

DOKTORI ÉRTEKEZÉS TÉZISEI

Kertész István

Budapest

2021



Magyar Agrár- és Élettudományi Egyetem

Élelmiszerek Minőségi Paramétereinek Matematikai Modellezése
Többváltozós és Spektrális Adatok Alapján

Kertész István

Budapest

2021

A doktori iskola megnevezése: Élelmiszertudományi Doktori Iskola

Tudományága: Élelmiszertudományok

vezetője: Simonné Prof. Dr. Sarkadi Livia
Egyetemi tanár, DSc
Magyar Agrár- és Élettudományi Egyetem
Élelmiszertudományi Intézet
Élelmiszerkémiai és Táplálkozástudományi Tanszék

Témavezetők: Dr. Felföldi József
Egyetemi tanár, PhD
Magyar Agrár- és Élettudományi Egyetem
Élelmiszertudományi Intézet
Élelmiszerpari Méréstechnika és Automatizálás Tanszék

Dr. Baranyai László
Egyetemi tanár, PhD
Magyar Agrár- és Élettudományi Egyetem
Élelmiszertudományi Intézet
Élelmiszerpari Méréstechnika és Automatizálás Tanszék

.....
Az iskolavezető jóváhagyása

.....
A témavezetők jóváhagyása

1. A munka előzményei, célkitűzések

Az élelmiszerek minősége a történelem kezdete óta lényeges tényező volt az emberek számára, ám a minőség vizsgálatának eszközei rengeteget fejlődtek. Az egyik igen fontos elem a minőség megállapításakor a vizsgálat objektivitása, és annak biztosítása, hogy az egységes összehasonlítás lehetséges legyen helytől és időtől függetlenül, lehetőleg a vizsgálatnak a termékre kifejtett maradandó hatásának minimalizálásával. Ez a szempont kedvezően hatott a roncsolásmentes (non-destructive testing, NDT) módszerek térnyerésére. Ezek közül számos a mechanikai rezgések vizsgálatában gyökerezik, ide tartoznak az akusztikus és ultrahangos vizsgálatok.

Célok

Elsődleges célom olyan eljárás kidolgozása volt, amely összetett anyagok, úgyszintén mint az élelmiszerek különböző minőségi jellemzőinek becslését teszi lehetővé roncsolásmentes vizsgálatok eredményei alapján. Disszertáciomban különböző többváltozós matematikai módszerek alkalmazásával igyekszem különböző élelmiszerek egyes tulajdonságairól információt nyerni, melyek összefüggenek azok minőségével.

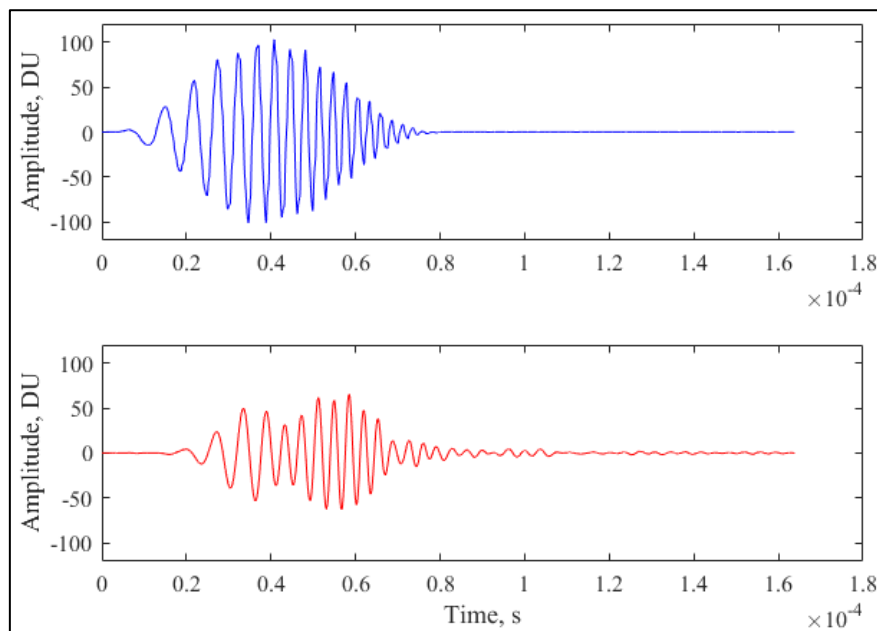
A disszertációban bemutatott kísérletek céljai a következők a következők voltak:

1. Jelfeldolgozási módszer kidolgozása spektrális, főként ultrahang jelekből, a lehető leghatékonyabban történő információkinyerésre, oly módon, hogy az ipari környezetben is alkalmazható legyen.
2. Megfelelő kísérleti elrendezés és kalibrációs valamint becslési eljárás kidolgozása tej oltós alvasztás során bekövetkező viszkozitás-változás közvetett, ultrahang alapú becslésére.
3. Kísérleti elrendezés kidolgozása tojásbél repedtségének detektálására akusztikus válaszjel alapján, és ipari szempontból megbízható osztályozási eljárás kifejlesztése.
4. A kidolgozott jelfeldolgozási eljárás alkalmazhatóságának vizsgálata többcsoportos osztályozási probléma esetén egy élelmiszeripari terméken.

2. Anyag és módszer

2.1. Jelforma és jelelemzés

A bemutatott spektrális vizsgálatok során gyors Fourier-transzformációt (FFT) és folytonos wavelet-transzformációt (CWT) alkalmaztam. Az ultrahangos mérések során, úgynevezett “chirp” jelet használtam, amely egy lineárisan (jelen esetben 50-től 450 kHz-ig) növekvő frekvenciájú pásztázó jel, Hanning-filterrel burkolva. A jellegzetes bemenő és válaszjelre egy példa az 1. ábrán látható. mindkét jel normalizálatlan és zajszerűt.



1. ábra Jellegzetes bemenő (a) és válaszjel (b)

2.2. Tejalvasztási kísérletek

Összesen 13 mérést végeztem 1,5%, 2,8%, és 3,5% zsírtartalmú tejekkel. A mintákba ultrahangos adó-vevő pár került bemeítésre, a mérések 120 percen keresztül történtek 38°C-on, borjúgyomor oltóenzim hozzáadását követően.

A tejalvadás során fellépő viszkozitás-változást HAAKE RotoVisco 1 (Haake Technik GmbH, Németország) rotációs viszkoziméterrel mértem, 38°C-on.

A viszkozitás-változás inverz negatív exponenciális alakot mutatott, ezen görbe illesztése történt mindegyik mérés esetén.

Az illesztés jóságát a korrigált determinációs együttható (R^2_{adj}) jelzi, ennek értékeit az 1. táblázat tartalmazza.

1. táblázat Illesztett viszkozitásgörbe korrigált determinációs együttható értékei

zsírtartalom (m/m%)	R²_{adj}
1,5	0,9985
	0,9948
	0,9980
	0,9792
	0,9984
2,8	0,9862
	0,9510
	0,9973
	0,9858
3,5	0,9869
	0,7251
	0,9101
	0,9917

Az így kapott görbék a továbbiakban az ultrahang jelek alapján történő viszkozitásbecslésben kerültek felhasználásra.

A kísérletek során piezoelektromos ultrahangos adó-vevő párral dolgoztam, melynek névleges frekvenciája 250 kHz volt. (The Ultrason Group, USA). A bemenő és válaszjelek rögzítésére Velleman PCSGU250 oszcilloszkópot használtam (Velleman NV, Belgium), az adatrögzítés percenként történt 120 percen keresztül. A további jelvizsgálatot Matlab 2017a (The Mathworks Inc., USA) környezetben végeztem. A jeleken zajszűrés végeztem, normalizáltam, és az eltolódásból fakadó hiba kizárása okán időben egymáshoz igazítottam.

2.3. Viszkozitásbecslés ultrahangjelek alapján

A viszkozitásgörbék ultrahang alapján történő becslésére CWT algoritmust használtam. Az így kapott együtthatók varianciáját idő függvényében kiszámítottam, és az ebből számított vektormezőn megjelenő csúcsokat azonosítottam. Feltételezve, hogy ezen wavelet-együtthatók jól jellemzik a környező együtthatókat, azok kovariancia-mátrixának számítása felesleges. Az így kiválasztott együtthatók száma 7 és 32 között volt. Ezen együtthatókat parciális legkisebb négyzetek (Partial Least Squares, PLS) regressziója során használtam fel. A PLS modell optimális látensváltozó-számát az átlagos négyzetes hiba gyökének (root mean squared error, RMSE) legalacsonyabb értéke alapján határoztam meg.

2.4. Sajtok osztályozása

Előkísérletek során 60 szeletelt natúr és füstölt mozzarella mintát (Szarvasi Mozzarella Kft., Magyarország) hasonlítottam össze az áthaladás idejének (Time-of-Flight, TOF) új kiszámítási lehetőségeit keresve. A TOF értékét keresztkorreláció, Short-Time Average/Long-Time Average (STA/LTA), és Autoregresszív Akaike Információs Kritérium (Autoregressive Akaike Information Criterion picker (AR-AIC picker) algoritmusokkal számoltam. Ezzel egyidőben a FFET s CWT együtthatók alapján is végeztem csoportosítást, egyenként 66 és 65 együttható felhasználásával. Az alkalmazott osztályozási módszerek között szerepeltek lineáris és nemlineáris módszerek egyaránt: döntési fa, lineáris és négyzetes diszkriminancia-analízis (Linear Discriminant Analysis, LDA, Quadratic Discriminant Analysis, QDA), Support Vector Machine (SVM), K-legközelebbi szomszéd (K-Nearest Neighbors, KNN) és Ensemble osztályozó módszerek.

Az ezt követő kísérletben három különböző típust vizsgáltam, normál, füstölt és laktózmentes típusokat, mivel a korábbi kísérletek nem szolgáltak kielégítő eredménnyel. Ezalkalommal AR-AIC és XC módszereket használtam a TOF számítására, valamint az FFT és CWT együtthatók is felhasználásra kerültek, amelyeket az eredeti jelből, és keresztkorrelált be- és kimenő jelből számítottam. A 2.3 pontban kifejtett eljárást alkalmaztam a leglényegesebb együtthatók kiválasztására.

2.5. Tojánhéj repedtségének detektálása

Hat héten keresztül végeztem akusztikus vizsgálatokat közepes méretű tojásokon, hetente 20 tojáson apró hajszáltrepedést hoztam létre a csúcsukon. Az akusztikus mérést ez előtt és után is elvégeztem, összesen 705 mérési hangfájlt gyűjtve. Magát a mérést a tojás óvatos, üreges fémpálcával történő megütésével végeztem, mely egyetlen ütést jelentett álló (North-South, NS) és elfektetett (East-West, EW) pozíciókban. Stabilizálás és zajcsillapítás céljából habzivacsból készült platformon végeztem a méréseket, amelybe egy mikrofon került beszerelésre. A rögzített hangfájlokat Matlab 2017a környezetben dolgoztam fel, ahol azok Fourier-transzformáltjait különböző ablakméretek, és nullával történő különböző hosszú feltöltések mellett végeztem a felbontásnövelés hatásának vizsgálatára. Az elemzés során a repedés jelenlétét LDA, QDA, SVM, KNN és Ensemble módszerekkel vizsgáltam.

3. Eredmények és értékelésük

3.1. Tejalvadási kísérletek

PLS regressziót végeztem három becslési elrendezésben:

1. Viskozitás becslése az egyes kísérletek esetén
2. Viskozitás becslése különböző zsírtartalmú tejeket csoportjaira
3. Minden mért viszkozitás egyegyidejű becslésére

Az egyedi mérési görbékre történt becslés keresztvalidált R^2_{adj} és maradó becslési eltérés (Residual Prediction Deviation, RPD) mutatószámait a 2. táblázat tartalmazza.

2. táblázat Az egyes viszkozitásgörbék becslésére épített modellek R^2_{adj} és RPD értékei

Zsírtartalom (m/m%)	R^2_{adj}	RPD
1,5	0,9944	5,07
	0,9697	4,49
	0,9876	6,96
	0,9966	10,88
2,8	0,9688	5,20
	0,9758	4,53
	0,9983	14,22
	0,9877	6,23
3,5	0,9632	4,38
	0,9943	10,09
	0,9973	13,99
	0,9684	4,71
	0,9910	8,27

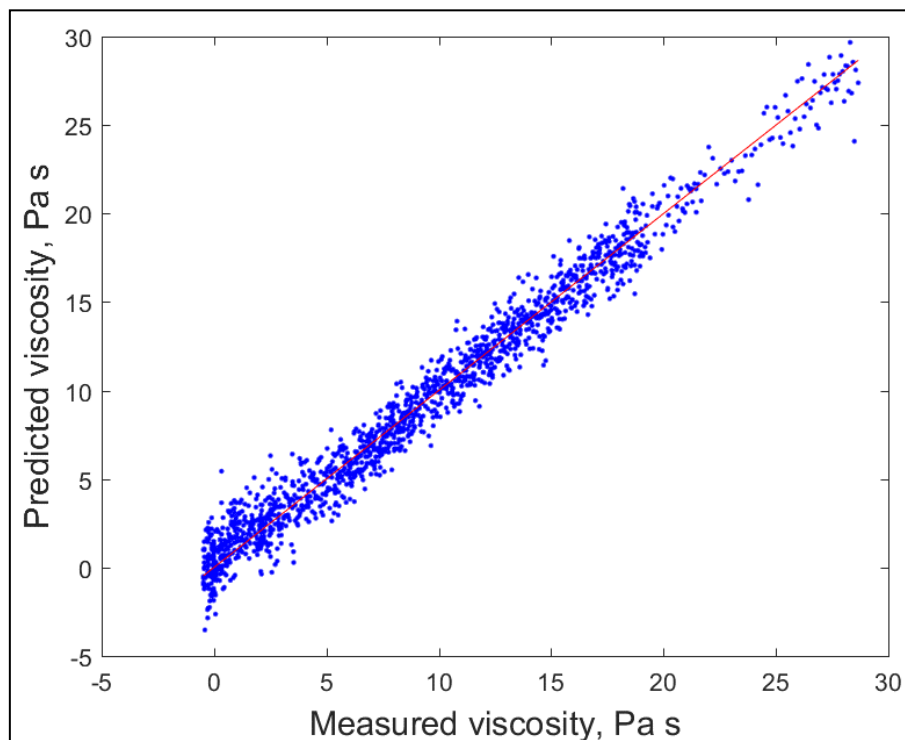
A második becslési sor illesztésének jóságát és pontosságát jellemző mutatószámokat a 3. táblázat tartalmazza.

3. táblázat A második becslési sorhoz tartozó modellek R^2_{adj} és RPD értékei a legalacsonyabb becslési hiába esetén

zsírtartalom (m/m%)	R^2_{adj}	RPD
1,5	0,8929	2,72
2,8	0,9464	2,73
3,5	0,8606	2,51

A harmadik, utolsó becslési sor esetén 135 látens változó mellett a korrigált R^2 értéke 0,9708, az RPD pedig 5,85, ami igen jónak minősül, magas leíróképességgel és becslési pontossággal.

A harmadik becslési sor eredménye a 4. ábrán látható.



4. ábra A harmadik becslési sor modellje 135 látens változóval

3.2. Sajtok osztályozása

Bár a teljes becslési hibák az előkísérletekben magasak voltak, de rámutatott, hogy a FFT-t és CWT-t használni előnyös a továbbiakban. A 4. táblázat összefoglalja az egyes osztályozó algoritmusokhoz tartozó helyes osztályozási arányokat az utófeldolgozási módszerek szerint (keresztvalidált eredmények).

4. táblázat Helyes osztályozási arányok a különböző utófeldolgozási és osztályozási eljárások szerint

	válaszjel-FFT	XC-FFT	válaszjel-CWT	XC-CWT
Döntési fa	97,8%	55,6%	97,8%	95,6%
LDA	96,7%	54,4%	98,9%	96,7%
QDA	84,4%	52,2%	98,9%	97,8%
SVM	96,7%	61,1%	98,9%	97,8%
KNN	97,8%	56,7%	100,0%	97,8%
Ensemble	97,8%	55,6%	100,0%	97,8%

3.3. Tojánhéj repedtségének detektálása

Elsőként egy általános becslés történt minden lehetséges beállítás mellett, hogy átfogó képet kapjunk, melyik módszerek alkalmazását érdemes a továbbiakban megfontolni. A három választott teljesítményt leíró mutatószámra vonatkozó eredményeket az 5., 6. és 7. táblázat tartalmazza.

5. táblázat Osztályozási pontosság (%) a különböző osztályozó algoritmusok szerint. A szintenzitás a jobb becslési pontosság szerint nő.

Szorzó		2 ¹	2 ²	2 ³	2 ¹	2 ²	2 ³	2 ¹	2 ²	2 ³
Ablak méret		2 ¹²			2 ¹³			2 ¹⁴		
Algoritmus	LDA	94,8	94,2	94,5	94,0	94,6	94,6	94,6	95,0	94,7
	QDA	95,5	95,0	94,3	96,3	95,6	95,6	97,2	96,2	97,0
	SVM	95,2	95,9	95,2	95,9	95,3	95,6	96,3	96,3	96,0
	KNN	93,2	92,8	93,3	94,7	94,2	94,9	94,9	94,9	94,7
	Ensemble	94,5	92,5	94,5	94,6	94,6	94,5	94,5	95,0	94,7

6. táblázat Téves pozitív eredmény (%) az egyes osztályozó módszerek esetén. A szintenzitás a jobb becslési pontosság szerint nő.

Szorzó		2 ¹	2 ²	2 ³	2 ¹	2 ²	2 ³	2 ¹	2 ²	2 ³
Ablak méret		2 ¹²			2 ¹³			2 ¹⁴		
Algoritmus	LDA	6	6	6	6	6	6	6	5	6
	QDA	3	3	3	3	3	3	2	3	2
	SVM	5	4	5	3	5	4	3	3	3
	KNN	6	6	6	5	5	4	4	5	5
	Ensemble	4	5	6	5	6	6	6	5	6

7. táblázat Az osztályozás sebessége (s^{-1}) az egyes osztályozó módszerek esetén. A szintenzitás a jobb becslési pontosság szerint nő.

Szorzó		2^1	2^2	2^3	2^1	2^2	2^3	2^1	2^2	2^3
Ablak méret		2^{12}			2^{13}			2^{14}		
Algoritmus	LDA	610	450	130	430	250	220	170	160	91
	QDA	610	500	270	410	210	230	210	100	92
	SVM	780	270	140	480	220	270	260	150	95
	KNN	750	380	140	470	270	260	260	100	91
	Ensemble	460	490	180	300	200	220	120	88	63

Az ablakméret nagyban hozzájárult az osztályozás javulásához, míg a szorzó láthatóan nem. Az osztályozás sebessége igen magas értékeket mutat, ez azonban az ablakméret és a szorzó növelésével egyaránt hamar csökken. a QDA és SVM osztályozók teljesítettek a legjobban, azonban az osztályozóképesség-beli különbség akkor válik igazán egyértelművé, ha az eredményeket megfelelően értelmezzük: minden második, a QDA algoritmus által hibásan épnek vélt tojásra az SVM algoritmus esetén jut egy harmadik hibás tojás is. Ez alapján egyértelmű, hogy a továbbiakban a QDA szigorú tesztelését érdemes elvégezni.

Ennek eredményeként 2,1%-os teljes becslési hibát és 0,87%-os téves pozitív becslési hibát sikerült elérni (8. táblázat) a teljes variancia 99,6%-át leíró 40 látens változó mellett.

8. táblázat Klasszifikációs mátrix a négyzetes diszkriminancia analízis esetén, 40 látens változó mellett.

		valós	
		ép	repedt
becsült	ép	566	6
	repedt	9	112

4. Következtetések és javaslatok

A kidolgozott wavelet-alapú változós szám-csökkentő eljárás hatásossága bizonyításra kerül, különböző anyagok esetén, osztályozási alkalmazás és fizikai jellemző időbeni változásának becslése során is. A lehető legkevesebb korrekció alkalmazása történt a módszer robusztusságának biztosításának érdekében, a való életben azonban a megfelelően adaptált módszer pontossága fontosabb az általános alkalmazhatóságnál, így bizonyos korrekciók alkalmazása javasolt.

A tejalvadási kísérletekben egy bemelegített, átmenő sugárzásos elrendezésben alkalmazott adó-vevő pár került felhasználásra, azonban érdemes lenne vizsgálni „pitch-catch” elrendezésben is az alkalmazhatóságot, hiszen a viszkozitás változása szinte biztosan befolyásolja a felületen történő hullámterjedést, és az alvadó tej állománya erősen befolyásolhatja a jelek viselkedését. Emellett a pitch-catch elrendezés érintésmentes vizsgálatot tesz lehetővé, ami előnyösebb a bemelegített eszközzel vizsgálatnál. A bemutatott ultrahangos módszer megfelelő, alkalmas és javasolt ipari környezetben történő alkalmazásra az élelmiszeripar több területén az adott technológiához történő adaptálás mellett. Elméletben a módszer alkalmas lehet további termékek, úgy mint joghurt, tejföl, puding, stb. állományváltozásának monitorozására.

A tojás héj repedésének vizsgálata igen magas validált pontosságot eredményezett FFT spektrum alapján. Ipari alkalmazásban a legvalószínűbb problémát a külső zajforrások jelenléte okozza, illetve az gerjesztés (megütés) automatizálása. A zajok egy kvázi-hangszigetelt alagúttal részben kezelhetőek, valamint egy nagy érzékenységgű, „shotgun” poláris érzékenységi mintát mutató kondenzátormikrofon alkalmazása javasolt. A kísérleti elrendezésben a habzivacs platform alkalmazásának fő célja a környezeti rezgések kizárása, de más, nem rideg anyagok is alkalmazhatóak a minta gerjesztésének idejére alátámasztás gyanánt.

5. Új tudományos eredmények

1. Új utófeldolgozási eljárást fejlesztettem többdoménes jelek változószámának csökkentésére.

A procedúra hangjelek folytonos wavelet-transzformációján alapszik, amely Morse anyawavelet segítségével idődoménról idő- és spektrumdomén mátrixszá transzformálja a jeleket. Ezen mátrixok a wavelet-együtthatókat tartalmazzák, amelyek a vizsgált mérések szerint egyenként kerülnek elemzésre a legnagyobb szórás-csúcsokat keresve, ezen csúcsok a minták közötti legnagyobb eltérésekre vonatkozó információt tartalmazó együtthatókhoz társíthatók. Az így kiválasztott együtthatók értékei minden mérés adataiból kinyerésre kerülnek, jelentősen csökkentve a további elemzéshez szükséges változók számát, ami a változók óriási száma miatt az eredeti mátrixok alapján nem lenne kivitelezhető.

A javasolt utófeldolgozási eljárás elméletben alkalmazható más többdoménes adathalmaz esetén is, ahol a domének változói folytonosan változtatják értéküket, így a változók halmazából a csúcsok reprezentálhatják a környező adatpontokat, úgy mint térbeli, vagy spektrális adatok esetén.

2. Ultrahangos mérési elrendezés került kidolgozásra a tej enzimes alvadása során jelentkező viszkozitásváltozás meghatározására, amely az alvadás státuszát jellemzi. Az elrendezés átmenő módon alkalmazott, mintába merített, 250 kHz névleges frekvenciájú piezoelektromos adó-vevő párból és egy $4 \cdot 10^{-8}$ másodperces felbontású oszcilloszkópból áll. Az alkalmazott jelforma egy lineáris chirp jel, amely Hanning-szűrővel burkolt, 50-450 kHz között lineárisan modulált pásztázó jel.

Új kalibrációs módszer került kidolgozásra a tej oltós alvasztása során történő ultrahangos viszkozitásmérésére. Az eljárás során az oltós alvasztás során mérhető viszkozitásértékek becslését igényli inverz negatív exponenciális görbével (R^2 : 0.7251 – 0.9985), majd az így számított viszkozitásértékek becslése történt a következő algoritmus alapján. Az ultrahangos jelek normalizáláson, zajscökkentésen estek át, majd wavelet-transzformáció után a mérések mentén, a wavelet-együtthatókra számolt korigált szórások mátrixában kikeresett, nagy információ-tartalmú, releváns csúcsok kinyerésével a viszkozitásértékek becslése történt parciális legkisebb négyzetek regressziójával. A kalibrációs eljárás magas

determinációs együttható (0,9632-0,9983) és maradó becslési eltérés (4,38-14,22) értékeket eredményezett, demonstrálva a módszer pontos becslésre való alkalmasságát.

Új eljárás került kidolgozásra 1,5-3,5%-os zsírtartalmú tejek alvasztásának közvetett ultrahangos méréseinek kalibrációjára. A kalibráció 50-450 kHz közti lineáris chirp jellel végzett mérések válaszjeleivel történt, átmenő módban alkalmazott, mintába merített, 250 kHz névleges frekvenciájú piezoelektromos adó-vevő párral mérve. A normalizált, zajcsökkentett jelek feldolgozása a bemutatott algoritmus alapján történt, és nagy predikciós képességű, jól illeszkedő ($R^2 = 0,9708$) és nagy pontosságú (RPD = 5,85) viszkozitásbecslést tesz lehetővé.

3. Új mérési elrendezés és akusztikus módszer került kidolgozásra tojánhéj repedésének detektálására. Az elrendezés egy habszivacs platformba épített nagy érzékenységgű mikrofonból áll, amelyre a vízszintesen fektetett tojás megütése üreges fémpálcával történt a legnagyobb átmérő mentén, a mikrofon irányával 0° -os szöget bezárva, egyszeri megütéssel. Az akusztikus válaszjel felvétele 96000 Hz-es mintavételezési frekvenciával történt.

A hajszáltrepedések okozta hibás tojások négyzetes diszkriminancia-analízissel kerültek azonosításra 2,1%-os teljes osztályozási hiba, és 0,87% téves elfogadási hiba mellett, negyvenszeres keresztvalidáció alkalmazásával. Az osztályozáshoz felhasznált változók az egyszeri, fektetett pozícióban történt megütés akusztikus válaszjeleinek Fourier-együtthatói voltak. Az együtthatók kinyerése során a megnövelt precizitás a jelek nulla értékekkel végzett, az eredeti jelhossz duplájára történő feltöltésének és 2^{14} hosszúságú Fourier-ablak alkalmazásának eredménye.

4. A bemutatott új változós szám-csökkentési módszer többcsoportos osztályozási probléma esetén is nagy sikerrel alkalmazhatónak bizonyult. Az alkalmazhatóság szemléltetésére három eltérő (normál, füstölt, laktózmentes) típusú szeletelt sajtminthán történt, csupán az átmenő módban chirp jellel történt ultrahangos mérések jelei alapján. A három típus közül kettő megegyező összetételű, és mindhárom eltérő vastagságú és kezelésű volt. Ötszörös keresztvalidáció mellett a bemutatott utófeldolgozási eljárás 100%-os helyes osztályozást eredményezett K legközelebbi szomszéd osztályozási módszert alkalmazva, szemben a Fourier-együtthatók alapján történő 97,8%-os helyes osztályozással, bizonyítva, hogy a módszer javítja az osztályozási pontosságot. A bemutatott módszer abban az esetben is

javította az osztályozási pontosságot a Fourier-együtthatók alapján történő osztályozással szemben, ha a bemenő és válaszjelek keresztkorreláltjai alapján történt a becslés.

6. Kutatási témához kötődő publikációs tevékenység

Kertész, I., Nagy, D., Baranyai, L., Pásztor-Huszár, K., Varsányi, K., Nguyen, L. P. L., Felföldi, J. (2021a): Development of a novel ultrasonic spectroscopy method for estimation of viscosity change during milk clotting. *Molecules*. (accepted for publication)

Kertész, I., Zsom-Muha, V., András, R., Horváth, F., Németh, C., Felföldi, J. (2021b): Development of a Novel Acoustic Spectroscopy Method for Detection of Eggshell Cracks. *Molecules*, 26(15), 4693;

Kertész István, Zsomné Muha Viktória, Zsom Tamás, András Rebeka, Nagy Dávid, Felföldi József (2019): Tojás minőségének és a héj repedésének roncsolásmentes vizsgálata, *Animal welfare, etológia és tartástechnológia*, 15. (1), 19-35.

Kertész, I., Felföldi, J. (2016): Comparison of sound velocity estimation and classification methods for ultrasonic testing of cheese, *Progress in Agricultural Engineering Sciences*, 12 (1), 51-62.

Kertész, I., Simon, N., Kaszab, T., Zsom Muha, V., Felföldi, J. (2015): Ultrasonic method for cheese quality evaluation, In: Walker, J. (Ed.) *American Society of Agricultural and Biological Engineers Annual International Meeting 2015*