



SZENT ISTVÁN EGYETEM

BÚZANEMESÍTÉSI INFORMÁCIÓS RENDSZER

DOKTORI ÉRTEKEZÉS TÉZISEI

KUTI CSABA

GÖDÖLLŐ

2007

A doktori iskola

Megnevezése: Növénytudományi Doktori Iskola

Tudományága: Növénytermesztési és Kertészeti

Vezetője: **Dr. Virányi Ferenc**
egyetemi tanár, az MTA doktora
SZIE Mezőgazdasági- és
Környezettudományi Kar
Növényvédelemtani Tanszék

Témavezetők: **Dr. Jolánkai Márton**
intézetigazgató, az MTA doktora
SZIE Növénytermesztési Intézet

Dr. Láng László
tudományos osztályvezető, az MTA
doktora
MTA Mezőgazdasági Kutatóintézete

.....
Dr. Virányi Ferenc
Az iskolavezetőjének
jóváhagyása

.....
Dr. Jolánkai Márton
A témavezető jóváhagyása

.....
Dr. Láng László
A témavezető jóváhagyása

Bevezetés

Az elmúlt két évtizedben a számítógépek használata a mezőgazdasági kutatással foglalkozók körében gyorsan terjedt és ezzel egyidőben kezdtek kialakulni az adatkezelési szokások is. Az általunk vizsgált nemzetközi gyakorlat és a hazai tapasztalatok is azt mutatják, hogy speciális szoftverek hiányában, a mezőgazdasági kutatásban a kis és nagyobb szervezetek egyaránt kénytelenek voltak „egyéni” megoldásokat kifejleszteni az évek során. Az adatkezelés és kiértékelés a legtöbb esetben a számítógépeken szövegszerkesztők, táblázatkezelők, adatbázis-kezelők, statisztikai programok segítségével történt.

A kutatási programok méretének növekedése, amely általában együtt járt a humán infrastruktúra optimalizálásával, megmutatta, hogy ezek a megoldások nem képesek a kutatási tevékenységek szervezésére, nagy adatszám esetén az adatok feldolgozása nehéz, a nem struktúrált adatok közötti összefüggések vizsgálata nem lehetséges, konkrét helyzetekben a döntés az azt alátámasztó információk nélkül bizonytalan, az elemzések, jelentések elkészítése mindennapos nehézségekbe ütközik.

Szükségessé vált olyan technológia kifejlesztése, amely egységbe integrálja a komplex forrásokból származó különféle adatokat, leír és eloszt tevékenységeket, meghatározza ezek időbeni sorrendjét és végül az eredmények feldolgozásával

adatokat szolgáltat a kutatási program teljesítményének jellemzésére.

Általánosságban elmondhatjuk, hogy még napjainkban is igaz az, hogy amíg az alapszoftvereket a szoftverpiacról lehet és célszerű beszerezni, addig a speciális mezőgazdasági kutatások és alkalmazási területek igényeit kielégítő felhasználói szoftverek, és információs rendszerek jelentős részét még mindig egyedileg kell kifejleszteni.

E dolgozat a martonvásári búzanemesítéssel foglalkozó mezőgazdasági kutatás hatékonyságának javításra létrehozott információs rendszert mutatja be. Célja a kutatói közösség nemesítési és szántóföldi kísérletezésekkel kapcsolatos közös ismeretrendszerének egységbe foglalása és az egyes tevékenységek információtechnológiai támogatása.

Célkitűzések

A kutatási tevékenységgel összefüggő adatok, információk egységbe foglalása és az egyes tevékenységek információtechnológiai támogatása olyan átfogó célkitűzések megfogalmazására készített, amelyek érintik a búzanemesítési és szántóföldi kísérletezési tevékenységek minden területét. Az információs rendszer létrehozásának fő célkitűzései:

1. egységes, integrált adatstruktúra megtervezése, és fizikai megvalósítása,
2. input adatok adatbázisokba történő egységes elhelyezése, automatizált adatgyűjtő és kézi beviteli eszközök létrehozása,
3. adatbázisból kiválogatott adatokból, rögzített és változó szerkezetű listák, különféle outputok előállítás,
4. összetett szántóföldi kísérletezési és nemesítési műveletek szervezése,
5. általános adatlekérdező funkciók beépítése,
6. teljes körű pedigre- és génbanki nyilvántartások létrehozása és felügyelete,
7. alapanyagcsere szervezése és a hozzátartozó címlista adatnyilvántartása,
8. alapvető elemzések készítésére alkalmas, az adatbázisban lévő adatokon közvetlenül futtatható, statisztikai modul létrehozása,
9. egyidejű hozzáférések és csoportmunka lehetővé tétele.

Anyag és módszer

Szoftver

A Breeder alkalmazást a Microsoft® VB-IDE (Visual Basic Integrated Development Environment) általános célú programfejlesztői környezetének segítségével fejlesztettük. Mintegy 86 MB-nyi területet foglal el a fejlesztő gépén, a felhasználók gépeire ezt nem kell telepíteni.

A Breeder által létrehozott (output) és a Breeder-be bemenő (input) adatok többsége olyan formátumokban van, amely a Windows® alatt futó Microsoft Office2000® és OfficeXP® csomagok részét képező Microsoft® Excel, Microsoft® Access és Microsoft® Word alkalmazásokban is megjeleníthető, így az optimális működéshez célszerű az Office irodai csomagot telepíteni a munkaállomásokra (kb. 150 MB).

Hardver

Munkaállomások

Azok a számítógépek, amelyeket a nemesítési információs rendszer fejlesztői és felhasználói (kutatók, intézeti mérnökök, technikusok) munkájuk során használnak. Tipikusnak mondott konfiguráció: Intel Pentium IV 1,5 GHz-től (vagy kompatibilis), memória: 256 MB RAM, merevlemez: 160 GB, monitor: SVGA

17", hálózati kártya (10/100 Ethernet LAN), optikai egység (CD/DVD), USB csatlakozó, egér, billentyűzet.

Automata adatgyűjtők

Azokat a számítógépeket, amelyek még működőképeseek, de teljesítményük a mindennapi munkavégzéshez már nem elegendő. Tipikusnak mondott konfiguráció: Intel Pentium III 800 MHz-től (vagy kompatibilis), memória: 128 MB RAM, merevlemez: 40 GB, monitor: SVGA 15" (800x600), hálózati kártya (10/100 Ethernet LAN), optikai egység (CD/DVD), soros (RS232) csatlakozó, egér, billentyűzet.

Hálózati eszközök

A munkaállomásokról érkező kérések kiszolgálásának biztosítására egy IBM eSeries 200 típusú központi adatbázisszerver üzemel az alábbi kiépítésben: 1,6 GHz-es processzor, 160 Gbyte-os merevlemez, 1 Gbyte RAM-al felszerelve.

Adatfájlok és tárigény

Az 1984-től kezdődően évente létrehozott adatbázisok együttes mérete kb. 200 MB. Ezek a fájlok nem egyforma méretűek, hanem arányban állnak a nemesítési anyag növekedésével és a gyűjtött

adatok típusának és számának évről-évre történő növekedésével. A felsoroltakon kívül még tárolunk egyszerűbb szöveges és Excel típusú alapfájlokat, paraméterfájlokat, ezekben kísérleti elrendezések, programbeállítások, eltárolt címkeformátumok vannak.

Vonalkódalapú eszközök

Címketípusok

A vonalkódok nyomtatása a kereskedelmi forgalomban kapható szabványos méretű öntapadós és műanyag címkékre történik. Az öntapadós címkéket hagyományos lézer (HP, CANON) nyomtatókkal nyomtatjuk Avery-Zweckform típusú szabványos (AZ3651, AZ3652) címkeívekre. Az általunk használt thermo-érzékeny műanyagcímkek speciális anyagból készülnek (HDPE), 0,25mm vastagságúak, kereszt alakú résezettel, kétféle méretben: 11cm x 2,5cm, ill. 11cm x 5cm.

Vonalkódnymtatók

A műanyag címkék nyomtatása speciális eszközök igénybevételével történik. Az általunk használt thermo-transzfer nyomtatók: TTX300 COBRA, TTX350 OCELOT (Avery Dennison Corporation, 1985) max. 110mm széles címkék

előállítására alkalmasak. Műszaki szempontból jelentősen eltérnek a hagyományos nyomtatóktól.

Vonalkódleolvasók

A mérési adatok automatikus azonosítása CCD (Charge Coupled Devices) típusú optikai leolvasókkal történik: BCH5X49, BS-L01

Adatgyűjtési eszközök

- NIR/NIT készülékek: FOSS Infratec® 1241, Perten Inframatic® 8611,
- digitális mérlegcsalád (Mettler): PB602-S, Viper SW, Spider SW,
- szemkeménységmérő készülék: Perten SKCS 4100,
- dagasztott tészta tágulásának vizsgálata (Chopin): Alveograph, Alveolink,
- reológiai tulajdonságok mérése (Brabender): Farinograph, Farinograph-E,
- textúra elemző rendszer (Stable Micro Systems): TA.XTPlus,
- kombájnra szerelhető automata adatgyűjtők: OMNIDATA Polycorder, HARVESTMASTER HM-Fieldbook,

Alkalmazott statisztikai módszerek és kísérleti elrendezések

Figyelembe véve, hogy az egy helyen, egy táblában beállított többismétléses kísérlet adja a fajta teljesítményének az adott táblára vonatkozó legmegbízhatóbb becslését (Matuz, 1987), az egytényezős többismétléses és többtényezős többismétléses kísérletek statisztikai értékelését beépítettük a rendszerünkbe. A véletlen blokkelrendezésű egytényezős és különböző elrendezésű kéttényezős (véletlen blokk, osztott parcellás) kísérletek statisztikai kiértékelését a Sváb által leírt módszer (Sváb, 1981) alapján végezzük és értelmezzük.

A kéttényezős varianciaanalízis esetében a két minőségi tényező kölcsönhatásának megléte vagy hiányának kimutatása fontos lehet számunkra. A kéttényezős variancia tábla értékelését szignifikáns kölcsönhatás esetén a Tóthné által leírt módszer (Tóthné, 2004) alapján végezzük és értelmezzük.

Az öntermékenyülő gabonafajok nemesítésére leggyakrabban alkalmazott nemesítési eljárások (pedigré- és tömegszelekciós módszer; Bedő, 2004) közül a pedigrémódszer sajátosságaihoz optimalizáltuk az adatszerkezetet.

Eredmények és következtetések

A nemesítési és szántóföldi kutatási program méretének és hatékonyságának növeléséhez információs infrastrukturális háttérrel hoztuk létre. Egy áttekinthető rendszert építettünk fel, amely az általánosságra való törekvés mellett, tükrözi a Martonvásáron sok éves tradíció során kialakult kutatási szemléletet.

1. Relációs adatmodell megtervezése és létrehozása

A martonvásári búzanemesítési adatmodellt a célkitűzésben megfogalmazott információs rendszer részeként hoztuk létre. Célja a nemesítéssel kapcsolatos összes adat egységes, integrált nyilvántartása; döntéstámogató információk, jelentések csoportosításának és teljes körű megjelenítésének lehetővé tétele; a nemesítési adatok kapcsolatának biztosítása a genealógiai és génbanki adatokkal, az alapanyagcsere és hozzátartozó címlista adatnyilvántartásokkal.

Az általunk tervezett rugalmas adatmodellre támaszkodva megvalósítottuk a teljes körű adatintegrációt, amely biztosítja hogy a nemesítés és szántóföldi kísérletezés folyamán keletkező összes releváns adatnak helye legyen az adatbázisokban, és rendelkezésre álljon valamilyen eszköz a kezeléséhez.

Adatainkat témakörök szerint csoportosítottuk, így jöttek létre az adatmodellen belül az egyes adatbázisok: az évről-évre megújuló nemesítési adatbázisok, a genealógiai adatbázis, az alapanyagcsere adatbázis, génbanki adatbázis. Az egyes adatbázisokban helyet foglaló táblák számát az adott témakör által determinált témák száma határozza meg, így a nemesítési adatbázisban (12 tábla) pl. külön táblákban helyezkednek el a kísérletek, a szántóföldi megfigyelési adatok, a minőségvizsgálatok eredményei, súlynyilvántartás, stb.

2. Tevékenységszervező programmodulok

Ahhoz, hogy az adatok az információs rendszeren belül felhasználhatóvá váljanak, szükség van a kutatási tevékenységek végzését segítő speciálisan tervezett alkalmazásokra. Ezek az alkalmazások az egyes tevékenységek technológiai leírásán alapulnak. A közel ötven alkalmazást a „Breeder” nevű felhasználói felület foglalja keretbe és nyújt egységes hozzáférést a kezelésükhöz az őket használó nemesítőknek, szakirányú mérnököknek és más mezőgazdasági szakembereknek a kutatási munkájuk elvégzése során. A különböző modulokat összefogó egységes keret tartalmaz menüket, almenüket és egy gyorsindító eszköztárat a leggyakrabban használat programmodulok

meghívására. A leggyakrabban használt modulok között első helyen kell említenünk a kísérletek tervezéséhez segítséget nyújtó alkalmazásokat: a keresztező, szelektáló és parcellázó modulokat.

Keresztezéskor a program automatikusan hozza létre az új pedigréket (Purdy et al., 1968), a szelekció történetét rögzítő „Kód” információt és keresztezési listákba menti az egyes keresztezési programokkal kapcsolatos fontos adatokat. Ezekből később elő lehet keresni az anyai és a beporzó összesítőket, tudni lehet hogy a szülőket mely kísérletekből kell elővenni és azt is, hogy az illető szülő összesen hányszor szerepel a keresztezési programban.

A szelekciós modul rögzíti a genotípus származását (előző évi: kísérlet, parcella), az elvetett sorok számát, automatikusan ír: kódot (szelekció történet), új törzsazonosítószámot (ha szükséges) és kapcsolatot tart a fajtafenntartással, fagy- és rezisztencia kísérletekkel. Automatikusan nyilvántart súlyt és lehetőség van az egyenkénti kijelölés és átszelektálás helyett akár több száz genotípust is automatikusan átszelektálni, előre beállított bonitálási értékek és egyéb feltételek megadásával.

A parcellázó modul nemcsak a több szempont alapján történő végső parcellasorrend kialakításának, hanem bizonyos kísérletek kontrollokkal való kiegészítésének, egyéb esetekben előre megadott méretű kísérletek automatikus előállításának hatékony

eszköze is. Ez attól függ, hogy a rendelkezésre álló (négyféle) parcellázási módok közül a felhasználó melyiket választja. A végleges sorrend kialakításakor a nemesítőnek lehetősége van elkülöníteni a különböző kalásztípushoz tartozó törzseket, sorba rakni őket az előző évi kalászosolási idő szerint, azonos kalászosolási idejű genotípusokon belül a testvér törzseket csoportosítani, stb.

Szintén a gyorsindító eszköztárról kezdeményezhető a génbanki nyilvántartást végző modul, amelynek segítségével megvalósítjuk a génbanki és a nemesítési adatbázisok között zajló adatcserét (tárolás, kitermesztés). Ugyaninnen indítjuk a nemesítési és génbanki alapanyagok bel- és külföldi elosztását és az onnan érkező tételek felügyeletét végző modult. Ennek segítségével tartjuk nyilván úgy a beérkező, mint az intézetből kiküldött magtételhez kapcsolódó összes információt.

A gyakran futtatott modulok közé tartozik a felvételezési adatok kézi bevitelét lehetővé tevő modul, a külső (Excel) adatfájlok tartalmát az információs rendszer adatstruktúrájába illesztő modul, a genealógiai adatok konzisztenciáját javító modul, a vetésterveket és felvételezési füzeteket előállító modul, a vonalkódok tervezését és nyomtatását végző modul, a legkülönbözőbb laboratóriumi minőségvizsgáló műszerekről és automata adatgyűjtővel ellátott kombájnokról származó adatokat automatikusan gyűjtő modul.

3. Pedigré és génbanki nyilvántartás

A nemesítési és szántóföldi adatok nyilvántartásának sarokpontja az egységes pedigré model. Az egységes pedigré model központi gondolata a genotípusok egyedi azonosítók hozzárendelésével történő azonosítása, a homonímák és szinonímák kezelése, amelyek természetesen módon adódnak a genotípusok fejlődése és használata során.

A jelenleg több mint 110.000 különböző genotípus elnevezést tartalmazó nyilvántartásunk számára létrehoztunk egy különálló adatstruktúrát. Minden pedigré bejegyzéshez két azonosítót rendelnek automatikusan a pedigré adatokat kezelő programmodulok. Az első (PAz) az eltérő kombinációkat azonosítja, a második (SID) az azonos keresztezési kombinációból származó testvértörzseket különbözteti meg egymástól.

Minden genotípus elnevezés csak egyszer, egy központi nyilvántartásban szerepel, és a genotípusra vonatkozó genealógiai információk innen csatolhatók a különböző évek, vagy kísérletek nyilvántartásához. Azoknál a genotípusoknál, amelyeknél több elnevezés is van forgalomban (pedigré, törzsnév, fajtanév, stb.) az adatbázist kezelő programmodulok eldöntik, hogy ezek közül melyik a legaktuálisabb és azt mutatják a felhasználónak.

4. Adatgyűjtés

Amikor adatgyűjtésről beszélünk elsősorban az automatizált, vagy félig automatizált funkciók megvalósítására gondolunk. Az adatgyűjtés automatizálásához szükség van olyan korszerű technikák alkalmazására, amelyek a mikroprocesszorral felszerelt és számítógéphez csatlakoztatható digitális vezérlésű mérési- és analitikai eszközök, széleskörű elterjedésével elérhetővé váltak. Az adatgyűjtési funkciók automatizálásával az adatgyűjtés és az adatfeldolgozás üteme (és minősége) azonos szintre emelhető, ez nagyon fontos olyan esetekben, ahol a megtakarított időt tekintjük meghatározó tényezőnek (pl. aratás és vetés közötti rövid időköz optimális kihasználásakor).

Az automatizált adatgyűjtéshez alapvetően két dolog szükséges: kommunikáció a megfelelő műszerrel és az automata adatazonosítás.

A kommunikációhoz műszerspecifikus szoftveres interfész szükséges, amelyet minden műszerhez, vagy műszer családhoz külön létre kell hozni. A befektetett munka megtérül, mert az adatok hibátlan, gyors gyűjtése a műszer hátralévő élettartamára megoldott.

Rendszerünkben az automata adatazonosítás alapja a vonalkód, ezért beépített vonalkódtervező és nyomtató modult

hoztunk létre. Ez több adatmezőt is össze tud vonni egyetlen vonalkódba, és ami nagyon fontos, az adatok közvetlenül az adatbázisból nyomtathatók öntapadós vagy műanyag alapú címkékre.

Az általunk alkalmazott vonalkódokkal ellátott címkéknek kettős szerepe van: egyrészt a humán olvasásra alkalmas felirat rendelkezésre áll felvételezéskor/vetőmag előkészítéskor, másrészt az alatta elhelyezkedő vonalkódokat a mérési eredmények adatbázisokba történő elhelyezéséhez használjuk azonosításhoz.

5. Génbank

A martonvásári génbank alapgyűjteményét mintegy 11-12.000 tétel képezi. Fontos célkitűzése volt a tervezett rendszernek, a génbanki adatok elkülönítése a nemesítési adatoktól. Az átjárás a nemesítési adatbázisok és a génbanki adatbázis között kétirányú: a nemesítési programból tételek helyezhetők el megőrzésre a génbankba és fordítva génbanki tételeket kitermeszthetünk a nemesítésbe való bevonással.

A génbanki modul a kiválasztott kísérlet genotípusait mindig szelektíven ajánlja fel az éppen aktuális műveletek végzéséhez. Ha egy adott kísérletből (kísérletcsoportból) új tételeket szeretnénk

felvinni a génbankba, akkor csak azokat a genotípusokat ajánlja fel, amelyek még nem szerepelnek a génbankban. Ellenkező esetben, amikor arra vagyunk kíváncsiak, hogy egy adott kísérlet genotípusai közül melyek azok, amelyekkel frissíteni tudjuk a génbank egy adott célkísérletét, akkor a kísérletből csak azokat a genotípusokat látjuk, amelyek már jelen vannak ott. Az ellentétes irányú tevékenység - génbankból a kitermesztésbe - adatkezelése nagymértékben hasonlít a fentihez. Fontos különbség, hogy hangsúlyosabbak a célirányos génbanki menedzselés eszközei: kérhetjük azoknak a génbanki tételeknek a listázását, amelyek régebbiek egy adott évnél, súlyuk, mennyiségük (zacskószám) egy adott kritikus érték alatt van. Ha egy konkrét kísérletet kiválasztottunk, akkor a modul mutatja, hogy abból milyen kritikus génbanki tételek frissíthetők.

6. Statisztikai modul

Bizonyos kutatási programok számára csak egy szűken kimért időkeret áll rendelkezésre a kutatások szakmai eredményeinek kiértékelésére, különösen igaz ez a nemesítési programok esetében, amelyeknek az aratási szezon végén általában napok állnak rendelkezésre a jelentések elkészítéséhez és hetek alatt kell előkészíteni a következő ciklus tenyészkertekben elvetésre kerülő

anyagait. Az egy vagy több termőhelyen elvetett azonos vagy különböző randomizációjú kísérletek alapvető statisztikai elemzéseinek elvégzéséhez nincsen szükség külön statisztikai csomagok használatára, sem az adatok előkészítésére ezekhez az elemzésekhez. A rendszerben megtalálható alapvető statisztikák:

- Leíró statisztika
 - átlag mutatószámai: számtani átlag, középső érték (Median), mértani (Geom), harmonikus (Harm) középárányos
 - változó adatainak: legnagyobb (Max), legkisebb (Min) értéke, összege (Sum)
 - szóródás mutatói: szórásnégyzet, szórás, átlagos hibaszórás, variációs koeff., átlag konfidencia intervalluma
- Véletlenblokk varianciaanalízis
 - egytényezős
 - kéttényezős
- Osztott parcellás (split-plot) elrendezés
- Lineáris regresszióanalízis
 - egyszeres
 - többszörös
- Korrelációs mátrix számítása

**Főbb tudományos eredményeimnek a következőket
tekintem:**

1. A nemesítési és szántóföldi kutatási program méretének és hatékonyságának növeléséhez információs infrastrukturális háttérrel hoztam létre.
2. Az általam tervezett rugalmas adatmodellre támaszkodva megvalósítottam a teljes körű adatintegrációt, amely biztosítja a rendszer hatékonyságát a nagyméretű nemesítési és szántóföldi kutatási programok esetében. A nemesítés és szántóföldi kísérletezés folyamán keletkező összes releváns adatnak helye van az adatbázisokban, és rendelkezésre áll valamilyen eszköz a kezelésükhöz.
3. Korszerű pedigre nyilvántartást fejlesztettem ki, amellyel megkülönböztethetők az eltérő kombinációk és együtt láthatók az azonos származású törzsek. Minden genotípus elnevezés csak egyszer, egy központi nyilvántartásban szerepel, a genotípusra vonatkozó genealógiai információk innen csatolhatók a különböző évek, vagy kísérletek nyilvántartásához.

4. Létrehoztam a főbb nemesítési tevékenységek technológiai leírásán alapuló tevékenységszervező programmodulokat. Segítségükkel lehetővé vált számos munkafolyamat automatizálása és a szelekciós döntések során a genotípusokra vonatkozó, korábban rögzített valamennyi információ elérhetősége. Átaluk valósul meg a hálózatos adatkezelés és a csoportmunka.

5. Megteremtettem az automatizált adatgyűjtés feltételeit. Beépített vonalkódgeneráló segítségével hozom létre az automata adatazonosításhoz szükséges jelöléseket és létrehoztam a műszerekkel kommunikáló adatgyűjtő programokat is.

Következtetések és javaslatok

A búzanemesítéssel foglalkozó kutatás versenyképességének fontos feltétele a kutatási programok méretének növelhetősége, az egyre nagyobb számban keletkező adatok hatékony kezelése, de mindenekelőtt a hasznos, döntéstámogatáshoz nélkülözhetetlen információk szintetizálása a több milliányi adatból, amelyet egy nagyobb méretű kutatási program generál évről-évre.

Ezen szempontok szerint tervezett információs rendszer megvalósítását ismerteti a dolgozat. Az eredmények bemutatása mellett igyekeztem felmérni a növénytermesztéssel és növénynemesítéssel kapcsolatos témakör szoftver ellátottságát is.

Információs rendszerünk létrehozásakor abból indultunk ki, hogy egy nagyméretű kutatási programban csapatmunka folyik. Egyidejűleg többen dolgoznak a feladatok megvalósításán, azt többször is megszakítva más időszerű részfeladatok kényszerű beütemezése miatt. A rendszertervező számára ez azt jelenti, hogy minél nagyobb szabadságot kell biztosítani a kutatóknak és a technikai személyzetnek ahhoz, hogy közös tevékenységeiket különböző időpontokban tudják végezni. A rendszer feladata fenntartani a mindenkori változtatások és a megkezdett műveletek folytatásának lehetőségét úgy, hogy előreláthatólag ne kelljen semmit előlről kezdeni. Ebből adódik az első lényeges különbség,

ami a rendszerünk és a vizsgált szoftveres megoldások között van. Utóbbiaknál már a tervezés legelején be kell állítani a kísérletek több olyan jellemzőjét, amelyeknek a rögzítése egy nagyméretű kutatási programban a genotípusok összeválogatásának megkezdésekor még nem áll rendelkezésre vagy jelentős további erőfeszítéseket igényelne.

Az adatszerkezet rugalmassága és az egyes tevékenységek megfelelő szervezése biztosítja számunkra azt, hogy egymásra épülő kutatási tevékenységek elkezdhetők legyenek akkor is, ha az alapul szolgáló műveletek még nincsenek végleges formában kidolgozva. Így pl. nem kell megvárni, amíg egy korai generációs kísérlet több ezer genotípusa véglegesen összeáll, hanem a válogatással párhuzamosan el lehet indítani a szelektált genotípusok számára a vonalkódok előállítását, majd ennek alapján el lehet kezdeni a vetőmagok előkészítését is.

A vizsgált szoftveres megoldásokban a genealógiai adatok kezelése nagyon eltérő, arányban áll a rendelkezésre álló genealógiai adatok mennyiségével és sokszínűségével. Azt mondhatjuk, hogy amint a genealógia adatok mennyisége és változatossága (több faj) nő egy rendszerben, úgy válik egyre bonyolultabbá és kap mind nagyobb és nagyobb jelentőséget a kezelésük az egyes szoftveres megoldásokban. A genealógiai adatok felépítése és kezelése úgy a rendelkezésre álló adatok

számát, mint a kezelési módszerek bonyolultságát tekintve, fontos szerepet kap a *Breeder*-ben. Eltérően az összes többi szoftveres megoldástól, a nagyméretű nemesítési és szántóföldi programok megvalósításához szükséges alapanyag nyilvántartáson túl szoros kapcsolatban áll a vele párhuzamosan működő génbankkal.

A *Breeder* statisztikai moduljának célja lehetőséget biztosítani a nemesítési és szántóföldi kísérletek gyors kiértékeléséhez, oly módon, hogy az adatbázisban tárolt mérési és megfigyelési adatokon közvetlenül és könnyen futtathatók legyenek a statisztikák. Az egy vagy több termőhelyen elvetett azonos vagy különböző randomizációjú kísérletek alapvető statisztikai elemzéseinek elvégzéséhez nincsen szükség külön statisztikai csomagok használatára, sem adatok előkészítésére az elemzésekhez.

Egyedi megoldások a rendszerben: a vonalkód technika beillesztése és az erre alapozott széleskörűen alkalmazott automata adatgyűjtés, a rendszer informatív jellegét biztosító többtáblás, többéves adatlekérdező funkciók, rögzített és változó formátumú és tartalmú vetéselrendezések és felvételező füzetek nyomtatása.

Eredményeim révén a dolgozatban bemutatott számítógéppel végzett adatkezelés és tevékenységszervezés bevezetésével költséges idő és munkaerő igényes munkafolyamatokat

egyszerűsítettünk le, miközben minimálisra csökkent a pontatlanságból, következetlenségből eredő hiba, ezáltal az alkalmazott informatika jelentős mértékben válhatott a hatékonyság és versenyképesség növelésének fontos eszközévé.

Irodalomjegyzék

- Bedő Z., Marton Cs. (2004): Növénynemesítési módszerek. 71-100 pp. In Bedő Z. (Szerk): *A vetőmag születése. A vetőmagtermesztés elmélete és gyakorlata*. Agroinform Kiadó, Budapest, 537 pp.
- Matuz J. (1987): Az üzemi kísérletezés jelentősége és problémái. In: Barabás Z. (Szerk.): *A búzatermesztés kézikönyve*, Mezőgazdasági Kiadó, Budapest, 223-236.
- Purdy H. L., Loegering W. Q., Konczak C. F., Peterson C. J. Allan R. E., (1968): A proposed standard method for illustrating pedigrees of small grain varieties. *Crop Science*, 8: 405-406.
- Sváb J. (1981): *Biometriai módszerek a kutatásban*, Mezőgazdasági Kiadó, Budapest. 557 pp.
- Tóthné L. K. (2004): Statisztikai hipotézisvizsgálatok. In: Szűcs István (Szerk.) *Alkalmazott statisztika*. Agroinform Kiadó, Budapest, 30-47.

A tézisek alapjául szolgáló, tudományos publikációk

Referált tudományos folyóiratban megjelent dolgozatok

- Láng L., **Kuti Cs.**, Bedő Z. (2001): Computerised data management system for cereal breeding. *Euphytica*, 119.1-2: 235-240.
- Marton L.Cs., **Kuti Cs.** (2002): A kukorica (*Zea Mays*) fiatalkori hidegtűrése öröklődésének értékelése módosított "joint scaling test"-tel. *Növénytermelés*, 51.4: 1-6.
- Marton L.Cs., **Kuti Cs.** (2002): Modified joint scaling test for evaluating the effect of level of heterozygosity of the female parent. *Acta Agronomica Hungarica*, 50.2: 185-190.
- Kuti Cs.**, Láng L., Bedő Z. (2003): Szántóföldi kísérletek tömegmérési adatainak számítógépes rögzítése és nyilvántartása. *Növénytermelés*, 52.3-4: 329-340.
- Kuti Cs.**, Láng L., Bedő Z. (2004): Use of barcodes and digital balances for the identification and measurement of field trial data. *Acta Agronomica Hungarica*, 52.4: 409-419
- Bedő Z., Láng L., Veisz O., Vida Gy., Karsai I., Mészáros K., Rakszegi M., Szűcs P., Puskás K., **Kuti Cs.**, Megyeri M., Bencze Sz., Cséplő M., Láng D., Bányai J. (2005): Breeding of winter wheat (*Triticum aestivum* L.) for different adaptation types in multifunctional agricultural production. *Turk. J. Agric. For.*, 29: 151-156.
- Kuti Cs.**, Láng L., Bedő Z. (2006): Pedigree records in plant breeding: from independent data to interdependent data structures. *Cereal Research Communications*, 34.2-3: 911-918.
- Klara Meszaros, Ildiko Karsai, **Csaba Kuti**, Judit Banyai, Laszlo Lang, Zoltan Bedo (2007): Efficiency of different marker systems for genotype fingerprinting and for genetic diversity studies in barley (*Hordeum vulgare* L.). *South African Journal of Botany* 73: 43-48.

Konferenciakiadványok, konferencia összefoglalók

- Láng L., **Kuti Cs.**, Bedő Z. (2000): Computerised data management system for cereal breeding. 6th International Wheat Conference 5-9 June 2000, Budapest, Hungary, Abstracts, 91
- Láng L., **Kuti Cs.**, Bedő Z. (2001): Computerised data management system for cereal breeding. Wheat in a Global Environment Proc. of the 6th International Wheat Conference, 5-9 June 2000, Budapest, Hungary, Kluwer Academic Publishers 561-569.
- Láng L., **Kuti Cs.**, Bedő Z. (2002): Martonvásár software in cereal breeding. Chinese-Hungarian Workshop on "Molecular Genetics and Breeding in wheat". Martonvásár, Szeged 21-26 May, 2002 Lecture Abstracts, 11-12.
- Láng L., **Kuti Cs.**, Bedő Z. (2003): Növénynevelési adatok számítógépes nyilvántartása és elemzése a martonvásári fejlesztésű "BREEDER" programcsomaggal. IX. Növénynevelési Tudományos Napok, Összefoglalók, 2003. Márc. 5-6. MTA Budapest, 18.
- Kuti Cs.**, Láng L., Bedő Z. (2004): Azonosítás és mérés vonalkód és digitális mérleg alkalmazásával. X. Növénynevelési Tudományos Napok, Poszter Összefoglalók, 2004. Február 18-19, MTA Budapest, 125.
- Kuti Cs.**, Láng L., Bedő Z. (2005): Pedigré nyilvántartás a növénynevelésben: rendezetlen adatoktól az EGYSEGES adatstruktúrákig XI. Növénynevelési Tudományos Napok Összefoglalók, 2005. márc. 3-4. MTA, Budapest, 47.
- Bányai J., Szűcs P., Karsai I., Mészáros K., **Kuti Cs.**, Láng L., Bedő Z. (2005): DNS szintű fajtaazonosítás búzában. XI. Növénynevelési Tudományos Napok Összefoglalók, 2005. márc. 3-4. MTA, Budapest, 60.

Egyéb publikációk

Bedő Z., Szunics L., Láng L., Veisz O., Karsai I., Juhász A., Rakszegi M., Vida Gy., Szűcs P., **Kuti Cs.**, Megyeri M., Gál M. (2002): Breeding. Annual Wheat Newsletter, Vol. 48:66-70.

Bedő Z., Láng L., Szunics L., Veisz O., Vida Gy., Karsai I., Mészáros K., Juhász A., Rakszegi M., Szűcs P., Puskás K., **Kuti Cs.**, Megyeri M., Gál M., Nagy IJ. (2003): Items from Hungary, Department of Wheat Breeding, Agricultural Research Institute, Martonvásár. Annual Wheat Newsletter, 49:30-34.