



Grafikus felhasználói felület NIR spektrumok statisztikai kiértékeléséhez

Kövér¹ Gy., Bázár² Gy.

¹Kaposvári Egyetem Gazdaságtudományi Kar, Matematika és Fizika Tanszék, 7400 Kaposvár Guba Sándor u. 40.,
²Kaposvári Egyetem Állattudományi Kar, Állattermék Minősítő Laboratórium, 7400 Kaposvár Guba Sándor u. 40.,

ÖSSZEFOGLALÁS

Az R nyílt forráskódú statisztikai programcsomag elsősorban parancssori üzemmódban használható. A bonyolult szintaktikával rendelkező parancsokat grafikus felhasználói felületek rejthetik el a felhasználó elől. Az R Commander kiterjesztését végeztük el úgy, hogy alkalmas legyen NIR spektrumok feldolgozására és kalibrációs becslő egyenletek meghatározására a PLS regresszió eszközével. Az elkészült modulok alkalmasak az egyedi szerkezetű adatfájlok beolvasására, a spektrumok grafikus megjelenítésére, összefüggésvizsgálatokra, valamint a kalibrációs egyenlet előállítására.
(Kulcsszavak: NIR, GUI, PLS regresszió, R)

ABSTRACT

A Graphical User Interface for Statistical Evaluation of NIR Spectra

Gy. Kövér¹, Gy. Bázár²

¹Department of Mathematics and Physics, Faculty of Economic Science, Kaposvár University, H-7400 Kaposvár, Guba S. u. 40.
²Laboratory of Animal Product Qualification, Faculty of Animal Science, Kaposvár University, H-7400 Kaposvár, Guba S. u. 40.

R is a free, open-source statistical environment and does not incorporate a statistical graphical user interface. The commands with complex syntax and hundreds of arguments are supposed to be hidden behind an easy to use GUI. The R Commander provides a basic-statistics which was extended by the authors to implement some functions to analyse NIR spectra and calculate calibration equations using PLS regression. The extensions provides the abilities to read the unique structure input files, graphical presentation of the spectra and calculating calibrations equations.

(Keywords: NIR, GUI, PLS regression, R)

BEVEZETÉS

Az R olyan nyelv és környezet, amely különösen alkalmas matematikai statisztikai számítások elvégzésére, adatelemzésre, az eredmények grafikus prezentálására. További előnyeként említhető, hogy az R szabad szoftver, GNU licenz feltételek mellett terjeszthető. Az R fejlesztését *Ihaka és Gentleman* (1996) kezdte, jelenleg az R Development Core Team (<http://www.R-project.org>) kezeli.

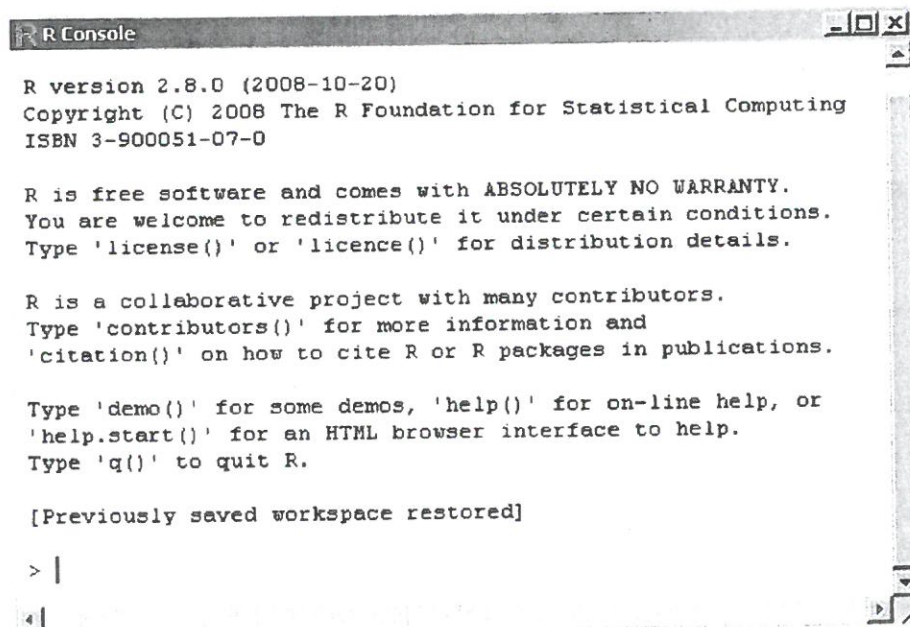
Sajnos az R nem tartalmaz olyan grafikus felhasználói felületet, amely segítségével az R alaprendszerébe épített nagyszámú statisztikai eljárás elérhető. Hasonló a helyzet az egyes fejlesztők által a felhasználók rendelkezésére bocsátott speciális csomagokkal. Az R felhasználója az R-szintaxis szerint megszerkesztett kódokat vagy parancssori vagy script módban juttatja el az R parancs-értelmezőjéhez. (1. ábra)

Az R parancssori használatát, számos munkából megismerhetjük, köztük megemlíthetjük *Solymosi Norbert* (2005) magyar nyelvű bevezetőjét.

Az R nyílt széleskörű alkalmazhatósága, nyílt forráskódja maga után vonta hogy egymással párhuzamosan több grafikus felhasználói felület kifejlesztése is megtörtént. Természetesen nem sorolhatjuk ebbe a kategóriába azokat az R csomagokat, melyek egy-egy szűk területtel foglalkoznak és csupán az eredmények szemléltetését végzik grafikus felületen, még ha interaktív grafikával is rendelkeznek. Ebbe a körbe tartozó példaként említhetjük a teljesség igénye nélkül az *igraph*, vagy a *neural* csomagot.

1. ábra

Az R környezet parancssori üzemmódja



```
R Console
R version 2.8.0 (2008-10-20)
Copyright (C) 2008 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> |
```

Figure 1: The command line interface of the R environment

A grafikus felhasználói felületek a <http://www.r-projekt.org/GUI> hivatkozason megtekinthetők. Kitűnő munka a *Rattle* (*Williams*, 2008), amely az adatbányászat témakörét dolgozta fel, jelenleg is folyamatos fejlesztés alatt áll.

A *SciViews-R* (*Grosjean et al.*, 2005) egy olyan projekt, amely a grafikus felhasználói felület kifejlesztését az általában használatos statisztikai számítások elvégzésének megkönnyítésére végezte el. Ugyanakkor a fejlesztők célkitűzései között az is szerepelt, hogy olyan grafikus felület jöjjön létre, amely nem kizárólag a statisztikai számítások elvégzését teszi lehetővé, hanem a felhasználó a program használata közben „kiképzést” kap az egyes statisztikai parancsok R nyelvű szintaxisából is. Nem elrejti a kezdők számára valóban borzalmasnak tűnő programrészleteket, hanem központi szerepet szán a létrejövő utasítássorozat képernyőn való megjelenítésének.

Az R Commander (Fox, 2005) fejlesztése során a szerző legfontosabb célkitűzései között szerepelt, hogy könnyen használható grafikus felületet készítsen, amelyen keresztül minden olyan statisztikai elemző módszer elérhető, amiket egy kezdőknek szánt statisztikai kurzus tartalmaz. El kívánta rejteni a felhasználó elől az R nyelv parancsutasításait, a bonyolult szintaktikát valamint az R adatszerkezeit.

Jelen dolgozat célja, hogy az R Commander továbbfejlesztési lehetőségeit kihasználva olyan grafikus felhasználói felületet hozzunk létre, amely alkalmas arra, hogy közeli infravörös (NIR) spektrumokat a PLS regresszió segítségével feldolgozzunk. A sikeres fejlesztés során olyan szabadon felhasználható szoftverhez jutunk, amely rugalmasan továbbfejleszhető változó kísérleti körülményekhez igazítva, másfelől tantermi oktatásban is változatosan felhasználható. A kereskedelemben kapható, nem paramétereztető, vagy script nyelven nem programozható spektrumfeldolgozó számítógépes programoknál jobban használható a nyílt forráskódú megoldás.

A kifejlesztett szoftver alkalmazhatóságát mangalica sertésből származó húsminta reflexiós spektrumának feldolgozásával mutatjuk be.

ANYAG ÉS MÓDSZER

A fejlesztői környezet

A NIR spektrumok feldolgozására alkalmas grafikus felhasználói felület kialakítása az R 2.7.0 (2008.04.22) verzióján alapul. A programozás a Tinn-R (1.19) grafikus integrált fejlesztői környezetben készült.

A fejlesztés során két lehetőség kínálkozott. Az egyik az, hogy teljesen önálló modult fejlesztünk ki és R csomagként tesszük használhatóvá. A másik lehetőség az, hogy egy már működő, funkciógazdag grafikus felhasználói felületet fejlesztünk tovább. A mások által készített program továbbfejlesztése azzal a hátránnyal jár, hogy egységes szerkezetéhez, belső adatstruktúrájához kell illeszteni az új funkciót. Előnyként jöhet számításba a felhasználható letesztelt eljárások széles skálája. A bevezetésben említett SciViews-R vagy R Commander továbbfejlesztése egyaránt azt az előnyt kínálja, hogy a szokásos statisztikai számítások már részét képezik e programok grafikus felületének. A felhasználó a spektrumok adatait a PLS regresszió mellett más eszközökkel is feldolgozhatja.

Mivel az R Commander (2. ábra) eredeti célkitűzése az, hogy teljesen elrejtse a felhasználó elől a bonyolult szerkezetű parancsokat, ezért ezt a programot választottuk arra, hogy a spektrumok feldolgozására alkalmas grafikus felhasználói felület céljára továbbfejlesszük.

2. ábra

A továbbfejlesztésre kiválasztott R Commander

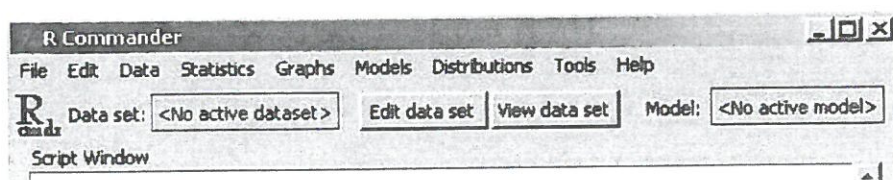


Figure 2: The R commander is chosen to be extended by a plug-in module

Az R Commander a Tcl/Tk nyelvet megvalósító tcltk R csomag lehetőségeit használja fel a grafikus felhasználói felület programozott megjelenítéséhez, az eseményvezérelt feladatok megvalósításához.

A demonstrációs adatbázis

A kifejlesztett grafikus felhasználói felület alkalmazhatóságát olyan adatbázison mutatjuk be, melyet 29 mangalica sertésből származó húsminta reflexiós spektruma és kémiai elemzés eredményeként kapott beltartalmi értékek (szárazanyag-, zsír- és fehérjetartalom) képezik.

Az állatokat hagyományos takarmányozási és tartási körülmények között hizlalták, a vágáskori átlagos testtömeg 157 kg volt. 24 órás hűtést követően a bal oldali hosszú hátizom (m. *longissimus dorsi*) utolsó bordatájékról származó szelete (kb 100g) került vizsgálatra. Minden mintát gondosan megtisztítottunk a kötőszövetől, hogy csak az intramuszkuláris zsírtartalommal kelljen számolni. Az egyes mintákat IKA A11 basic berendezéssel homogenizáltuk majd Christ Alpha fagyasztva szárítóval liofilizáltuk.

A fagyasztva szárított (liofilizált) mangalica húsminták közeli infravörös vizsgálatát NIRSystem 6500 (Foss NIRSystem, Silver Spring, MD, USA) spektrométerrel végeztük el. A reflexiós spektrumokat az 1100-2500 nm-es tartományban rögzítettük (log 1/R), 2nm-es lépésközzel. „Small ring cup” mintatartó küvettát (IH-0307) és „Sample transport” egységet használtunk a vizsgálat során. A műszer üzemeltetéséhez és az elsődleges adatkezeléshez a WinISI II version 1.5 szoftvert alkalmaztuk (InfraSoft International, Port Matilda, PA, USA). A küvettákat minden minta után elmostuk, majd szárazra töröltük.

A kémiai analízis során a liofilizált minták szárazanyag-tartalmát az MSZ ISO 1442 szabvány, a zsírtartalmat *Folch és mtsai* (1957) szerint határoztuk meg. Sósavas emésztést és Kjeld-Foss Fast Nitrogen Analyzer készüléket alkalmaztunk a nitrogén tartalom meghatározására; a nitrogén tartalmat 6,25-dal szorozva fejeztük ki a fehérjetartalom. A beltartalmi értékeket (zsír- és fehérjetartalom) 100% szárazanyagra vonatkoztatva adtuk meg.

A NIR spektrum adatszerkezete

A reflexiós spektrum felvétele során egy minta 1100-2498 nm tartományban szolgáltató mérési adatot.

3. ábra

A PLS regresszió alkalmazására megfelelő adatszerkezet

	y_1	y_2	y_3	x_1	x_n	x_{700}
minta	Száraz- anyag	nyers- zsír	fehérje	1100 nm						2498n m
1										
...										
m										
...										
29										

Figure 3: The data structure required by the PLS regression

A PLS regresszió módszerével olyan kalibrációs egyenletet hozunk létre, amely a minták meghatározott szárazanyag-, zsír- és fehérjetartalmát megfelelő pontossággal becsli (Siesler és mtsai, 2002). A statisztikai modell a 29 elemű minta három függő változója és a spektrumonként 700 reflexiók értékét jelentő független változók (3. ábra) között teremt kapcsolatot a következő formában:

$$Y = XB + \varepsilon, \text{ ahol } \varepsilon \text{ a véletlen hibák mátrixa.} \quad (1)$$

A PLS regressziót tartalmazó csomag Ron Wehrens és Bjørn-Helge Mevik munkája (Mevik-Wehrens, 2007).

A NIR spektrumot tartalmazó adatfájl szerkezete

A NIRSystem 6500 (Foss NIRSystem, Silver Spring, MD, USA) spektrométer adatait a WINISI 1.5 (FOSS TECATOR Intrasoft International, LLC) szoftver segítségével lehet ASCII kódolású szöveges állománnyá alakítani.

Az eredményfájl első tizenkét sora az adat-felvételezés körülményeit jellemző információkat tartalmazza. A PLS regresszió végrehajtásához ezekre az adatokra nincs szükség. Az eredményfájl többi sora tartalmazza a mintákra vonatkozó összes adatot, soronként egy-egy minta szerepel:

- A spektrum értékei 1100-2498 nm tartományban
- A kémiai analízis adatai: szárazanyag-, zsír- és fehérjetartalom
- A minta sorszáma
- A minta egyéb azonosítója

Az adatsorokban az egyes adatelemeket vessző választja el egymástól, a szöveges adatokat idézőjel határolja. Az adatsorokat megelőzi egy oszlopfejléc információt tartalmazó sor.

EREDMÉNYEK

A spektrumfeldolgozás beillesztése az R Commander menüszerkezetébe

Az R Commander menüszerkezete a program futásideje alatt jön létre. A pontos felépítés a „\etc\Rcmdr-menu.txt” szövegfájlban található. A fájl tartalmazza a menüpontok és almenüpontok feliratait, az indítandó függvény nevét.

A „\etc\Rcmdr-menu.txt” szerkezete lehetőséget kínál egy logikai feltétel megadására is, ami a menüpont kiválaszthatóságát engedélyezi. Mint Fox (2005) kifejti a menüpontok kiválaszthatóságát azzal a szándékkal kívánta engedélyezni, illetve tiltani, hogy csak a beolvasott adatok természetének, szerkezetének megfelelő statisztikai funkciókat választhassa ki a felhasználó. Az engedélyezést az „activateMenus” függvény végzi. A NIR spektrumok adatainak beolvasási tesztjei során kiderült, hogy az „activateMenus” függvény végrehajtási időigénye túl sok, interaktív programfelhasználás során nem fogadható el, különösen azért nem, mert a függvényt az R Commander elég gyakran elindítja. Feltételezi ugyanis, hogy a statisztikai adatfeldolgozás közben megváltozhat az adatszerkezet ezért a menüpontok engedélyezhetősége megváltozhat. A NIR spektrumok feldolgozását végző modulok beillesztését követően szükségessé vált az „activateMenus” függvény funkciójának felfüggesztése, hogy az R commander futási sebessége elfogadhatóvá váljon.

A spektrumfeldolgozás beillesztése az R Commander programszerkezetébe

Az R Commander 1.3 (2008.04.09) verziója messzemenőig támogatja a továbbfejlesztetőséget. A forrásprogramot az „R” környezet „\library\Rcmdr\R” mappa tartalmazza. A továbbfejlesztés során készített modulok a „\library\Rcmdr\etc” mappába kerülnek.

A spektrumadatok beolvasása

Az R Commander „Data \ Import data \ from text file or clipboard” menüpontja szöveges adatállományok beolvasását végzi el a „readDataSet” függvény indításával. Bár a „WINIsi” program által szolgáltatott adatok is szöveges formátumban kerülnek exportálásra, nehézséget okoz az adatok előtt elhelyezkedő tizenkét sor amely számunkra nem releváns adatot tartalmaz. A „readDataSet” függvény az „R” „read.table” utasítását használja adatbeolvasásra, amely alkalmas az első néhány felesleges sor átlépésére, de a „readDataSet” ezt a lehetőséget nem használja ki, sőt elfedi. A „readDataSet” függvény célszerű átalakításával a 12 sor átlépése megoldásra került. A spektrumadatok beolvasását végző modul grafikus megjelenése nem különbözik a „readDataSet” függvényétől (4. ábra). Az adatbeolvasás eredményeként a spektrumadatok az „R” környezetben szokásos „data frame” szerkezetbe kerülnek.

4. ábra

A WINISI által szolgáltatott adatállomány beolvasó modulja, grafikus felület.

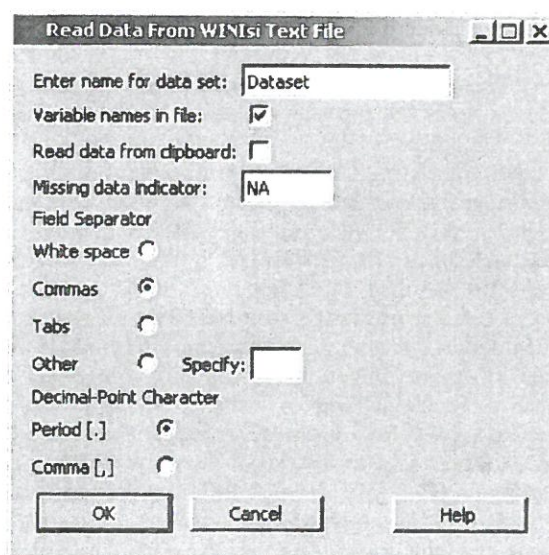


Figure 4: GUI to read the data provided by the WINIsi

A spektrumadatok grafikus megjelenítése

A matematikai statisztikai módszerek alkalmazása előtt ajánlatos az adatok grafikus megjelenítése. Az adatelemzés buktatóinak nagy részét ki lehet kerülni az adatábrázolás segítségével. Az „R” környezet alapvető grafikus lehetőségei a „plot” parancs köré épülnek. A spektrumok ábrázolása a mátrixok ábrázolására létrehozott „matplot” utasítással történik. A „matplot” két márix (X,Y) oszlopvektoraiban tárolt adatok ábrázolását hajtja végre. Jelen esetben Y a spektrumok transzponált adatait jelenti, X a hullámhosszokból álló oszlopvektor. A grafikon további elemei, mint a tengelyek, megjegyzések, feliratok, stb a grafikus rendszer alacsonyabb szintű utasításainak használatával kerülnek beállításra (5. ábra).

5. ábra

A spektrumok grafikus ábrázolása. Az interaktív felület lehetőséget ad egyes minták kiemelésére

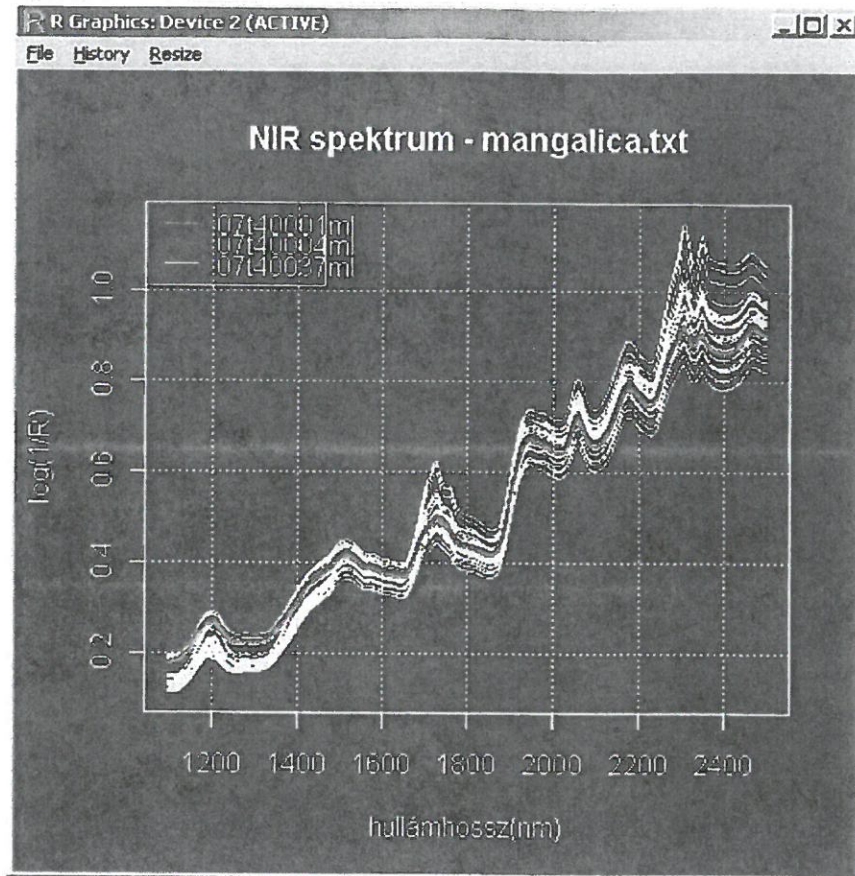


Figure 5: Spectra visualized by the GUI. Interactively selected curves are highlighted.

Statisztikai műveletek és a grafikus megjelenítés

Az „R” környezet előnyei elsősorban a statisztikai számítások területén mutatkoznak meg. A NIR spektrumok elemzése során az általában alkalmazott eszközök egyike a korrelációs számítás. A korrelációs együtthatót a kémiai analízis egy kiválasztott mérési eredménye (pl. szárazanyag) és a spektrumok egy adott hullámhosszon mért értékei között lehet kiszámolni. Az összes hullámhosszon elvégezve a számítást a kapott korrelációs együtthatókból görbe rajzolható (6. ábra). Mivel a szomszédos hullámhosszakon számított együtthatók értéke nyilvánvalóan nem független egymástól ezért az ábrából közvetlen következtetések levonása nem ajánlatos, ugyanakkor hasznos eszköze az adatelemzésnek. A görbék megjelenítése az előzőekben részletezett módon történik.

6. ábra

Korreláció elemzés. A szárazanyag, zsír és fehérje összefüggése az egyes hullámhosszakon mért intenzitásértékekkel

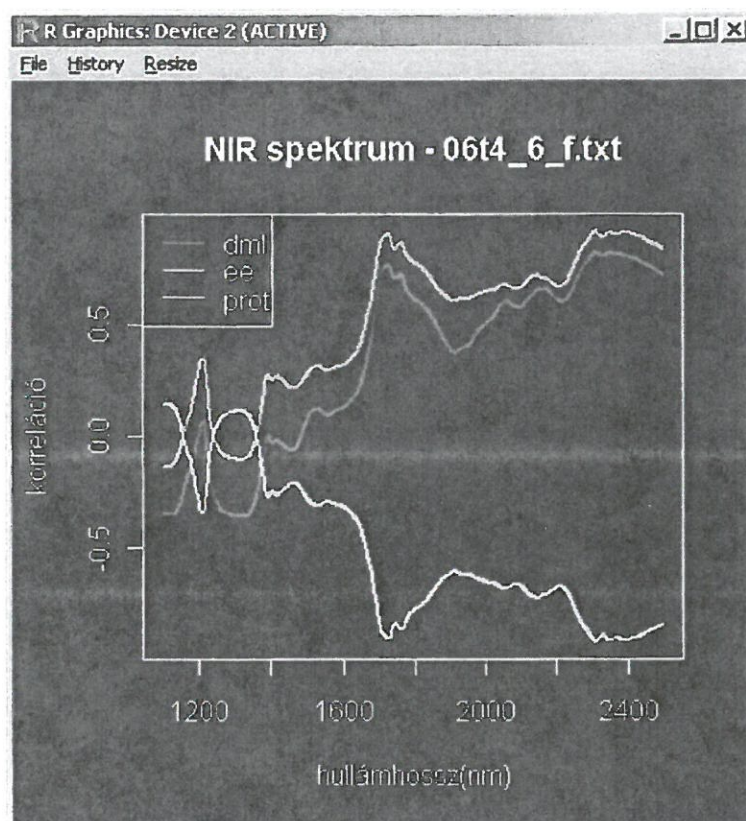


Figure 6: Correlation analysis. Association between dry matter, protein, fat and intensity depending on the wavelength.

A spektrumok matematikai előkezelése

Egyrészt a NIR spektrumok felvételének metodikája, másrészt a jobb kalibrációs egyenlet létrehozásának célja indokolja azt, hogy a spektrumokat matematikai módszerekkel kezeljük még a modellalkotás előtt (7. ábra).

A „R” környezet nyílt, moduláris felépítése lehetővé teszi, hogy mások által csomagokba foglalt, tesztelt eljárásokat alkalmazzunk. A NIR spektrumok előkezelésére a használt algoritmusokat a „pls” (Wehrens, Mevik) és az „RTisen” (Hegger, Kantz, Schreiber) csomagban találhatjuk. Jelen munka keretében kialakított interaktív, grafikus környezet feladatául azt is kiegészítettük, hogy a felhasználó elől elrejtjük azokat az „R” utasításokat is, melyek a csomagok kezelését (telepítés, betöltés) végzik.

A matematikai módszerekkel, mint például a deriválással előkezelt spektrumok grafikus megjelenítése a helyes következtetések levonását segíti elő (8. ábra).

7. ábra

A spektrumok előkezelése matematikai módszerekkel

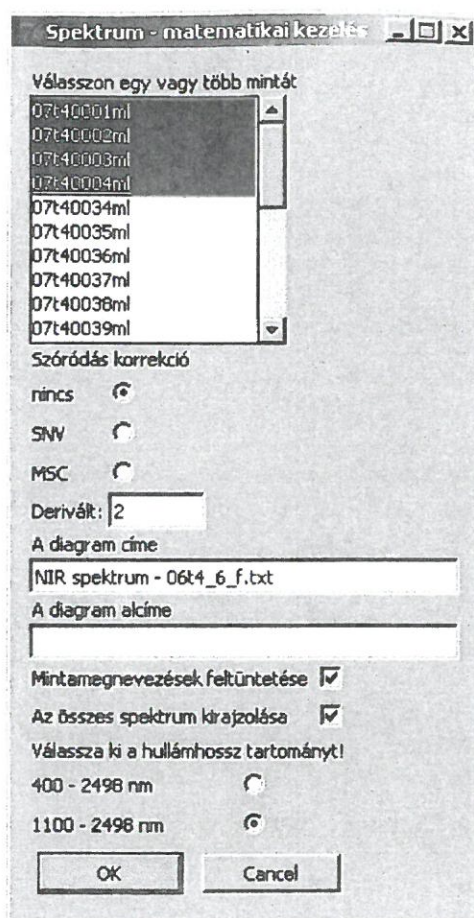


Figure 7: Mathematical treatment of spectra

A kalibrációs egyenlet modellje

Kísérletek során a NIR spektrumok felvételének gyakori célja az, hogy a párhuzamosan végzett kémiai analízis által származtatott beltartalmi értékek becslésére alkalmas, megfelelő pontosságot biztosító kalibrációs egyenletet hozzunk létre.

A NIR spektrumok kiértékelését biztosító grafikus felhasználói felület létrehozásakor kitűzött egyik legfontosabb cél az volt, hogy a felhasználó elől elrejtjük az „R” környezet parancssori használatának kényelmetlenségeit, elsősorban az utasítások szintaktikájának ismeretét tegyük feleslegessé. A kalibrációs egyenlet létrehozását a „pls” (Wehrens, Mevik) csomag „pls” utasításával végezhetjük el. A parancs szintaktikája a gyakorlatlan felhasználó számára áttekinthetetlenül bonyolult (9. ábra).

8. ábra

Négy kiválasztott spektrum második deriváltja

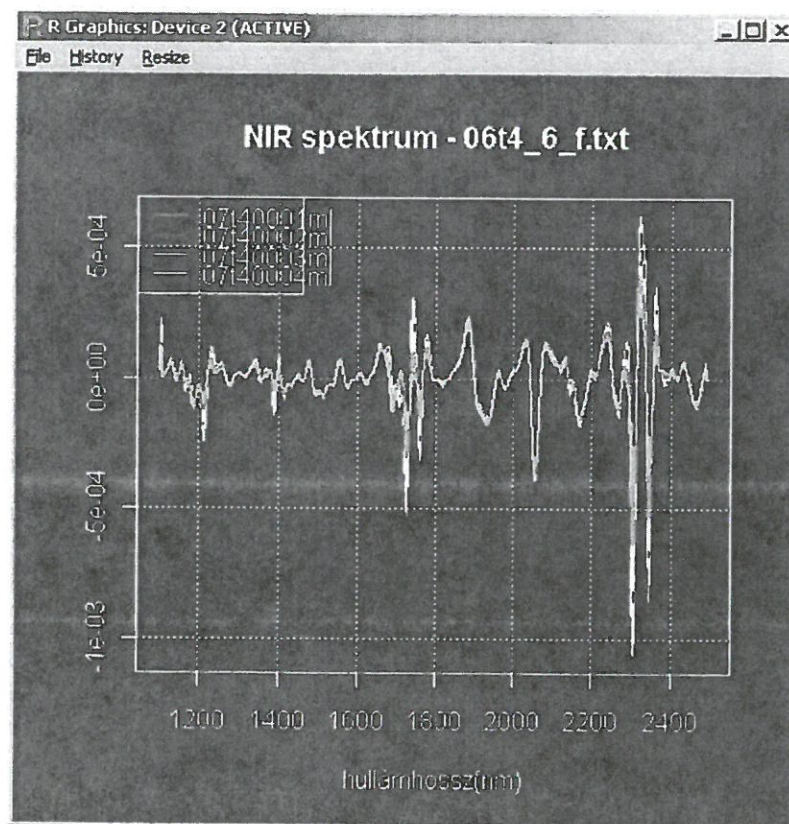


Figure 8: Second derivative of the spectra of four selected samples

9. ábra

A PLS utasítás szintaktikája

```
mvr(formula, ncomp, data, subset, na.action,  
     method = pls.options()$mvralg,  
     scale = FALSE, validation = c("none", "CV", "LOO"),  
     model = TRUE, x = FALSE, y = FALSE, ...)  
plsr(..., method = pls.options()$plsralg)  
pcr(..., method = pls.options()$pcralg)
```

Figure 9: Syntax of the PLS command in R

A 10. ábrán látható grafikus felület teszi lehetővé, hogy a megfelelő szerkezetű „pls” utasítás bemenő paraméterei a gyakorlatlan felhasználó számára is érthető, kezelhető módon kerüljenek meghatározásra. A parancs végső formájának előállítására már programutasítások feladata.

A NIR spektrumok feldolgozását elősegítő grafikus felhasználói felület használhatóságát jelen dolgozatban az Anyag és módszer fejezetben ismertetett demonstrációs adatbázison mutattuk be. A 11. ábrán az alkalmazott pls regresszió futási eredményeit láthatjuk. A grafikus megjelenítés nem helyettesíti az eredmények numerikus megjelenítését, csupán kiegészíti azt. A numerikus megjelenítésben az „R commander” meglévő lehetőségeit használtuk fel.

Bár nem célja a dolgozatnak, a teljesség kedvéért értelmezzük a 11. ábra legfontosabb eredményeit. Hat komponens alkalmazása esetén a kalibrációs modell a szárazanyag (dml), zsírtartalom (ee) nyersfehérje (prot) variációját rendre 93,09, 98,99, 99,07 százalékban becsli. Az eredmények magyarázatában a további részletek iránt érdeklődők figyelmébe a szakirodalmat ajánljuk: *Siesler és mtsai (2002), Mevik és Wehrens (2007)*.

10. ábra

A PLS regresszió paramétereinek beállítása.
A grafikus felület a keresztvalidáció előírására is alkalmas.

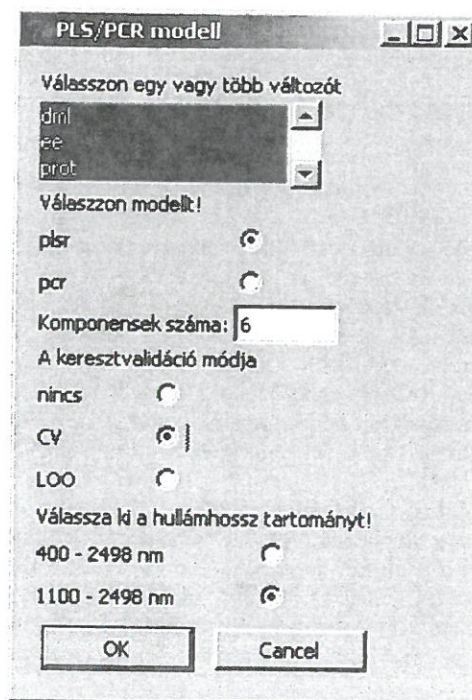


Figure 10: GUI to enter the parameters of the PLS regression including the cross validation

11. ábra

A PLS regresszió numerikus végeredménye, az R Commander eredményablakában.

VALIDATION: RMSEP							
Cross-validated using 10 random segments.							
Response: dml							
	(Intercept)	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps
CV	0.6393	0.4613	0.2753	0.2818	0.2482	0.2162	0.2271
adjCV	0.6393	0.4597	0.2736	0.2798	0.2437	0.2138	0.2242
Response: ee							
	(Intercept)	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps
CV	6.826	3.894	2.989	1.187	1.180	1.159	1.226
adjCV	6.826	3.882	2.971	1.178	1.173	1.145	1.172
Response: prot							
	(Intercept)	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps
CV	6.856	3.876	2.885	1.203	1.182	1.190	1.260
adjCV	6.856	3.863	2.867	1.191	1.172	1.174	1.203
TRAINING: % variance explained							
	1 comps	2 comps	3 comps	4 comps	5 comps	6 comps	
X	81.85	98.76	99.70	99.88	99.96	99.97	
dml	52.58	84.36	84.43	88.69	92.90	93.09	
ee	70.32	83.98	97.69	98.01	98.28	98.98	
prot	70.97	85.41	97.82	98.14	98.33	99.07	

Figure 11: Computational results of the PLS regression presented in the modified R Commander

KÖVETKEZTETÉSEK

A NIR spektrumok feldolgozására és kalibrációs egyenletek meghatározására alkalmas PLS regresszió használatát jelentősen egyszerűsítő grafikus felhasználói felülettel rendelkező programmodul létrehozására irányuló célkitűzésünk sikeres volt. A továbbfejlesztés alapjául választott R Commander alkalmasnak bizonyult arra, hogy a meglévő menüstruktúráját kibővítsük a szükséges menüelemekkel. Az adatfájl beolvasását végző eredeti R Commander modul sikeresen átalakítható volt. A beolvasott spektrumadatok R Commanderben belüli tárolási módja kompatibilis a program korábbi funkcióival így az általános statisztikai eljárások szintén felhasználhatók a spektrumok elemzésére.

Az elkészült modulok szerkezeti továbbfejlesztéseként a „plug-in” csomagként való átalakítás lehetősége kínálkozik. Az R Commander jelen dolgozatban felhasznált 1.3 verziója már ismeri a „plug-in” fogalmát (Fox, 2007). Ez a technika lehetővé teszi, hogy az R Commander eredeti forrásprogramjának változatlanul hagyjuk. A menüstruktúra megfelelő kiterjesztésével olyan R csomagok futtatása válik lehetővé, melyeket úgy fejlesztettek ki, hogy az R Commanderrel képesek legyenek együttműködni.

Ugyanakkor a spektrumadatok kiugróan magas száma (700 db) rendkívül lelassítja az R Commander futását. Az eredeti forráskód elemzése azonosította az időigényes

kódrészletet, csupán kiiktatásával vált az R Commander használhatóvá NIR spektrumok feldolgozására. Ez a „plug-in” csomagként való átalakítás lehetőségét megkérdőjelezi. A NIR spektrumok feldolgozására létrehozott grafikus felhasználó felület elsősorban a graduális és posztgraduális képzésben alkalmazható eredményesen.

IRODALOM

- Solymosi N. (2005): <...erre, erre...! (Bevezetés az R-nyelv és környezet használatába). <http://cran.r-project.org/doc/contrib/Solymosi-Rjegyzet.pdf>
- Ihaka R., Gentleman R. (1996): R: a language for data analysis and graphics. *J. Comput. Graphic. Stat.*, 5: 299-314.
- Fox J. (2005): The R Commander: A Basic-Statistics Graphical User Interface to R. *Journal of Statistical Software*, September 2005. Volume 14, Issue 9.
- Grosjean P., Lecoutre E., Faria J.C. (2005): SciViews-R. A GUI API and a suite of applications for R. <http://www.sciviews.org/SciViews-R>
- Williams G. (2008): Rattle: A graphical user interface for data mining in R using GTK. <http://rattle.togaware.com>
- Fox J. (2007): Extending the R Commander by „Plug-In” Packages. *R News* Vol 7/3, December 2007. http://www.r-project.org/doc/Rnews/Rnews_2007-3.pdf
- Siesler, H.W., Ozaki, Y., Kawata, S., Heise, H.M. (2002): *Near-Infrared Spectroscopy* Wiley-VCH GmbH, 2002. 132-136.
- Mevik, B.H., Wehrens, R. (2007): The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software*, January, 2007, Volume 18, Issue 2.

Levelezési cím (*Corresponding author*):

Kövér György
Kaposvári Egyetem, Gazdaságtudományi Kar
Matematika és Fizika Tanszék
7400 Kaposvár Guba Sándor u. 40.,
Kaposvár University, Faculty of Economic Science
Department of Mathematics and Physics
H-7400 Kaposvár, Guba S. u. 40.
Tel.:36-82-505-596
e-mail: kover.gyorgy@ke.hu

