
Átlátszó k. PLS	0	2
Fehér k. PLS	2	2

4.4.1. táblázatban látható, hogy a transzmissziós Raman spektrometria ha kevéssel is, de jobb eredményt ért el mindkét tartomány esetében (az SRD módszer szerint ez a különbség szinte elhanyagolható, tehát a két módszer ugyanolyan jól teljesített). Ez azt jelenti, hogy a transzmissziós Raman spektrometria már egy mérési pont felhasználásával is képes ugyanolyan jó eredményt elérni a polimorfok mennyiségének becslésében, amire a reflexiós Raman spektrometria csak több mérési pont és ezért hosszabb egy mintára vonatkozó mérési idő felhasználásával képes. Fontos megjegyezni, hogy emellett a transzmissziós Raman spektrometria fehér zselatinkapszulák alkalmazásakor is közel ugyanolyan eredményt szolgáltatott.

5. Összefoglalás

A transzmissziós Raman spektrometria új tudományterületén ez ideig kisszámú eredmény született. A technika előtt azonban biztató jövő áll, mivel magába foglalja a hagyományos (reflexiós) Raman spektrometria előnyeit (gyors és roncsolásmentes minőségi és mennyiségi elemzés, nagy érzékenység), s emellett a módszer reprezentatív képet ad a teljes vizsgált mintatérfogatról, könnyen automatizálható, és könnyen beépíthető a készítmények gyártási technológiájába mint sokoldalúan felhasználható eszköz gyártásközi ellenőrzés megvalósítására. Dolgozatomban ez utóbbit vizsgáltam meg részletesen.

A dolgozatban rutinszerűen alkalmaztam egy a spektrometriában még nem alkalmazott, de nagyon jó eredményeket adó alapvonal-korrigáló algoritmust, az asszimmetrikus legkisebb négyzetek módszerét. A módszert kromatogramok alapvonalának simításához fejlesztették ki, de eredményeim tanulsága szerint kitűnően alkalmazható a Raman-spektrometriában is.

Megállapítottam, hogy a tablettabevonás ipari körülmények között megvalósítható nyommonkövetésére a transzmissziós Raman spektrometria alkalmas. Dolgozatomban Panangin tabletták korpuszok bevonását vizsgáltam. A két különböző bevonattal készült bevonási folyamatok eredményei alapján megállapítható, hogy az intenzitás-bevonási idő összefüggés exponenciális, amely azonban könnyen linearizálható. Egyszerre nyerhetünk információt a bevonás időbeli lefolyásáról, valamint a dózis-egyenletességről. A bevonás nyommonkövetése a kezdeti szakaszban pontatlan a nagy szórás miatt, de a transzmissziós Raman spektrometria kiválóan alkalmazható gyors és roncsolásmentes vizsgálati módszer a bevonás végpontjelzésére. A módszer a későbbiekben alkalmas lehet az eddig szinte egyeduralkodó tömegnövekedés-mérési bevonás végpont-jelzés leváltására, s alkalmazásával eleget tehetünk a PAT irányelveinek.

A famotidine hatóanyag polimorf módosulatainak keverékekben történő meghatározásának eredményeiről elmondható, hogy a transzmissziós Raman spektrometria alkalmas az egyik polimorf módosulat, mint szennyező meghatározására a másik módosulat mellett. Az összes mintasorozatot összevetve a következő eredményekre jutottam. A legjobb modelleket a nagy mérési idővel készült, nagyobb bemért anyagmennyiséget tartalmazó kapszulák alkalmazásával sikerült készítenem. Kevesebb bemért anyagmennyiséget alkalmazva, alacsonyabb mérési időnél messze elmaradnak a mérési eredmények a nagyobb mérési idővel

mért társaikétól. Nagyon fontos megjegyezni, hogy a kapszulák anyaga kevésbé befolyásolja a modellek eredményeit, a mérési idő és az akkumulációk hatása sokkal nagyobb a modellek jóságára nézve.

A modellezéshez felhasznált spektrális tartományok szerepe is jelentős a modellek jósága szempontjából. Az eredmények tanulsága szerint a nem teljes spektrális tartományt felhasználó modellek alkalmazása nem minden esetben eredményezi a modellek eredményeinek javulását. A modellek érzékenysége a kis mérési időre ezen csökkentett tartomány esetében sokkal jobban megfigyelhető, mint a teljes tartományt felhasználó modellek esetében. A teljes spektrális tartományt felhasználó modell kevésbé érzékeny a spektrumokat terhelő zajra. Nagyobb mérési idővel és több akkumulációval történő mérés esetében azonban a csökkentett spektrális tartomány felhasználása jó választás lehet.

A kétféle Raman technikát összehasonlítva a dolgozatom eredményei alapján elmondható, hogy a transzmissziós Raman képes elérni, sőt meghaladni a reflexiós Raman spektrometria által elért robusztusságot és megbízhatóságot. További előnyei révén pedig alkalmassá válik arra, hogy a készítménytechnológiai gyártásban rutinszerűen alkalmazható minőség-ellenőrző berendezés legyen.

6. Köszönetnyilvánítás

Szeretném megköszönni a munkám során kapott rengeteg segítséget a kutatócsoport tagjainak. Külön köszönöm türelmüket és megértésüket, amivel irántam viseltettek a dolgozat megírása közben. Köszönettel tartozom Marosi Attilának a hatalmas mennyiségű famotidine minták elkészítéséért és méréséért. Köszönet illeti Nagy Zsombort a Panangin tabletták bevonásában nyújtott segítségéért. Köszönetemet szeretném kifejezni Dr. Demeter Ádámnak és Dr. Német Zoltánnak a famotidine minták elkészítéséért és az FT-Raman vizsgálatok elvégzéséért. Köszönet illeti a Richter Gedeon Nyrt.-t a Panangin és famotidine mintákért. Köszönöm Dr. Héberger Károlynak, hogy átnézte a dolgozatom egy részét, és építő jellegű észrevételeivel hozzájárult a végső változathoz. Végül, de nem utolsósorban szeretném megköszönni a rengeteg támogatást családomnak, barátaimnak.

7. Függelék

A fejezet tartalmazza a dolgozatban bemutatott mintasorozatok felhasználásával készült modellek összes paraméterét. A mintasorozatok és tartományokat jelölő rövidítések feloldása megtalálható a dolgozat 4.3.1.4. fejezetében a 4.3.6. táblázatban.

AM15s100T:

	PLS		PCR		CLS		SVM		PLS-GA		Int	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Komponens	2	2	2	2	1	1	12	12	2	2	-	-
RMSEC	0,8654	0,8654	0,9037	0,9037	1,7995	1,7995	0,3506	0,3502	0,7982	0,7982	1,4398	1,6133
RMSECV	1,0421	1,0421	1,0417	1,0417	2,0676	2,0676	1,5085	1,1635	1,0252	1,0252	-	-
RMSEP	0,3061	0,3061	0,2954	0,2954	0,9282	0,9282	0,9056	0,8252	0,7484	0,7484	0,9184	1,2775
Bias	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	-0,0460	0,1023	0,0000	0,0000	0,0114	0,0042
CV Bias	0,0754	0,0754	-0,0704	0,0704	0,0264	-0,0264	0,1322	0,2142	-0,1229	0,1229	-	-
Pred Bias	0,1870	-0,1870	0,1826	-0,1826	0,6886	-0,6886	0,3413	-0,2376	0,0816	-0,0816	0,0105	0,0249
R ² Kal	0,9995	0,9995	0,9994	0,9994	0,9978	0,9978	0,9999	0,9999	0,9996	0,9996	0,9986	0,9982
R ² CV	0,9993	0,9993	0,9993	0,9993	0,9971	0,9971	0,9986	0,9991	0,9993	0,9993	-	-
R ² Pred	1,0000	1,0000	1,0000	1,0000	0,9999	0,9999	0,9998	0,9998	0,9997	0,9997	0,9996	0,9998
1.komp.%	81,67	81,67	81,68	81,68	81,65	81,60	-	-	73,77	73,70	-	-
Rez eloszl.	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés												
RMSEP	1,4929	1,4929	1,4562	1,4562	2,3229	2,3229	2,7456	2,2557	3,5733	3,5733	14,9087	11,2755
Pred Bias	1,0682	-1,0682	1,0005	-1,0005	1,5872	-1,5872	-2,5124	2,0389	3,3460	-3,3460	0,0099	0,0280
R ² Pred	0,9996	0,9996	0,9996	0,9996	0,9984	0,9984	0,9998	0,9998	0,9995	0,9995	0,9995	0,9985

AM15s100R:

	PLS		PCR		CLS		SVM		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Komponens	1	1	1	1	1	1	10	17	2	2
RMSEC	1,3280	1,3280	1,3490	1,3490	1,3288	1,3288	0,3471	0,4921	0,9354	0,9354
RMSECV	1,4954	1,4954	1,4999	1,4999	1,4913	1,4913	1,3532	1,2967	1,1284	1,1284
RMSEP	0,5109	0,5109	0,5058	0,5058	0,5439	0,5439	0,5131	0,6495	0,4228	0,4228
Bias	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	-0,0393	-0,0379	0,0000	0,0000
CV Bias	0,0114	-0,0114	0,0118	-0,0118	0,0099	-0,0099	-0,3821	0,1262	0,0194	-0,0194
Pred Bias	-0,1923	0,1923	-0,1805	0,1805	-0,1996	0,1996	-0,0867	0,0875	-0,0507	0,0507
R ² Kal	0,9988	0,9988	0,9988	0,9988	0,9988	0,9988	0,9999	0,9998	0,9994	0,9994
R ² CV	0,9985	0,9985	0,9985	0,9985	0,9985	0,9985	0,9989	0,9989	0,9991	0,9991
R ² Pred	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9998	0,9999	0,9999
1.komp%	92,3600	92,3600	92,3600	92,3600	92,3600	92,3600	-	-	95,2400	95,2400
Rez eloszl.	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés										
RMSEP	1,4834	1,4834	1,5006	1,5006	1,5139	1,5139	1,1013	1,8560	1,6186	1,6186
Pred Bias	-0,2379	0,2379	-0,2361	0,2361	-0,2401	0,2401	-0,3620	0,9049	0,0454	-0,0454
R ² Pred	0,9991	0,9991	0,9991	0,9991	0,9991	0,9991	0,9998	0,9996	0,9993	0,9993

AL5s50T:

	PLS Form A	Form B	PCR Form A	Form B	CLS Form A	Form B	SVM Form A	Form B	PLS-GA Form A	Form B	Int Form A	Form B
Komponens	2	2	2	2	-	-	12	17	2	2	-	-
RMSEC	1,0329	1,0329	1,2057	1,2057	1,4695	1,4695	0,5886	0,5800	0,8749	0,8749	1,0955	2,0890
RMSECV	1,2725	1,2725	1,4635	1,4635	1,5713	1,5713	1,1475	1,1258	1,3079	1,3079	-	-
RMSEP	1,4960	1,4960	1,4476	1,4476	1,5434	1,5434	1,8924	1,8717	2,3343	2,3343	1,8232	2,1758
Bias	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0648	-0,0944	0,0000	0,0000	0,0046	0,0204
CV Bias	0,0712	-0,0712	-0,0872	0,0872	0,0757	-0,0757	-0,0094	-0,0572	-0,0777	0,0777	-	-
Pred Bias	-0,6502	0,6502	-0,6249	0,6249	-0,4266	0,4266	-0,5534	0,6100	-0,5309	0,5309	0,0035	0,0263
R ² Kal	0,9993	0,9993	0,9990	0,9990	0,9985	0,9985	0,9998	0,9998	0,9995	0,9995	0,9993	0,9972
R ² CV	0,9989	0,9989	0,9985	0,9985	0,9983	0,9983	0,9991	0,9991	0,9988	0,9988	-	-
R ² p	0,9991	0,9991	0,9992	0,9992	0,9991	0,9991	0,9985	0,9984	0,9984	0,9984	0,9989	0,9992
1.komp%	87,9400	87,9400	87,9400	87,9400	87,9300	87,9300	-	-	84,3200	84,3200	-	-
Rez eloszl.	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	rand	nem rand	nem rand	nem rand	nem rand
Ismétlés												
RMSEP	0,9974	0,9974	1,0341	1,0341	1,2353	1,2353	1,1116	0,9749	1,5385	1,5385	1,4596	1,8431
Pred Bias	0,0057	-0,0057	-0,0135	0,0135	-0,0338	0,0338	-0,0595	0,0144	0,1361	-0,1361	0,0066	0,0210
R ² Pred	0,9994	0,9994	0,9994	0,9994	0,9991	0,9991	0,9994	0,9995	0,9987	0,9987	0,9990	0,9981
Új bemérés 1												
RMSEP	0,8987	0,8987	0,9866	0,9866	1,1209	1,1209	1,4840	1,3093	1,2203	1,2203	1,4832	1,5982
Pred Bias	-0,3923	0,3923	-0,3836	0,3836	-0,2851	0,2851	-0,8746	0,8011	-0,8053	0,8053	0,0044	0,0248
R ² Pred	0,9997	0,9997	0,9996	0,9996	0,9997	0,9997	0,9995	0,9995	0,9995	0,9995	0,9995	0,9990
Új bemérés 2												
RMSEP	0,8693	0,8693	0,9673	0,9673	1,0967	1,0967	1,4663	1,2929	1,6880	1,6880	1,6124	1,6978
Pred Bias	-0,4866	0,4866	-0,5088	0,5088	-0,3877	0,3877	-1,0265	0,9659	-0,9845	0,9845	0,0032	0,0250
R ² Pred	0,9997	0,9997	0,9996	0,9996	0,9997	0,9997	0,9996	0,9996	0,9988	0,9988	0,9994	0,9986

AL5s50R:

	PLS		PCR		CLS		SVM Form		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	A	Form B	Form A	Form B
Komponens	2	2	2	2	1	1	12	13	1	1
RMSEC	0,7508	0,7508	0,7549	0,7549	1,7378	1,7378	0,7777	0,7408	1,3142	1,3142
RMSECV	0,9183	0,9183	0,9109	0,9109	1,8341	1,8341	1,0041	1,1195	1,4704	1,4704
RMSEP	1,6814	1,6814	1,6756	1,6756	1,5837	1,5837	1,6266	1,6612	1,2817	1,2817
Bias	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0486	-0,0694	0,0000	0,0000
CV Bias	0,0446	-0,0446	0,0419	-0,0419	0,0715	-0,0715	0,0985	-0,1985	-0,0286	0,0286
Pred Bias	-0,9619	0,9619	-0,9672	0,9672	-0,7224	0,7224	-0,8592	0,8374	-0,6162	0,6162
R ² Kal	0,9996	0,9996	0,9996	0,9996	0,9979	0,9979	0,9996	0,9996	0,9988	0,9988
R ² CV	0,9994	0,9994	0,9994	0,9994	0,9977	0,9977	0,9993	0,9992	0,9985	0,9985
R ² Pred	0,9990	0,9990	0,9990	0,9990	0,9994	0,9994	0,9991	0,9990	0,9995	0,9995
1. komp%	93,8900	93,8900	93,8900	93,8900	93,7500	93,7500	-	-	99,5300	99,5300
Rez eloszl.	rand	rand	rand	rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés										
RMSEP	1,5417	1,5417	1,5009	1,5009	1,5618	1,5618	0,9293	0,9311	1,3025	1,3025
Pred Bias	1,1158	-1,1158	1,0754	-1,0754	0,0369	-0,0369	0,4144	-0,4581	0,2579	-0,2579
R ² Pred	0,9996	0,9996	0,9996	0,9996	0,9985	0,9985	0,9996	0,9996	0,9990	0,9990
Új bemérés 1										
RMSEP	2,4148	2,4148	2,3540	2,3540	1,8705	1,8705	1,6373	1,7573	1,4299	1,4299
Pred Bias	2,2325	-2,2325	2,1695	-2,1695	-0,2475	0,2475	1,3561	-1,4107	1,2089	-1,2089
R ² Pred	0,9995	0,9995	0,9995	0,9995	0,9990	0,9990	0,9995	0,9995	0,9996	0,9996
Új bemérés 2										
RMSEP	2,7650	2,7650	2,6987	2,6987	2,0387	2,0387	1,9258	2,0619	1,6286	1,6286
Pred Bias	2,5989	-2,5989	2,5302	-2,5302	-0,2966	0,2966	1,6753	-1,7294	1,4073	-1,4073
R ² Pred	0,9995	0,9995	0,9995	0,9995	0,9989	0,9989	0,9995	0,9994	0,9996	0,9996

TL1s1T:

	PLS Form A	Form B	PCR Form A	Form B	CLS Form A	Form B	SVM Form A	Form B	PLS-GA Form A	Form B	Int Form A	Form B
Komponens	2	2	2	2	-	-	17	17	1	1	-	-
RMSEC	1,1679	1,1679	3,3685	3,3685	3,1855	3,1855	0,0402	0,0433	3,3332	3,3332	8,7003	8,3136
RMSECV	3,1924	3,1924	4,075	4,075	3,8852	3,8852	3,5759	4,0365	3,9259	3,9259	-	-
RMSEP	4,779	4,779	3,5902	3,5902	3,7476	3,7476	5,0837	5,0838	4,5382	4,5382	11,0602	7,1463
Bias	0	0	0	0	0	0	0,0069	-0,0069	0	0	0,0649	0,0409
CV Bias	0,4818	-0,4818	0,3391	-0,3391	0,4127	-0,4127	0,6949	-1,3552	0,3376	-0,3376	-	-
Pred Bias	-1,1142	1,1142	-1,431	1,431	-1,3879	1,3879	-0,7715	0,7709	0,4406	-0,4406	0,0709	0,064
R ² Kal	0,9991	0,9991	0,9922	0,9922	0,9931	0,9931	1	1	0,9924	0,9924	0,9518	0,9558
R ² CV	0,9938	0,9938	0,9898	0,9898	0,9905	0,9905	0,9937	0,9912	0,9904	0,9904	-	-
R ² p	0,9904	0,9904	0,9944	0,9944	0,994	0,994	0,9889	0,9889	0,9941	0,9941	0,9351	0,9828
1. komp%	52,59	52,59	52,59	52,59	52,56	52,56	-	-	52,81	52,81	-	-
Rez eloszl.	random	random	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés												
RMSEP	2,7486	2,7486	3,0909	3,0909	2,9143	2,9143	2,3809	2,3804	4,7013	4,7013	9,9722	10,8518
Pred Bias	0,724	-0,724	0,6385	-0,6385	0,6448	-0,6448	0,5202	-0,5176	0,7987	-0,7987	0,0644	0,0459
R ² Pred	0,9962	0,9962	0,9957	0,9957	0,9958	0,9958	0,997	0,997	0,9878	0,9878	0,9634	0,9375

TL1s1R:

	PLS		PCR		CLS		SVM		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Komponens	2	2	2	2	-	-	17	17	2	2
RMSEC	3,2936	3,2936	4,8003	4,8003	6,9265	6,9265	0,0133	0,0197	6,3364	6,3364
RMSECV	5,0508	5,0508	5,5137	5,5137	8,3774	8,3774	3,9352	2,1809	8,7611	8,7611
RMSEP	6,5944	6,5944	7,4036	7,4036	8,3807	8,3807	7,0624	6,6987	4,1001	4,1001
Bias	0	0	0	0	0	0	0	0,0021	0	0
CV Bias	-0,1854	0,1854	-0,0194	0,0194	-0,0328	0,0328	1,452	-0,9354	-0,1346	0,1346
Pred Bias	-3,8706	3,8706	-4,7496	4,7496	-5,138	5,138	-2,3912	2,2822	-0,4811	0,4811
R ² Kal	0,9926	0,9926	0,9842	0,9842	0,9681	0,9681	1	1	0,9724	0,9724
R ² CV	0,9828	0,9828	0,9792	0,9792	0,9527	0,9527	0,9909	0,9973	0,9478	0,9478
R ² Pred	0,9917	0,9917	0,9839	0,9839	0,9774	0,9774	0,9966	0,9966	0,9942	0,9942
1. komp%	70,38	70,38	70,42	70,42	70,12	70,12	-	-	68,39	68,39
Rez eloszl.	nem rand	nem rand	rand	rand	nem rand	nem rand	nem random	nem random	nem rand	nem rand
Ismétlés										
RMSEP	6,6226	6,6226	6,7809	6,7809	7,0904	7,0904	6,5524	6,7412	6,7569	6,7569
Pred Bias	-3,7787	3,7787	-4,2186	4,2186	-4,7508	4,7508	-2,7289	2,8242	-1,5707	1,5707
R ² Pred	0,9819	0,9819	0,982	0,982	0,9839	0,9839	0,9816	0,9805	0,9732	0,9732

TL15s100T:

	PLS		PCR		CLS		SVM		PLS-GA		Int	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Komponens	2	2	2	2	-	-	16	17	2	2	-	-
RMSEC	1,0653	1,0653	1,266	1,266	1,4393	1,4393	0,8613	0,608	1,0833	1,0833	1,2693	2,0256
RMSECV	1,3075	1,3075	1,4912	1,4912	1,6427	1,6427	1,218	1,3885	1,5922	1,5922	-	-
RMSEP	1,2733	1,2733	1,3	1,3	1,4085	1,4085	1,1311	1,007	0,9461	0,9461	1,3255	1,6524
Bias	0	0	0	0	0	0	0,1696	-0,1077	0	0	0,0047	0,0228
CV Bias	-0,0484	0,0484	0,021	-0,021	0,0146	-0,0146	0,2043	-0,1001	-0,0859	0,0859	-	-
Pred Bias	-0,2724	0,2724	-0,246	0,246	0,093	-0,093	-0,0049	0,0183	0,1155	-0,1155	0,0006	0,0249
R ² Kal	0,9992	0,9992	0,9989	0,9989	0,9986	0,9986	0,9995	0,9998	0,9992	0,9992	0,9989	0,9973
R ² CV	0,9988	0,9988	0,9985	0,9985	0,9981	0,9981	0,999	0,9987	0,9983	0,9983	-	-
R ² Pred	0,9992	0,9992	0,9991	0,9991	0,9992	0,9992	0,9993	0,9995	0,9995	0,9995	0,9993	0,9986
1. komp%	88,1	88,1	88,11	88,11	88,1	88,1	-	-	86,55	86,55	-	-
Rez eloszl.	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés												
RMSEP	1,2393	1,2393	1,3682	1,3682	1,5389	1,5389	1,158	1,1157	1,47	1,47	1,4099	2,0375
Pred Bias	-0,152	0,152	-0,0463	0,0463	0,2355	-0,2355	0,0438	0,1658	-0,0251	0,0251	0,0013	0,0221
R ² Pred	0,999	0,999	0,9988	0,9988	0,9986	0,9986	0,9991	0,9992	0,9986	0,9986	0,9991	0,9977

TL15s100R:

	PLS		PCR		CLS		SVM		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Komponens	2	2	2	2	-	-	13	14	2	2
RMSEC	1,1087	1,1087	1,3204	1,3204	1,3673	1,3673	0,9788	0,8933	0,9617	0,9617
RMSECV	1,3352	1,3352	1,6443	1,6443	1,5533	1,5533	1,248	1,4421	1,1832	1,1832
RMSEP	1,2952	1,2952	1,1911	1,1911	1,1856	1,1856	1,3692	1,4145	1,1559	1,1559
Bias	0	0	0	0	0	0	0,2517	-0,1831	0	0
CV Bias	-0,084	0,084	-0,1095	0,1095	0,0996	-0,0996	0,1928	0,1944	-0,0094	0,0094
Pred Bias	-0,6085	0,6085	-0,357	0,357	-0,2361	0,2361	-0,0682	-0,0214	0,0305	-0,0305
R ² Kal	0,9992	0,9992	0,9988	0,9988	0,9987	0,9987	0,9994	0,9995	0,9994	0,9994
R ² CV	0,9988	0,9988	0,9982	0,9982	0,9984	0,9984	0,999	0,9986	0,999	0,999
R ² Pred	0,9994	0,9994	0,9993	0,9993	0,9993	0,9993	0,9992	0,9991	0,9994	0,9994
1. komp%	92,77	92,77	92,77	92,77	92,77	92,77	-	-	99,55	99,55
Rez eloszl.	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés										
RMSEP	1,3796	1,3796	1,5256	1,5256	1,4892	1,4892	1,182	1,1288	1,4154	1,4154
Pred Bias	-0,1926	0,1926	0,1331	-0,3309	0,1092	-0,1092	0,0338	0,0487	-0,3794	0,3794
R ² Pred	0,9988	0,9988	0,9986	0,9986	0,9986	0,9986	0,9991	0,9992	0,9989	0,9989

TM1s1T:

	PLS		PCR		CLS		SVM		PLS-GA		Int	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Comp	2	2	2	2	-	-	17	17	2	2	-	-
RMSEC	1,8117	1,8117	2,9763	2,9763	3,3086	3,3086	0,0089	0,0089	3,2185	3,2185	11,8621	5,3339
RMSECV	4,8629	4,8629	4,6654	4,6654	4,7859	4,7859	4,8842	5,35	6,4889	6,4889	-	-
RMSEP	3,0344	3,0344	3,03	3,03	2,8133	2,8133	3,3158	3,2725	2,7507	2,7507	18,3433	6,8126
Bias	0	0	0	0	0	0	0,0016	-0,0001	0	0	0,0895	0,0438
CV Bias	1,2092	-1,2092	0,7927	-0,7927	0,5107	-0,5107	1,2928	-1,4086	-0,0697	0,0697	-	-
Pred Bias	2,5444	-2,5444	2,3221	-2,3221	2,2072	-2,2072	2,9928	-2,9275	2,4049	-2,4049	0,0415	0,0252
R ² Kal	0,9977	0,9977	0,9939	0,9939	0,9925	0,9925	1	1	0,9929	0,9929	0,9215	0,9808
R ² CV	0,9861	0,9861	0,9865	0,9865	0,985	0,985	0,9865	0,9839	0,9711	0,9711	-	-
R ² Pred	0,9989	0,9989	0,9988	0,9988	0,9988	0,9988	0,999	0,999	0,9991	0,9991	0,9585	0,9822
1. komp%	51,12	51,12	51,13	51,13	51,08	51,08	-	-	47,76	47,76	-	-
Rez eloszl.	nem rand	nem rand	nem rand	nem rand	rand	rand	nem rand	nem rand	nem rand	nem rand	nem rand	nem rand
Ismétlés												
RMSEP	8,8757	8,8757	9,6911	9,6911	9,7847	9,7847	13,4824	12,7367	6,2694	6,2694	11,8955	6,8195
Pred Bias	2,4452	-2,4452	2,6404	-2,6404	2,4195	-2,4195	-0,5417	1,0358	2,2698	-2,2698	0,0607	0,0422
R ² Pred	0,9569	0,9569	0,9512	0,9512	0,9515	0,9515	0,8873	0,8983	0,9782	0,9782	0,9485	0,9724
hibás spektrum												
Ismétlés												
RMSEP	4,5076	4,5076	5,9271	5,9271	6,6582	6,6582	12,4333	12,4603	6,2694	6,2694	11,8955	6,8195
Pred Bias	1,1576	-1,1576	1,3393	-1,3393	1,2284	-1,2284	-1,3377	1,3395	2,2698	-2,2698	0,0607	0,0422
R ² Pred	0,9881	0,9881	0,9802	0,9802	0,9758	0,9758	0,9034	0,9028	0,9782	0,9782	0,9485	0,9724
javított spektrum												

TM1s1R:

	PLS		PCR		CLS		SVM		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Comp	2	2	2	2	1	1	11	8	1	1
RMSEC	4,1983	4,1983	5,0362	5,0362	5,0807	5,0807	3,3592	3,0825	5,1126	5,1126
RMSECV	6,334	6,334	5,9949	5,9949	5,8539	5,8539	9,4801	6,2517	5,966	5,966
RMSEP	3,3666	3,3666	4,2331	4,2331	4,0734	4,0734	2,7492	3,0212	9,4475	9,4475
Bias	0	0	0	0	0	0	-0,056	0,207	0	0
CV Bias	0,297	-0,297	0,0394	-0,0394	0,2835	-0,2835	2,0794	0,4417	-0,0809	0,0809
Pred Bias	-0,1432	0,1432	1,3029	-1,3029	1,4434	-1,4434	1,2537	0,035	6,2552	-6,2552
R ² Kal	0,9879	0,9879	0,9826	0,9826	0,9826	0,9826	0,9924	0,9935	0,9821	0,9821
R ² CV	0,9725	0,9725	0,9753	0,9753	0,9766	0,9766	0,9492	0,9733	0,9756	0,9756
R ² Pred	0,9976	0,9976	0,9915	0,9915	0,9928	0,9928	0,9968	0,9978	0,974	0,974
1.komp%	74,7	74,7	74,71	74,71	74,68	74,68	-	-	66,05	66,05
Rez	nem	nem	nem	nem	nem	nem	nem	nem	nem	nem
eloszl.	rand	rand	rand	rand	rand	rand	rand	rand	rand	rand
Ismétlés										
RMSEP	5,6157	5,6157	5,9103	5,9103	6,5759	6,5759	4,43	5,3569	10,8679	10,8679
Pred Bias	1,7038	-1,7038	2,7187	-2,7187	3,1584	-3,1584	1,7079	-0,8918	3,4705	-3,4705
R ² Pred	0,9848	0,9848	0,9834	0,9834	0,9811	0,9811	0,9896	0,986	0,9539	0,9539

TM15s100T:

	PLS		PCR		CLS		SVM		PLS-GA		Int	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Comp	2	2	2	2	-	-	12	12	2	2	-	-
RMSEC	0,9356	0,9356	0,9962	0,9962	1,2245	1,2245	0,6798	0,6291	0,9442	0,9442	0,8554	1,9542
RMSECV	1,4798	1,4798	1,4607	1,4607	1,4497	1,4497	2,2785	1,7319	1,4828	1,4828	-	-
RMSEP	0,4243	0,4243	0,398	0,398	0,2481	0,2481	0,4637	0,4098	0,6439	0,6439	0,8317	2,0364
Bias	0	0	0	0	0	0	0,1334	-0,1524	0	0	0,0068	0,0043
CV Bias	0,0065	-0,0065	0,0279	-0,0279	0,0747	-0,0747	-0,1099	0,0098	0,2113	-0,2113	-	-
Pred Bias	-0,2156	0,2156	-0,201	0,201	-0,0307	0,0307	0,1942	-0,192	-0,4364	0,4364	0,0044	0,0222
R ² Kal	0,9994	0,9994	0,9993	0,9993	0,999	0,999	0,9997	0,9998	0,9994	0,9994	0,9995	0,9975
R ² CV	0,9985	0,9985	0,9985	0,9985	0,9986	0,9986	0,9967	0,9981	0,9985	0,9985	-	-
R ² Pred	1	1	1	1	1	1	0,9999	0,9999	0,9999	0,9999	0,9997	0,9994
1.komp%	84,07	84,07	84,07	84,07	84,06	84,06	-	-	78,63	78,63	-	-
Rez	nem	nem	nem	nem	nem	nem	nem	nem	nem	nem	nem	nem
eloszl.	rand	rand	rand	rand	rand	rand	rand	rand	rand	rand	rand	rand
Ismétlés												
RMSEP	1,0939	1,0939	1,1604	1,1604	1,5156	1,5156	0,9504	0,8642	1,0525	1,0525	1,1168	1,7303
Pred Bias	-0,1428	0,1428	-0,139	0,139	-0,0432	0,0432	0,1243	-0,1321	-0,0336	0,0336	0,0055	0,0205
R ² Pred	0,9993	0,9993	0,9992	0,9992	0,9985	0,9985	0,9994	0,9995	0,9993	0,9993	0,9993	0,9983

TM15s100R:

	PLS		PCR		CLS		SVM		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Comp	1	1	1	1	1	1	16	17	1	1
RMSEC	1,23	1,23	1,2318	1,2318	1,2306	1,2306	0,0416	0,0517	1,2535	1,2535
RMSECV	1,4086	1,4086	1,4046	1,4046	1,4067	1,4067	0,7105	1,0182	1,4038	1,4038
RMSEP	1,2498	1,2498	1,2454	1,2454	1,2733	1,2733	1,4757	1,4509	1,1158	1,1158
Bias	0	0	0	0	0	0	0,0082	0,0083	0	0
CV Bias	0,0178	-0,0178	0,017	-0,017	0,0161	-0,0161	0,0766	-0,3673	-0,0346	0,0346
Pred Bias	-0,9287	0,9287	-0,9251	0,9251	-0,9358	0,9358	0,6649	-0,5185	-0,78	0,78
R ² Kal	0,999	0,999	0,999	0,999	0,999	0,999	1	1	0,9989	0,9989
R ² CV	0,9986	0,9986	0,9986	0,9986	0,9986	0,9986	0,9997	0,9994	0,9986	0,9986
R ² Pred	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9992	0,9992	0,9997	0,9997
1. komp%	93,95	93,95	93,95	93,95	93,95	93,95	-	-	99,58	99,58
Rez	nem	nem	nem	nem	nem	nem	nem	nem	nem	nem
eloszl.	rand	rand	rand	rand	rand	rand	rand	rand	rand	rand
Ismétlés										
RMSEP	1,109	1,109	1,108	1,108	1,11	1,11	0,7622	0,8868	1,2897	1,2897
Pred Bias	-0,2806	0,2806	-0,2783	0,2783	-0,2825	0,2825	-0,2181	0,3545	-0,2656	0,2656
R ² Pred	0,9993	0,9993	0,9993	0,9993	0,9993	0,9993	0,9997	0,9996	0,999	0,999

FT-Raman spektrometria eredményei:

Teljes spektrális tartományt felhasználó modellek eredményei:

	PLS Form A	Form B	PCR Form A	Form B	CLS Form A	Form B	SVM Form A	Form B	PLS-GA Form A	Form B	Int Form A	Form B
Comp	1	1	1	1	-	-	15	13	1	1	-	-
RMSEC	1,1579	1,1579	1,16406	1,1641	1,1585	1,1585	0,3765	0,3578	0,8558	0,8558	1,2626	1,9311
RMSECV	1,3184	1,3185	1,3143	1,3143	1,3164	1,3164	1,7621	1,7658	0,9627	0,9627	-	-
RMSEP	2,1522	2,1523	2,1487	2,1487	2,1706	2,1706	2,5471	2,6382	3,40497	3,4050	2,1542	2,7673
Bias	0	0	0	0	0	0	-0,0341	0,02906	0	0	0,0032	0,0032
CV Bias	-0,0049	0,0049	-0,0044	0,0045	-0,0059	0,00599	0,1454	0,02114	-0,0139	0,0140	-	-
Pred Bias	-1,6967	1,6967	-1,6931	1,6932	-1,7036	1,70367	-1,9125	1,96006	-2,1365	2,1366	-0,0057	0,0028
R2 cal	0,99907	0,9991	0,99907	0,9990	0,99907	0,99907	0,9999	0,9999	0,9994	0,9995	0,9989	0,9977
R2 CV	0,9988	0,9988	0,9988	0,9988	0,9988	0,9988	0,9978	0,9978	0,9993	0,9994	-	-
R2 p	0,9992	0,9993	0,9992	0,9993	0,9992	0,9992	0,9993	0,9993	0,9996	0,9996	0,9989	0,9999
1st comp%	89,53	89,53	89,53	89,53	89,53	89,53	-	-	90,84	90,84	-	-
Rez eloszl	nem random	nem random	nem random	nem random	nem random	nem random	nem random	nem random	nem random	nem random	nem random	nem random

Redukált spektrális tartományt felhasználó modellek eredményei:

	PLS		PCR		CLS		SVM		PLS-GA	
	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B	Form A	Form B
Comp	2	2	2	2	-	-	13	16	2	2
RMSEC	1,0705	1,0705	1,07836	1,07836	1,28045	1,28045	1,07626	1,04295	1,11807	1,118
RMSECV	1,3381	1,3381	1,30628	1,30628	1,4756	1,4756	1,43207	1,3897	1,3794	1,3794
RMSEP	1,2078	1,2078	1,22086	1,22086	1,3662	1,3662	1,1672	1,4758	1,5643	1,5643
Bias	0	0	0	0	0	0	-0,0608	0,03528	0	0
CV Bias	0,00166	-0,0017	0,01886	-0,0189	0,00265	-0,0026	-0,05004	-0,1006	-0,0329	0,03294
Pred Bias	-0,928	0,9288	-0,9162	0,9162	-0,8144	0,8144	-0,9705	1,2817	-1,4213	1,4213
R2 cal	0,9992	0,9992	0,9992	0,9992	0,9988	0,9988	0,9992	0,9992	0,9991	0,9991
R2 CV	0,9987	0,9987	0,9988	0,9988	0,9985	0,9985	0,9985	0,9987	0,9986	0,9986
R2 p	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9996	0,9997	0,9997
1st										
comp%	91,24	91,24	91,24	91,24	91,24	91,24	-	-	97,7	97,7
	nem	nem	nem	nem	nem	nem		nem	nem	nem
rez eloszl	random	random	random	random	random	random	nem random	random	random	random

8. Irodalomjegyzék

- ¹ L. Brillouin *Ann. Phys.* 88, 17 (1922 Paris)
- ² A. Smekal *Naturwissenschaften* 11, 873 (1923)
- ³ C.V. Raman, K.S. Krishnan: A New Type of Secondary Radiation, *Nature* 121, 501-2 (1928)
- ⁴ A. Schawlow, C. H. Townes *Phys. Rev.* 122, 1940 (1958)
- ⁵ T. H. Maiman, *Nature* 187, 493 (1960 London)
- ⁶ G. J. Rosasco, E. S. Etz, W. A. Cassatt *IVth Int. Conf. Raman Spectrosc.* Brunswick USA (1974)
- ⁷ M. Delhay, P. Dhamelinourt *IVth Int. Conf. Raman Spectrosc.* Brunswick USA (1974)
- ⁸ P. R. Griffiths, J. A. de Haseth: Fourier Transform Infrared Spectrometry, *John Wiley & Sons*, (1986 New York)
- ⁹ J. R. Ferraro, K. Nakamoto: Introductory Raman Spectroscopy, *Academic Press*, (1994 San Diego)
- ¹⁰ N. B. Colthup, L. H. Daly, S. E. Wiberley: Introduction to Infrared and Raman spectroscopy, *Academic Press* (1990)
- ¹¹ F. Billes: Rezgési spektroszkópia (*kiadás előtt álló egyetemi jegyzet*)
- ¹² G. Keresztury: Rezgési spektroszkópiai módszerek (IR, Raman), in: I. Bertóti, Gy. Marosi, A. Tóth (Eds.): Műszaki felülettudomány és orvosi biológiai alkalmazásai, *B+V Lap- és könyvkiadó KFT, Budapest*, (2003), 166-186
- ¹³ R.G. Brereton: Chemometrics: Data analysis for the laboratory and chemical plant, *John Wiley & Sons*, (2003) 292-297
- ¹⁴ R. Matousek, M.D. Morris: Emerging Raman Applications and Techniques in Biomedical and Pharmaceutical Fields, *Springer* (2010)
- ¹⁵ B. Schrader, G. Bergmann: Die Intensität des Ramanspektrums Polykristalliner Substanzen. I. Strahlungsbilanz von Substanz und Probenanordnung, *Zeitschrift für Analytische Chemie Fresenius* 225, 230-247 (1967)
- ¹⁶ P. Matousek, N. Everall: Depth profiling in diffusely scattering media using Raman spectroscopy and picosensored kerr gating, *Appl. Spectrosc.*, 59, 200-205 (2005)
- ¹⁷ N. Everall, T. Hahn: Photon migration in Raman spectroscopy, *Appl. Spectrosc.*, 58 591-597 (2004)
- ¹⁸ H. Wikstrom, I.R. Lewis: Comparison of sampling techniques for in-line monitoring using Raman spectroscopy, *Appl. Spectrosc.*, 59 934-941 (2005)
- ¹⁹ P. Matousek, A.W. Parker: Bulk Raman analysis of pharmaceutical tablets, *Appl. Spectrosc.*, 60 1353-1357 (2006)
- ²⁰ N. Everall, P. Matousek: Temporal and spatial resolution in transmission Raman spectroscopy, *Appl. Spectrosc.*, 64 52-60 (2010)
- ²¹ P. Matousek, N. Everall: Dependence of signal depth in transmission Raman spectroscopy, *Appl. Spectrosc.*, 65 724-733 (2011)
- ²² P. Matousek, N. Stone: Prospects for the diagnosis of breast cancer by noninvasive probing of calcifications using transmission Raman spectroscopy, *J. Biomed. Optics*, 12 1-8 (2007)
- ²³ N. Stone, P. Matousek: Advances transmission Raman spectroscopy: a promising tool for breast cancer disease diagnosis, *Cancer Res.*, 68 4424-4430 (2008)
- ²⁴ P. Matousek, A. Parker: Non-invasive probing of pharmaceutical capsules using transmission Raman spectroscopy, *J. Raman Spectrosc.*, 38 563-567 (2007)
- ²⁵ A. Sparén, J. Johansson: Transmission Raman spectroscopy for quantitative analysis of pharmaceutical solids, *American Pharmaceutical Review*, 12 62-71 (2009)
- ²⁶ M. D. Hargreaves, N. A. Macleod: Characterisation of transmission Raman spectroscopy for rapid quantitative analysis of intact multi-component pharmaceutical capsules, *J. Pharm. Biomed. Anal.*, 54 463-468 (2011)
- ²⁷ Y. Lee, J. Kim: Simple transmission Raman measurements using a single multivariate model for analysis of pharmaceutical samples contained in capsules of different colors, *Talanta*, 89 109-116 (2012)
- ²⁸ J. Johansson, A. Sparén: Quantitative transmission Raman spectroscopy of pharmaceutical tablets and capsules, *Appl. Spectrosc.*, 61 1211-1218 (2007)
- ²⁹ C. Eliasson, N. A. Macleod: non-invasive quantitative assessment of the content of pharmaceutical capsules using transmission Raman spectroscopy, *J. Pharm. Biomed. Anal.*, 47 221-229 (2008)
- ³⁰ M. Fransson, J. Johansson: Comparison of multivariate methods for quantitative determination with transmission Raman spectroscopy in pharmaceutical formulations, *J. Chemometrics* 24 674-680 (2010)
- ³¹ A. Aina, M. D. Hargreaves: Transmission Raman spectroscopy as a tool for quantifying polymorphic content of pharmaceutical formulations, *Analyst*, 135 2328-2333 (2010)

- ³² C. M. McGovern, M. D. Hargreaves: Pharmaceutical Polymorphs quantified with transmission Raman spectroscopy, *J. Raman Spectrosc.*, 43 280-285 (2012)
- ³³ M. A. Elbagerma, H.G.M. Edwards: Characterization of new cocrystals by Raman spectrometry, Powder X-ray diffraction, differential scanning calorimetry and transmission Raman spectroscopy, *Cryst. Growth & Design*, 10 2360-2371 (2010)
- ³⁴ M. Kim, H. Chung: Accurate determination of polyethylene pellet density using transmission Raman spectroscopy, *J. Raman Spectrosc.*, 42 1967-1976 (2011)
- ³⁵ M. A. Macleod, P. Matousek: Emerging non-invasive Raman methods in process control and forensic applications, *Pharmaceutical Research*, 25 2205-2215 (2008)
- ³⁶ K. Buckley, P. Matousek: Recent advances in the application of transmission Raman spectroscopy to pharmaceutical analysis, *J. Pharm. Biomed. Anal.*, 55 645-652 (2011)
- ³⁷ A. Palou, J. Cruz: Determination of drug, excipients and coating distribution in pharmaceutical tablets using NIR-CI, *J. Pharm. Anal.*, 2 90-97 (2012)
- ³⁸ S. R. Torres, J. D. P. Ramos: Raman spectroscopy for tablet coating thickness quantification and coating characterization in the presence of strong fluorescent interference, *J. Pharm. Biomed. Anal.*, 41 811-819 (2006)
- ³⁹ T. Sovány, K. Nikowitz: Raman spectroscopic investigation of film thickness, *Polymer Testing*, 28 770-772 (2009)
- ⁴⁰ J. F. Kauffmann, M. Dellibovi: Raman spectroscopy of coated pharmaceutical tablets and physical models for multivariate calibration to tablet coating thickness, *J. Pharm. Biomed. Anal.*, 43 39-48 (2007)
- ⁴¹ S.R. Torres, J.D. P. Ramos: Raman spectroscopic measurement of tablet-to-tablet coating variability, *J. Pharm. Biomed. Anal.*, 38 270-274 (2005)
- ⁴² J. Müller, D. Brock: Prediction of dissolution time and coating thickness of sustained release formulations using Raman spectroscopy and terahertz pulsed imaging, *Eur. J. Pharm. Biopharm.*, 80 690-697 (2012)
- ⁴³ T. De Beer, A. Burggraef: Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes, *Int. J. Pharm.*, 417 32-47 (2011)
- ⁴⁴ K. Héberger: Sum of ranking differences compares methods or models fairly, *TrAC, Trends Anal. Chem.*, 29, 101-109 (2010)
- ⁴⁵ D.E. Bugay, A.W. Newman: Quantitation of cefepime $\times$ 2HCl dihydrate in cefepime $\times$ 2HCl monohydrate by diffuse reflectance IR and powder X-ray diffraction techniques, *J. Pharm. Biomed. Anal.*, 15, 49-61 (1996)
- ⁴⁶ Kemény S., Deák A.: Kísérletek tervezése és értékelése, *Műszaki Könyvkiadó* (2000) 96-97
- ⁴⁷ R.D. Cattell: The scree test for the number of factors, *Multivar. Behav. Res.*, 1, 245-276 (1966)
- ⁴⁸ E.R. Malinowski: Theory of error in factor analysis, *Anal. Chem.*, 49, 606-612 (1977)
- ⁴⁹ A. Elbergali, J. Nygren, M. Kubista: An automated procedure to predict the number of components in spectroscopic data, *Anal. Chim. Acta*, 379 143-158 (1999)
- ⁵⁰ H. Horvai (szerk.), A.P. Borosy, K. Héberger, I. Kolossváry, A. Lengyel, L. Paksy, R. Rajkó, P. Szepesváry: Sokváltozós adatelemzés (Kemometria), *Nemzeti Tankönyvkiadó*, (2001)
- ⁵¹ H. Mark, J. Workman: Chemometrics in Spectroscopy, *Academic Press*, (2007)
- ⁵² P. Geladi, B. R. Kowalski Partial Least-squares Regression: A tutorial, *Anal. Chim. Acta*, 185 1-17 (1986)
- ⁵³ T. Hastie, R. Tibshirani, J. Friedman: The Elements of Statistical Learning: Data Mining, Inference and Prediction, *Springer*, (2009)
- ⁵⁴ U. Thissen, M. Pepers, B. Üstün, W.J. Melssen, L.M.C. Buydens: Comparing support vector machines to PLS for spectral regression applications, *Chemom. Intell. Lab. Syst.*, 73, 169-179 (2004)
- ⁵⁵ B.M. Wise, N. B. Gallagher: Chemometrics Tutorial, *Eigenvector Research Inc*, (2006)
- ⁵⁶ R. Leardi, M.B. Seasholtz, R.J. Pell: Variable selection for multivariate calibration using a genetic algorithm: prediction of additive concentrations in polymer films from Fourier transform-infrared spectral data, *Anal. Chim. Acta* 461, 189-200 (2002)
- ⁵⁷ J.H. Holland: Adaptation in Natural and Artificial Systems, *University of Michigan Press*, (1975)
- ⁵⁸ R. Wehrens, L.M.C. Buydens: Evolutionary optimisation: a tutorial, *Trends Anal. Chem.*, 17, 193-203 (1998)
- ⁵⁹ K. Héberger, K. Kollár-Hunek: Sum of ranking differences for method discrimination and its validation: comparison of ranks with random numbers, *J. Chemom.* 25 151-158 (2011)
- ⁶⁰ Z. Németh, Gy. Csonka Kis: Quantitative determination of famotidine polymorphs: X-ray powder diffractometric and Raman spectrometric study, *J. Pharm. Biomed. Anal.*, 49 338-346 (2009)