

*Maatalouden
tutkimuskeskuksen
julkaisuja*

S A R J A B

2

*Jukka Öfversten
Hannele Nikander*

**Lajikekoesarjojen
analysointi**

*Jukka Öfversten
Hannele Nikander*

Maatalouden tutkimuskeskus, tietopalveluyksikkö, 31600 Jokioinen, puh. (916) 41 881

Lajikekoesarjojen analysointi

Analysis of series of crop variety trials

ISSN 1238-9943

Copyright

Maatalouden tutkimuskeskus (MTT) 1996

Julkaisija

Maatalouden tutkimuskeskus (MTT), 31600 Jokioinen

Jakelu ja myynti

MTT, tietopalveluyksikkö, 31600 Jokioinen

Puh. (916) 41 881, telekopio (916) 418 8339

Sisäsivujen painopaperina käytetylle G-Printille on myönnetty pohjoismainen joutsenmerkki.
Kansimateriaalina käytetty Countryside on 75-prosenttisesti uusiokuitua.

Tiivistelmä

Avainsanat: peltokasvitutkimus, lajikevertailu, mallintaminen, sekamalli, sopeutuminen

Kasvinjalostajat ja siemenkauppiaat testauttavat uusien peltokasvilajikkeiden viljelyarvon puolueettomien tutkimuslaitosten järjestämissä kenttäkokeissa ennen lajikkeiden kaupallista viljelyä. Suomessa kokeita koordinoi maa- ja metsätalousministeriön toimeksiannosta Maatalouden tutkimuskeskus.

Lajikekokeista koottuja tiivistettyjä lajikekoeaineistoja on Maatalouden tutkimuskeskuksen käytössä noin kolmen vuosikymmenen ajalta. Tiivistettyjen koeaineistojen tilastollinen analyysi on helppoa, jos jokainen vertailussa tarvittava lajike on ollut mukana kaikissa vertailuun käytettävissä kokeissa. Paljon lajikkeita sisältävät monivuotiset koesarjat eivät yleensä ole tällaisia. Suurista aineistoista voidaan kuitenkin aina valita osa-aineistoja, joissa tietyn kasvin kaksi eri lajiketta ovat olleet yhtä aikaa samoissa kokeissa. Tällaista aineistoa voidaan käyttää kätevästi kyseisten kahden lajikkeen vertaamiseen. Osa-aineistojen valintaa on viime aikoihin asti käytetty lähtökohtana lajikekoesarjojen analysoimiseen. Tämä parivertailuksi kutsuttu menettely on kuitenkin monessa mie-

lessä kyseenalainen, sillä osa-aineistojen käyttö voi aiheuttaa ristiriitaisia johtopäätöksiä ja heikentää tilastollisten testien voimakkuutta. Sen avulla ei myöskään saada juuri mitään viitteitä lajikkeiden herkkyydestä erilaisille ulkoisille olosuhteille.

Tämän raportin tarkoituksena on kuvata lineaaristen sekamallien käyttöä lajikekoesarjojen analysoinnin lähtökohtana. Raportissa osoitetaan, että sekamalleilla saadaan päätöksenteon kannalta käyttökelpoisempia ja joustavampia tuloksia kuin perinteistä parivertailua käyttämällä. Johtopäätökset voidaan tehdä kaiken käytettävissä olevan aineiston perusteella. Keskiarvoparametrien lisäksi voidaan mallintaa myös vastemuuttujien hajontaa. Hajonnan mallintamisen avulla tehtävät johtopäätökset voidaan ulottaa koskemaan joko koko viljelyaluetta tai vain yhtä tai muutamaa koepaikkaa. Raportin esimerkeissä käytetään virallisista syysvehnän ja ohran lajikekokeista koottuja aineistoja. Kaikissa malliratkaisuissa käytetään vastemuuttujana koetuloksina saatuja jyväsatoja.

Summary

Key words: crop research, variety tests, modelling, mixed model, adaptation

Before introducing new agricultural crop varieties into commercial cultivation, plant breeders and seed merchants enter them into field trials organized by independent authorities. In Finland, the Agricultural Research Centre, authorized by the Ministry of Agriculture and Forestry, coordinates such trials.

The Agricultural Research Centre has gathered aggregated variety trial data from series of field trials for the past three decades. Analysing this kind of data is straightforward when every variety is grown in every trial. Series of trials lasting several years and including several varieties are not usually like this. From a large set of data we can, however, always select a subset, where two different varieties of a particular crop have been present in all the experiments. The resulting data can easily be used to compare the two varieties. Until recently, selecting this kind of a subset has formed the basis for the statistical analyses of the statutory Finnish variety trials. This method, strictly directed to pairwise comparisons, is questionable in many

respects. It may lead to contradictory results, decrease the power of the statistical tests and, moreover, fail to render any information on the sensitivity of the varieties to different environmental effects.

The aim of this report is to describe the statistical analysis of series of variety trials using mixed linear models. We demonstrate that we get more effective and flexible results to base our decisions on by using mixed linear models instead of the traditional pairwise comparisons. Furthermore, our conclusions can be based on all the data available. The use of the mixed linear models not only enables us to model the mean effects of the respective response variable, but also to model the variance of the response variable. This lets us extend our conclusions to cover either the whole growing region or, if needed, only one or few of the experimental sites. In the examples we use the statutory Finnish variety trial data for winter wheat and spring barley. As a response variable we use the recorded measurement of grain yield.

Sisällys

Tiivistelmä.....	3
Summary.....	4
Johdanto.....	7
Lajikekoeaineistot.....	7
Lajikkeiden parittainen vertailu.....	8
Lajikkeiden samanaikainen vertailu.....	9
Mallintamisen mahdollisuudet.....	9
Esimerkkiaineisto 1: syysvehnä.....	11
Mallin tarkentaminen.....	12
Poikkeuksellisten kokeiden tunnistaminen.....	12
Estimointi, ennustaminen ja testaaminen.....	14
Herkkien lajikkeiden tunnistaminen.....	15
Esimerkkiaineisto 2: ohra.....	18
Ohralajikkeiden satotulosten ennustaminen.....	19
Ohralajikkeiden vertaaminen.....	21
Mallintamiseen liittyvät oletukset.....	21
Yhteenveto.....	26
Kirjallisuus.....	27
LITTEET	

Figure 1



Johdanto

Kasvinjalostajat ja siemenkauppiat testauttavat uusien kasvilajikkeiden tai vielä nimeä vailla olevien uusien linjojen viljelyarvon puolueettomien tutkimuslaitosten järjestämissä kenttäkokeissa ennen kuin ottavat uudet lajikkeet kaupalliseen viljelyyn. Tarkoituksensa mukaisesti näitä kokeita sanotaan lajikekokeiksi. Niiden tuloksia tarvitsevat sekä kasvinjalostajat, siemenkauppiat, viljelijät että viranomaiset, minkä takia kasvintuotantoon liittyvistä kenttäkokeista huomattavan suuri osa onkin juuri lajikekokeita. Suomessa tärkeimmän osan lajikekokeista muodostavat ns. viralliset lajikekokeet, joita koordinoi maaja metsätalousministeriön toimeksiannosta Maatalouden tutkimuskeskus. Suomalaista lajikekoekäytäntöä ja sen viimeisimpiä tuloksia ovat esitelleet mm. Järvi *et al.* (1995). Patterson & Silvey (1980) ovat kirjoittaneet katsauksen vastaavasta brittiläisestä lajikekoekäytännöstä (statutory variety trials in the UK).

Lajikekokeet ovat yleensä yksinkertaisia koejärjestelyiltään, koska niissä tutkitaan vain yhtä lajikevaikutuksista muodostuvaa koetekijää. Useimmiten koeruudut muodostavat jo peltokuviolina muuttaman täydellisen kerranteen. Niistä jokainen erikseen muodostaa ruodun pitkänomaisia, leveämmältä sivultaan yhteen liittyviä kapeita koeruutuja. Tavallisesti lajikekokeissa käytetään kolmea tai neljää täydellistä kerrannetta. Lajikkeiden sijoitelemiseksi kussakin kerranteessa oleviin koeruutuihin on olemassa monia erilaisia käytäntöjä. Lajikkeet voidaan sijoittaa kerranteen sisältämiin koeruutuihin täysin satunnaisesti (randomized complete blocks). Tavallista on myös satunnaistaa lajikkeet kerranteen sisältämiin koeruutuihin jonkin valitun lohkorakenteen puitteissa. Kokonaisia kerranteita pienempiä lohkoja sisältäviä koekaavioita kutsutaan epätäydellisten lohkojen koekaavioiksi (incomplete block designs). Niitä käyttämällä pyritään kontrolloimaan maan kasvuominaisuuksi-

en vaihtelevuutta. Suomalaisissa kokeissa on yleensä käytetty Cochranin & Coxin (1950) suosittelemia koekaavioita.

Yksittäisten lajikekokeiden suunnittelussa ja analysoinnissa voidaan käyttää apuna esimerkiksi AGROBASE-ohjelmistoa. Lajike- ja lohko vaikutusten estimointi ja niitä koskevat testit ovat usein helposti tehtävissä myös SAS-ohjelmiston avulla. Tällaisista apuneuvoista ja pitkäaikaisesta käytännön harjaantumisesta johtuen yksittäisten kokeiden suunnittelu ja analysointi on jokseenkin yksinkertaista ja rutiininomaista (Mattila & Vähämäki 1994; Mattila *et al.* 1995). Yksittäisten kenttäkokeiden suunnittelu, tekeminen ja analysointi on kuitenkin vain osa lajikekoetoimintaa. Jotta lajikkeiden menestyminen erilaisissa olosuhteissa saataisiin kunnolla testatuksi, tarvitaan yleensä useita vuosia kestäviä ja usealle koepaikalle hajautettuja koesarjoja. Yksittäisten kokeiden sijoituspaikat valitaan siltä alueelta, jolla uusilla testattavilla lajikkeilla uskotaan olevan menestymisen mahdollisuuksia. Käyttökelpoisimmat testaustulokset saadaan yleensä silloin, kun koepaikat muodostavat edustavan otoksen testattavien lajikkeiden koko viljelyalueelta. Yksittäisiä lajikkeita koskevia johtopäätöksiä voidaan yleensä tehdä vasta sitten, kun lajike on ollut kokeissa usean koevuoden ajan. Tavallisesti lajikkeen kokeilu virallisissa lajikekokeissa kestää 4–6 vuotta (Järvi *et al.*, 1995).

Lajikekoeaineistot

Virallisten lajikekoesarjojen analysoinnin lähtökohtana on yleensä yksittäisten kenttäkokeiden analysoinnin lopputuloksena saatu tiivistetty tulosaineisto. Kokeiden tuloksina saatuihin lajikekeskiarvoihin tehdään käytetyn koekaavion mukaiset laskennalliset lohkokorjaukset ja näin saatu aineisto kootaan yhteen. Maatalouden tutkimuskeskuksella on käytettävissään pelto-kasvien virallisista lajikekokeista kootut tiivistetyt tulosaineistot noin kolmen vuo-

sikymmenen ajalta. Nimettyjä lajikkeita koskevien koetulosten lisäksi aineistot sisältävät myös nimeämättömiä kasvilinjoja koskevia koetuloksia. Aineistot on jaoteltu kasvikohteisesti ja niihin viitataan tavallisesti yhteisnimityksellä lajikekoeaineistot.

Edellä kuvatulla tavalla kootun lajikekoeaineiston analysointi on yksinkertaista, jos jokainen samaan vertailuun osallistuva lajike on ollut mukana kaikissa vertailuun käytettävissä kokeissa. Yksinkertainen keskiarvojen laskenta antaa lajikekeskiarvot, sekä koepaikkakohtaiset että vuosittaiset lajikekeskiarvot ja tarvittaessa lajikekeskiarvot vuosien ja koepaikkojen sisällä. Myös lajikevaikutusten eroja on helppo testata yksinkertaisilla varianssianalyysiin liittyvillä lineaarisilla testeillä. Analysoinnin helppoudesta johtuen erityisasemassa ovat sellaiset lajikekoeaineistot, joissa kaikki vertailtavat koeyäsenet ovat mukana kaikissa vertailuun käytettävissä kokeissa. Tilastollisen analyysin kannalta niitä voidaan kutsua täydellisiksi aineistoiksi (complete data). Vastaavasti aineistoa sanotaan epätäydelliseksi (incomplete data), jos kaikki vertailtavat koeyäsenet eivät ole mukana kaikissa kokeissa. Virallisiin lajikekokeisiin tulee vuosittain lisää uusia testattavia lajikkeita ja toisaalta riittävästi testattuja lajikkeita jätetään kokeista pois. Virallisten lajikekokeiden tulosaineistot ovat siten edellä olevassa mielessä kaikkien peltokasvin osalta epätäydellisiä, kun otetaan huomioon kaikki koevuodet, -paikat ja lajikkeet.

Lajikkeiden parittainen vertailu

Edellä määritelty tulosaineiston täydellisyys on suhteellista ja riippuu täysin siitä lajikejoukosta, mikä halutaan keskinäiseen vertailuun. Niinpä virallisten lajikekokeiden tulosaineistosta on helppo muodostaa melkoisen suuriakin osa-aineistoja, joissa jonkin peltokasvin kaksi eri lajiketta ovat

olleet samoissa kokeissa useana vuonna ja usealla koepaikalla. Tällainen osa-aineisto on edellä mainitussa mielessä täydellinen ja sitä voidaan kätevästi käyttää näiden kahden lajikkeen vertaamiseen. Jos kahden lajikkeen aineisto valitaan siten, että toisena on uusi testattava lajike ja toisena jo koeteltu ja tunnettu lajike (mittarilajike), saadaan uuden lajikkeen ominaisuuksia helposti selville vertaamalla tuloksia tunnettuun mittarilajikkeeseen.

Testattavan lajikkeen ja tunnetun mittarilajikkeen muodostama osa-aineisto on ollut tähän asti Maatalouden tutkimuskeskuksen koordinoimien virallisten lajikekokeiden analysoinnin pääasiallinen lähtökohta (Järvi *et al.*, 1995). Tällaisen osa-aineiston tilastollinen analyysi pelkistyy yksinkertaiseksi keskiarvojen vertailuksi, jota kutsutaan parivertailuksi. Johtopäätösten tekemisen kannalta parivertailu on kuitenkin puutteellinen analysointitapa. Parivertailussa käytettävien osa-aineistojen valinta riippuu sekä mittarilajikkeesta että vertailtavana olevasta toisesta lajikkeesta, jolloin vertailu perustuu kunkin parin osalta aina erilaiseen koemateriaaliin. Tästä johtuen vertailutulokset saattavat olla keskenään hyvinkin ristiriitaisia, minkä lisäksi ristiriitaisia tuloksia saattaa tulla myös verrattaessa testattavaa lajiketta moniin vaihtoehtisiin mittarilajikkeisiin. Lisäksi, osa-aineistojen käyttö jättää aina huomiotta osia alkuperäisestä aineistosta, eivätkä saavat testitulokset yleensä ole niin voimakkaita kuin ne voisivat olla koko alkuperäistä aineistoa käytettäessä. Testien heikentyminen vaikeuttaa johtopäätösten tekoa eikä todellisuudessa merkittäviäkään eroja voida aina tilastollisesti vahvistaa. Parivertailun kolmas puute on se, että häiritsevien ulkoisten tekijöiden vaikutus testaustilanteissa voidaan kyllä poistaa, mutta sen avulla ei saada mitään viitteitä lajikkeiden herkyydestä erilaisille ulkoisille olosuhteille. Todellisuudessa kasvinjalostajat ja päätöksentekijät tarvitsisivat usein myös tällaista tietoa.

Lajikkeiden samanaikainen vertailu

Parivertailumenetelmän valinta virallisten lajikekekoesarjojen pääasialliseksi vertailukeinoksi perustui aikanaan sen näennäisen helppoon ymmärrettävyyteen ja laskentatyön yksinkertaisuuteen. Nykyisin nämä tekijät eivät enää ole yhtä merkittäviä. Menetelmätieteiden kehityksen ja tietokoneiden laskentakapasiteetin lisääntymisen myötä kokeellisessa tutkimuksessa on totuttu käyttämään tutkittavia ilmiöitä yhä tarkemmin kuvaavia tilastollisia malleja. Erityisesti lineaarisia sekamalleja (mixed linear models) on viime vuosikymmeninä kehitetty yhdeksi tarkkuuteen pyrkivän mallintamisen perusvälineistöksi. Lineaaristen sekamallien perusominaisuuksiin kuuluu, että ne antavat mahdollisuuden mallintaa valitun vastemuuttujan sekä odotusarvon että hajonnan riippuvuutta valituista selittävistä muuttujista samanaikaisesti. Lineaariset sekamallit antavat myös mahdollisuuden käsitellä suuria muuttuja-joukkoja samanaikaisesti. Näiden ominaisuuksiensa ansiosta lineaariset sekamallit oikein käytettyinä soveltuvat hyvinkin monimuotoisten tutkimusaineistojen analysointiin. Talbot (1984) oli ensimmäisten joukossa soveltamassa sekamalleja lajikekoesarjojen analysointiin. Patterson & Nabugoomu (1992) ovat kirjoittaneet hyvän katsauksen analysointitapojen kehityksestä ja nykyisistä mahdollisuuksista. Seuraavassa on tarkoitettu esitellä yksityiskohtaisten esimerkkien avulla niitä keinoja ja mahdollisuuksia, joita lineaaristen sekamallien käyttö antaa suomalaisten virallisten lajikekoesarjojen analysoimiseen.

Lineaaristen sekamallien käytön tekemisenä esteenä viime vuosiin asti on ollut niiden käytössä usein tarvittava suuri laskentatyö ja helppokäyttöisten valmisohjelmistojen puute. Nykyisillä tietokoneilla helposti ajettavat lineaaristen sekamallien laskenta-algoritmit sisältyvät kuitenkin jomman GENSTAT-, BMDP- ja SAS-sovellus-

paketteihin. Valmiit algoritmit soveltuvat hyvin myös virallisten lajikekokeiden tulosaineistojen analysointiin. Öfversten (1995) on esittänyt myös kokonaan uudentyyppisiä estimointi- ja testausmenetelmiä ja esimerkin niiden soveltamisesta suomalaisten vehnälaajikkeiden satotulosten ennustamiseen ja varianssikomponenttien testaamiseen. Teknisiä esteitä lineaaristen sekamallien käytölle ei siten enää ole. Lähemmin selvitettäväksi jää vain se, miten tarkasti lineaariset sekamallit käytännössä taipuvat erilaisten lajikekoesarjojen perusteella saatujen tulosaineistojen mallintamiseen ja miten hyvin näiden mallien antamia ennuste- ja testituloksia voidaan soveltaa päätöksenteossa.

Mallintamisen mahdollisuudet

Peltokasvien virallisia kokeita on viime vuosina järjestetty noin 20 eri koepaikalla, joista useimmat ovat Maatalouden tutkimuskeskuksen tutkimusasemia. Kokeita järjestetään nykyään noin sata vuodessa, aiemmin niitä on ollut enemmänkin. Lajikekoeaineistoissa on niiden alkuperäisestä keräystavasta johtuen luokittelutekijöinä ainakin lajiketekijä (lajike), vuositekijä (vuosi), koepaikkatekijä (paikka) ja koetunniste eli koetekijä (koe). Koetekijää käytetään ensisijaisesti koepaikan sisäisenä tunnisteena, jonka avulla voidaan jäljittää samaan koesarjaan kuuluvat kokeet. Tämä tunniste saattaa pysyä samana useita vuosia, joten sellaisenaan se ei erottele vuosittaisia kokeita. Tiettyinä vuonna suoritettun yksittäisen kokeen tunnistamiseksi tulee määrittellä sekä vuosi, koepaikka (paikka) että kokeen tunnistetiedot (koe). Kaikki lajikekoeaineistot sisältävät myös lajikkeiden ominaisuuksia kuvaavia jatkuvaluonteisia muuttujia, jotka on laskettu kullekin lajikkeelle koekohtaisesti, lohko- ja kokekohtaisesti, lohkovaikutuksista johtuvat korjaukset huomioon ottavina kerranteiden keskiarvoina. Testaus- ja ver-

tailutilanteissa tavallisimmin käytetty lajikeominaisuus on näin korjaamalla saatu jyväsato (sato).

Lineaarisiin sekamalleihin liittyvät selittävät tekijät määritellään aineiston kokoaamistavan ja mallien käyttötarkoituksen mukaisesti joko kiinteävaikutteisiksi (fixed effects) tai satunnaisvaikutteisiksi (random effects). Lajikekoeaineistojen kokoaamistavasta takia lajiketekijä on aina kiinteävaikutteinen ja vuositekijä satunnaisvaikutteinen. Koepaikka- ja koetekijöiden satunnaisuus ei ole aivan yhtä yksiselitteinen. Lajikekoetoiminnan lähtökohtana on aina ollut se, että koepaikat valitaan jossain määrin satunnaisesti edustamaan koko tarkastelun alaista koealuetta ja yksittäisten kokeiden paikat valitaan satunnaisesti kunkin koepaikan koetoimintaan käyttämien peltoalueiden sisältä. Tällöin on perusteltua ainakin aloittaa mallintamisen pitämällä koe- ja koepaikkatekijöitä satunnaisvaikutteisina (vrt. Patterson & Nabugoomu 1992). Tilastollisen mallintamisen peruseräitä on myös se, että yhdysvaikutustekijää pidetään aina satunnaisvaikutteisena, jos sen yksikin osatekijä on satunnaisvaikutteinen. Täten tavanomaisten lajikekoeaineistojen mallintaminen voidaan yleensä aloittaa kokeilemalla lineaarisia sekamalleja, joissa vain lajiketekijä on kiinteävaikutteinen ja kaikki muut mahdolliset selittävät tekijät ovat satunnaisvaikutteisia. Kaikki lajikekoaineistoihin liittyvät satunnaisvaikutteiset tekijät voidaan perustellusti olettaa myös toisistaan riippumattomiksi. Erittäin hyvä lähtökohta kaikkia vuosia ja koko viljelyaluetta koskevien estimointi- ja vertailutulosten aikaansaamiseksi on se, että mallin ainoa kiinteä tekijä on lajiketekijä ja kaikki muut ovat satunnaistekijöitä.

Lajikekoeaineistojen kuvaamisen lähtökohtana voidaan siis pitää sekamalleja, joissa lajiketekijä on ainoa mahdollinen kiinteävaikutteinen tekijä. Malliin sisällytettävien satunnaistekijöiden valinnassa on kuitenkin valinnan varaa. Periaatteellinen maksimimäärä selittäviä tekijöitä saadaan määrittelemällä mahdollisiksi vaihtelun

lähteiksi kaikki kolme satunnaistekijää (vuosi, paikka, koe) ja kaikki mahdolliset neljän tekijän (lajike, vuosi, paikka, koe) yhdysvaikutuskombinaatiot.

Käyttökelpoisten mallien etsintä perustuu eri mallivaihtoehtojen sovittamiseen ja näihin liittyvien varianssikomponenttien testaamiseen. Erilaisten kokeiltavien vaihtoehtojen valinta taas riippuu tutkijan ennako-odotuksista sekä estimoinnin, ennustamisen ja testaamisen tarpeista. Mikään malli sinänsä ei ole ainoa oikea, vaan hyvinkin erilaisia malleja voidaan käyttää päätöksenteon pohjana.

Mallien kokeilemisen ensimmäisessä vaiheessa on tarkoitus tunnistaa selitettävän muuttujan eli vastemuuttujan erilaiset vaihtelun lähteet. Mallintamisen toisessa vaiheessa valitaan ja määritellään jokin käyttökelpoinen lineaarinen malli, joka ottaa huomioon tunnistetut vaihtelun lähteet ja estimoidaan tähän malliin liittyvät tuntemattomat parametrit. Käytännössä kaikkien mahdollisten vaihtelun lähteiden huomioiminen heti mallintamisen alkuvaiheessa johtaa usein turhan monimutkaisuuteen ja käyttöä selittäjinä vain jäljelle jääviä kolmea koetekijää. Näin tekemällä voidaan SAS-ohjelmiston (SAS Technical Report P-229) MIXED-proseduurin määrittelykieltä käyttäen kirjoittaa satotulosten odotusarvon selittämiseksi seuraava esimerkinomainen lineaarinen sekamalli:

```
PROC MIXED;  
CLASS lajike vuosi paikka;  
MODEL sato = lajike;  
RANDOM vuosi paikka vuosi*paikka  
        lajike*vuosi lajike*paikka;  
RUN;
```

Koevaikutusten samaistaminen jäännösvirheeseen on esimerkissä toteutettu teknisesti yksinkertaisesti jättämällä aineistoon sinänsä kuuluva koetekijä (koe) kokonaan pois mallin määrittelystä. Jo tämän

alustavan mallin avulla voidaan tehdä monia johtopäätöksiä satotuloksiin vaikuttavista vaihtelun lähteistä. Koevaikutuksia sisältävien tekijöiden lisääminen malliin on kuitenkin yleensä tarpeen mallin tarkentamiseksi. Se on paikallaan erityisesti silloin, jos mallin avulla pyritään löytämään muuhun aineistoon verrattuna poikkeuksellisia kokeita. Esimerkkejä mallien tarkentamisesta ja niiden avulla tehtävistä johtopäätöksistä esitetään jäljempänä.

Esimerkkiaineisto 1: syysvehnä

Viljelytekniikan ja lajikkeiden kehityksestä johtuen satotulokset nousevat jatkuvasti. Tästä syystä käytännön lajikevertailuihin ei yleensä oteta kaikkein vanhimpia kokeita. Testaustilanteissa kannattaa keskittyä vertailemaan vain todella kiinnostavia lajikkeita, koska turhien vertailujen tekeminen heikentää testituloksia. Seuraavassa tarkastellaan esimerkinomaisesti näiden näkökohtien mukaisesti valittua syysvehnän lajikekoeaineiston osa-aineistoa vuosilta 1981–1994. Valittu osa-aineisto (koesarja) sisältää kahdeksaa lajiketta ja kahdeksaa koepaikkaa koskevia lajikekokeiden tuloksia 14 vuoden ajalta. Soveltamalla edellä esitettyä ajomääritystä tähän osa-aineistoon saadaan SAS MIXED-proseduurin avulla liitteellä 1 oleva tulostus. Tästä tulostuksesta voidaan heti tehdä mm. seuraavat johtopäätökset:

- Lajikkeiden satoisuudessa on selkeitä eroja. Tämä näkyy siitä, että lajikkeiden satoeroja testaava F-testi on merkitsevä ($p = 0,0001$).
- Lajikkeiden satotulosten erot ovat kuitenkin herkkiä säätekijöille, ts. lajikkeiden väliset erot vaihtelevat vuosittain säätekijöistä johtuen. Tämä näkyy siitä, että lajike x vuosi -yhdysvaikutukseen liittyvä varianssikomponentti on tilastollisesti nolasta eroava ($p = 0,0429$).

- Lajikkeiden satotulosten vaihtelua ei voida selittää koko koeluetta koskevilla yhteisillä sääoloilla. Tämä näkyy siitä, että vuositekijään liittyvä varianssikomponentti on tilastollisesti merkityksetön ($p = 0,2052$). Edellä mainittuun tulokseen yhdistettynä tämä voitaneen tulkita siten, että koepaikkakohtaiset sääerot ovat sen verran suuria, että ne peittävät alleen keskimääräiset vuosivaikutukset.
- Lajikkeiden satotulosten vaihtelua ei voida selittää koepaikkojen erilaisuudella. Tämä näkyy siitä, että koepaikkatekijään liittyvä varianssikomponentti ($p = 0,1582$) ja lajike x koepaikka -tekijään liittyvä varianssikomponentti ($p = 0,3238$) ovat tilastollisesti merkityksettömiä. Jälkimmäisen tuloksen mukaisesti voidaan todeta, että lajikkeet eivät ole herkkiä koepaikkojen erilaisuudelle. Tämä tulos saattaa olla merkki siitä, että kokeet on suoritettu turhan samankaltaisissa koeolosuhteissa. Suurempi vaihtelu koepaikkojen välillä saattaisi paljastaa nyt piiloon jääviä herkkyyseroja.
- Suurin vaihtelun lähde on vuosi x paikka -yhdysvaikutustekijä ($p = 0,0000$). Tämä voi johtua siitä, että paikalliset sääolot todella vaihtelevat suuresti vuodenajan mukaan tai siitä, että samantapaisetkin sääolot kunkin koepaikan paikallisiin maalaajiin tai muihin kasvutekijöihin yhdistettynä saavat aikaan sadonmuodostuksen kannalta erittäin vaihtelevat kasvuolosuhteet.
- Edellä kuvattujen testitulosten mukaisesti kokeiltavana oleva malli sisältää satunnaistekijöitä, joiden varianssikomponenttien ei voida väittää eroavan nolasta. Siksi malli sellaisenaan ei ole paras mahdollinen. Jatkoanalyysyjä varten sitä on syytä tarkentaa.

Mallin tarkentaminen

Lajikkeen herkkyys ympäristövaikutuksille on tärkeä ominaisuus. Siksi edellä oli perusteltua aloittaa mallin etsintä kokeilemalla aluksi mallia, jossa olivat mukana ainakin herkkyyttä mittaavat yhdysvaikutustekijät lajike x vuosi ja lajike x paikka. Mikäli nämä tekijät voidaan perustellusti pitää mallissa mukana, siitä on yleensä hyötyä jatkoanalyysissä ja päätöksenteossa. Mallinetsinnän varsinaisena tarkoituksena on kuitenkin löytää mahdollisimman yksinkertainen malli, joka ei sisällä muita kuin vaihtelua todella aiheuttavia tekijöitä. Vain tällaista aineiston kanssa yhteensopivaksi todettua mallia voidaan jatkossa käyttää estimointiin, ennustamiseen ja tilastollisiin testeihin.

Koska edellä olevassa esimerkkitapauksessa vuosi-, koepaikka- ja lajike x koepaikka -tekijöillä ei varianssikomponentteihin liittyvien testien mukaan näyttänyt olevan merkitystä, voidaan jatkoanalyysissä mallia tarkentaa jättämällä nämä turhiksi todetut tekijät pois. Kun malliin lisätään uusia mahdollisesti vaihtelun lähteinä toimivia koetekijään liittyviä yhdysvaikutustekijöitä päädytään seuraavaan kokeiltavaan malliin:

```
PROC MIXED;  
CLASS lajike vuosi paikka koe;  
MODEL sato = lajike;  
RANDOM lajike*vuosi vuosi*paikka  
paikka*koe vuosi*paikka*koe;  
RUN;
```

Yksittäiset koevaikutukset ovat koepaikkojen sisäisiä satunnaisvaikutuksia, joten niiden itsenäisiä päävaikutuksia ei ole järkevää sisällyttää malliin. Malliin ei myöskään kannata enää ottaa mukaan uusia lajikevaikutuksiin liittyviä korkeampiasteisia yhdysvaikutuksia, koska mallin tarkentaminen ei sitä välttämättä vaadi eikä näillä olisi käyttöä tulosten tulkinnassa tai päätöksenteossa. Liitteellä 2 on ylläolevan SAS MIXED -proseduurin avulla saatu tulostus.

Tämän tulostuksen mukaisesti yhdysvaikutustermiin paikka x koe liittyvän varianssikomponentin ($p = 0,2684$) ei voida väittää eroavan nolasta. Tämä viittaa taas hyvin homogeenisiin koeolosuhteisiin. Jättämällä mallin kannalta turha yhdysvaikutustermi pois saadaan esimerkkiaineiston satotulosten mallintamiseksi seuraava lineaarinen sekamalli:

```
PROC MIXED;  
CLASS lajike vuosi paikka koe;  
MODEL sato = lajike;  
RANDOM lajike*vuosi vuosi*paikka  
vuosi*paikka*koe;  
RUN;
```

Liitteellä 3 on tätä SAS-proseduuria vastaava ajotulos. Ajotulokseen sisältyvien testiarvojen mukaisesti tämä malli on syysvehnäaineiston kanssa siinä määrin yhteensopiva, että sitä voidaan perustellusti käyttää jatkoanalyysiin.

Poikkeuksellisten kokeiden tunnistaminen

Lajikekokeita tehdään paljon. Niinpä yksittäisten kokeiden tulosten vaihtelua ei aina voida selittää pelkästään lajike-, vuosi-, koepaikka- tai koetekijöillä. Syynä poikkeukselliseen tulokseen saattaa olla esimerkiksi täysin epäonnistunut kokeen sijoituspaikan valinta tai koeprosessin aikana tapahtunut työvirhe. Tavanomaisen laatukontrolliperiaatteiden mukaisesti tällaiset kokeet olisi voitava tunnistaa. Päätöksenteon kannalta on usein tarkoituksenmukaista jättää muusta aineistosta selvästi poikkeaviksi todetut osat pois kokonaisanalyysistä.

Virallisten lajikekokeiden työprosessissa poikkeavat kokeet voidaan tunnistaa ainakin kolmessa työvaiheessa. Jo pellolla kenttäkokeiden aikana voidaan havaita koetuloksiin vaikuttava työvirhe. Jokaisen yksittäisen kokeen erillisen tilastollisen

analyysin tekovaiheessa voidaan joskus havaita, että koe on selvästi muista poikkeava. Tämä voi ilmetä esimerkiksi poikkeuksellisen matalana satotasona, suurena jäännösvirheenä tai suurena kerranteiden tai lohkojen välisenä vaihteluna. Syyinä voi olla esimerkiksi poikkeukselliset sääolot tai liian heterogeeninen kokeen suorituspaiikka. Poikkeavia kokeita voidaan tunnistaa myös analysoitaessa yhteenkootuja lajikekoeaineistoja ja tutkimalla niissä käytettyihin tilastollisiin malleihin sisältyviä estimoituja parametrejä.

Edellä kaksi ensiksi mainittua laatu-kontrollin keinoja eivät kuulu tämän kirjoituksen aihepiiriin. Sen sijaan näytämme seuraavassa esimerkinomaisesti miten poikkeavien kokeiden jäljittäminen voidaan tehdä kolmantena mainittua keinoja käyttäen, ts. koesarjan analysoimiseksi valitun sekamallin avulla. Esimerkissä jatketaan jo aloitetun syysvehnäaineiston analysointia.

Poikkeavien kokeiden tunnistamiseen voidaan käyttää vuosi x paikka x koe -tekijään liittyviä estimoituja satunnaisvaikutuksia. Oikealla tavalla suoritetuissa kokeissa tämän tekijän satunnaisvaikutukset ovat toisistaan riippumattomia, normaalijakautuman mukaisesti jakautuneita satunnaisuuttujia, joiden keskiarvo on nolla. Jos jonkun yksittäisen satunnaisvaikutuksen jakauman voidaan todeta poikkeavan tästä yleiskaavasta, kyseessä on poikkeava vuosi x paikka x koe -vaikutus eli poikkeava koe. Jotta estimoidut satunnaisvaikutukset saataisiin näkyville, tulee edellä olevan SAS MIXED-proseduurin RANDOM-komentoon lisätä /SOLUTION-optio seuraavasti:

```
PROC MIXED;
CLASS lajike vuosi paikka koe;
MODEL sato = lajike;
RANDOM lajike*vuosi vuosi*paikka
      vuosi*paikka* koe / SOLUTION;
RUN;
```

Liitteellä 4 on tätä ajomääritystä vastaava tulostus. Tässä tulostuksessa on jokaista erillistä satunnaisvaikutusta vastaava t-arvo

ja p-arvo, joita käyttämällä voidaan tunnistaa poikkeukselliset kokeet. Kyseessä oleva monivertailutilanne pitää tietysti ottaa huomioon. Yksinkertainen tapa on käyttää Bonferronin esittämää testausmallia (vrt. Miller 1981). Tällöin tulee aluksi valita testauksen kriittinen taso (experimentwise error rate) ja laskea uusi yksittäisissä testauksissa tarvittava kriittinen taso (comparisonwise error rate) tätä ja estimoitujen satunnaisvaikutusten lukumäärää vastaavaksi. Poikkeavat kokeet voidaan nyt tunnistaa vertaamalla satunnaisvaikutuksiin liittyviä alkuperäisiä p-arvoja uuteen laske-malla saatuun kriittiseen tasoon.

Testausutilanteen kriittiseksi kokonaistasoksi valitaan usein 0,05 eli 5 %. Vuosi x paikka x koe -vaikutuksia eli yksittäisiä kokeita on valitussa koesarjassa yhteensä 158. Bonferronin testin mukaisesti voidaan yksittäiset kokeet todeta poikkeaviksi, jos niihin liittyvä p - arvo on pienempi kuin $0,05/158 = 0,00032$ (liite 4). Esimerkkiajossa tällaisia kokeita ei ole. Tämä on osoitus siitä, että vuosien ja koepaikkojen sisällä kokeet ovat hyvin tasalaatuisia ja samansuuntaisesti käyttäytyviä. Tosiasia lienee myös se, että poikkeaviksi epäiltyjä kokeita on jo muilla perusteilla poistettu lajikekoeaineistosta.

Mielenkiintoinen sivutuloksena saadaan kun tutkitaan edellä kuvatulla tavalla vuosi x koepaikka -vaikutuksia. Näitä on yhteensä 76. Valitsemalla päätöstilanteen kriittiseksi kokonaistasoksi 5 % voidaan vuosi x koepaikka -vaikutus todeta poikkeavaksi, jos sitä liitteellä 4 vastaava p - arvo on pienempi kuin Bonferronin menetelmän mukaisesti laskettu kriittinen arvo $0,05/76 = 0,00066$. Syysvehnäaineistossa on ainakin 17 tämän kriteerin mukaisesti poikkeavaa vuosi x paikka -yhdistelmää. Tämä osoittaa, että valitut koepaikat eivät muodosta yhtenäistä viljelyaluetta. Jatkoanalyysien ja viljelysuositusten tekemistä varten nyt tarkasteltavaa syysvehnäaineistoa pitäisi vielä alueellisesti rajata. Erityisesti koepaikka 19 näyttää poikkeavan muusta viljelyalueesta.

Estimointi, ennustaminen ja testaaminen

Edellä on tarkasteltu syysvehnän lajikekoeaineiston kuvaamiseksi sopivan lineaarisen mallin etsintää ja poikkeavien kokeiden tunnistamista. Kun käyttökelpoinen malli on lopullisesti määritelty ja käytettävä aineisto valittu, voidaan lopullisesti estimoida myös kiinteät lajikevaikutukset. Käyttämällä lisäksi koepaikkoihin liittyviä estimoituja satunnaisvaikutuksia voidaan ennustaa esimerkiksi testattavien lajikkeiden satoisuutta eri koepaikoilla. Tämänlaisia ennusteita voidaan tehdä SAS MIXED -proseduuriin liittyvän ESTIMATE -komennon avulla. Kuvatussa esimerkkitaapauksessa ei kannata lähteä ennustamaan eri koepaikkojen satutuloksia, koska tehtyjen testausten mukaisesti koepaikoilla ei ole vaikutusta. Tässä tapauksessa riittää, että lajikekeskiarvot saadaan harhattomasti estimoiduiksi. SAS MIXED -proseduurin avulla tämä voidaan tehdä kaikkein yksinkertaisimmin lisäämällä MODEL -komentoon / NOINT- ja SOLUTION -optiot.

Pelkkä lajikekeskiarvojen estimointi ei usein riitä, vaan lajikekeskiarvoja on myös voitava vertailla. SAS MIXED -proseduuriin automaattisesti liittyvä F-testi kertoo heti, eroavatko kiinteät lajikevaikutukset toisistaan. Yleisluontoinen johtopäätös lajikkeiden eroista ei analyysissä useinkaan riitä, vaan lajikkeita on voitava verrata sekä toisiinsa että tunnettuihin mittarilajikkeisiin. Tällaisia tarkentavia testejä voidaan suorittaa, kun käytetään estimoituja lajikekeskiarvoja. Usein näitä jatkotestejä kutsutaan joko kontrasteiksi tai lineaarisiksi testeiksi. Niitä voidaan tehdä joko yksittäin tai kytkemällä useita lineaarisia testejä samaan päätöstilanteeseen. Testaustavan valinta tehdään sen mukaan miten päätöstilanteen virhetodennäköisyyksiä halutaan kontrolloida.

Käytännössä vertailuja voidaan tehdä jo edellä mainittua ESTIMATE -komentoa

käyttäen. Tavallisin lajikevertailujen päätöstilanne lienee monivertailutilanne, jossa kaikkien testaustilanteeseen liittyvien värijohtopäätösten todennäköisyyden (experimentwise error rate) halutaan samanaikaisesti pysyvän aikaisemmin sovitua virherajaa pienempänä. Yksittäisten testien voimakkuus heikkenee nopeasti päätöstilanteeseen liittyvien testien lukumäärän lisääntyessä, joten monivertailutilanne edellyttää päätöksentekijöiltä tarkkuutta ja harkintaa. Kaikki yksittäiset vertailut kannattaa määritellä edeltä käsin ja harkita niiden todellinen tarve. Testien voimakkuus on huonoimmillaan silloin, kun jokaista lajiketta verrataan samanaikaisesti kaikkiin muihin. Testien lukumäärää voidaan vähentää ja voimakkuutta parantaa mm. vertaamalla kokeiltavia uusia lajikkeita vain yhteen tai muutamaaan tunnettuun mittarilajikkeeseen.

Seuraavaan ajomääritykseen on lisätty jo edellä mainitut NOINT- ja SOLUTION -optiot lajikekeskiarvojen estimoimiseksi ja ESTIMATE -komentoja, joilla verrataan mittarilajike Auraa (103045) seitsemään muuhun koesarjaan kuuluvaan lajikkeeseen.

```
PROC MIXED;
CLASS lajike vuosi paikka koe;
MODEL sato = lajike / NOINT
SOLUTION;
RANDOM lajike*vuosi
vuosi*paikka vuosi*paikka* koe;
ESTIMATE ' 100835 - 103045 '
lajike 1 0 -1 0 0 0 0 0;
ESTIMATE ' 101851 - 103045 '
lajike 0 1 -1 0 0 0 0 0;
ESTIMATE ' 103063 - 103045 '
lajike 0 0 -1 1 0 0 0 0;
ESTIMATE ' 103077 - 103045 '
lajike 0 0 -1 0 1 0 0 0;
ESTIMATE ' 103079 - 103045 '
lajike 0 0 -1 0 0 1 0 0;
ESTIMATE ' 222585 - 103045 '
lajike 0 0 -1 0 0 0 1 0;
ESTIMATE ' 523977 - 103045 '
lajike 0 0 -1 0 0 0 0 1;
RUN;
```

Liitteellä 5 on edellä olevaa ajomääritystä vastaava tulostus. Tähän tulostukseen on SOLUTION-option vaikutuksesta tullut myös tarpeettomia testiarvoja sille, eroavatko lajikekeskiarvot nolasta. ESTIMATE-komennon avulla on saatu lajikekeskiarvojen erojen estimaatit sekä t-arvot ja p-arvot yksittäisten testien suorittamiseksi. Oikeat lajikenimet käyttöön ottamalla voidaan liitteen 5 avulla laatia jäljempänä oleva vertailutaulukko (Taulukko 1).

Jos kaikkien vertailutaulukossa esitettyjen seitsemän testin kriittisen tason halutaan samanaikaisesti olevan 5 %, tulee yksittäisten vertailujen kriittistä tasoa tiukentaa esimerkiksi edellä kuvatulla Bonferronin esittämällä tavalla, ts. valitsemalla yksittäisten testien kriittiseksi tasoksi $0,05/7 = 0,0071$. Näin menetelmällä huomataan, että vain lajikkeen Vakka -71 (101851) satotulos on mittarilajikkeesta (Aura) eroava. Uusimmassa SAS-versiossa (SAS 6.10) on valmiiksi ohjelmoitu myös monia muita monivertailutestejä. Näistä monet (esim. Tukey-Kramer) ovat Bonferronin menettelyä voimakkaampia mutta lajikevertailuissa tuskin tarpeellisia.

Herkkien lajikkeiden tunnistaminen

Lajikkeet ovat usein eritavoin herkkiä (sensitive) ympäristövaikutuksille, toisin sanoen jotkut lajikkeet ovat viljelyolosuhteiden suhteen toisia sopeutuvaisempia (adaptive). Viljelysuosituksiin ja viljelypäättöksiin sisältyvien riskien arvioimiseksi on hyödyllistä tunnistaa lajikkeiden herkkyyserot. Herkkyyttä voidaan testata usealla eri tavalla. Esimerkiksi tutkimalla lajike x koepaikka -tekijään liittyvää estimoitua satunnaisvektoria voidaan selvittää, mitkä lajike ja koepaikkayhdistelmät ovat erityisen hyviä tai huonoja. Vastaavasti lajike x vuosi -tekijään liittyvää satunnaistekijää tutkimalla voidaan päätellä, mitkä lajike ja vuosiyhdistelmät ovat olleet erityisen oikeukkaita. Näin voidaan selvittää lajikkeiden herkkyyttä erikseen koepaikka- ja säävaikutuksille. Varsinainen selvitystyö voidaan tehdä käyttämällä samoja tilastollisia testejä kuin poikkeavien kokeiden jäljittäminen (katso liite 4).

Edellä olevien varianssikomponentteja koskevien testien mukaisesti vehnälaajikkeet ovat kuitenkin herkkiä vain vuosi/säävaikutuksille mutta eivät koevaikutuksille. Liitettä 4 tutkimalla havaitaan lisäksi, että

Taulukko 1. Eräiden syysvehnälaajikkeiden vertailu Aura-lajikkeeseen.

Lajike	Satoero (t/ha)	Hajonta	Vapausasteet	t-arvo	p-arvo
Nisu	-0,130	0,147	584	-0,88	0,379
Vakka -71	-0,518	0,107	584	-4,81	0,000
Pitko	-0,166	0,114	584	-1,46	0,144
Jo 3077	0,290	0,218	584	1,33	0,184
Jo 3079	-0,169	0,179	584	-0,94	0,346
Limna P	0,014	0,130	584	0,11	0,911
Folke	0,049	0,154	584	0,32	0,746

yhdenkään yksittäisen lajike x vuosi -vaikutuksen ei voida väittää olevan nollasta poikkeava. Jos tällaisia poikkeavia vaikutuksia löytyisi, pitäisi seuraavaksi selvittää säätilastoista, mitä poikkeavaa sisältyy näiden vuosien säätietoihin. Joissakin tapauksissa tällä tavalla voitaisiin ehkä jäljittää esimerkiksi talvenkestävyyttä ja/tai kuivuuden sietokykyä.

Perinteinen sopeutumisen (adaptation) ja herkkyuden (sensitivity) määrittäminen perustuu regressioanalyysin käyttöön. Tätä keinoa ovat suositelleet mm. Finlay & Wilkinson (1963). Keinon esittivät ensimmäisinä Yates & Cochran (1938). Regressioanalyysiin pohjautuva herkkyuden mittaaminen perustuu ensisijaisesti siihen, miten muut kokeissa mukana olleet lajikkeet ovat menestyneet. Se ei erottele sää- ja koepaikkavaikutuksia. Sen sijaan sen avulla voidaan tunnistaa kasvuolosuhteille yleisesti ottaen eritavoin herkkät kasvilajikkeet. Se antaa myös keinon verrata eri lajikkeiden sopeutumiskykyä.

Regressioanalyysiin perustuvassa herkkyuden määrittämisessä käytetään vaste-muuttujana kokeessa kullekin lajikkeelle saatua satoa ja selittävänä regressiomuuttujana vastaavan kokeen keskisatoa. Ykköstä suurempi regressiokerroin viittaa keskimääräistä suurempaan ja ykköstä pienempi keskimääräistä pienempään herkkyuteen ympäristövaikutuksille. Herkkyyttä mittaava regressiokerroin (herkkyyskerroin) ei kuitenkaan yksinään riitä lajikkeen viljelyarvon määrittämiseksi. Sen lisäksi tarvitaan tunnusluku, joka mittaa lajikkeen keskimääräistä satoisuutta. Tällaisena tunnuslukuna voidaan käyttää liitteellä 4 olevia lajikkeiden keskisatoestimaatteja tai regressioanalyysin avulla saatavia vakiotermien estimaatteja. Kunkin lajikkeen sopeutuminen voidaan siten ilmaista aina kahden tunnusluvun avulla.

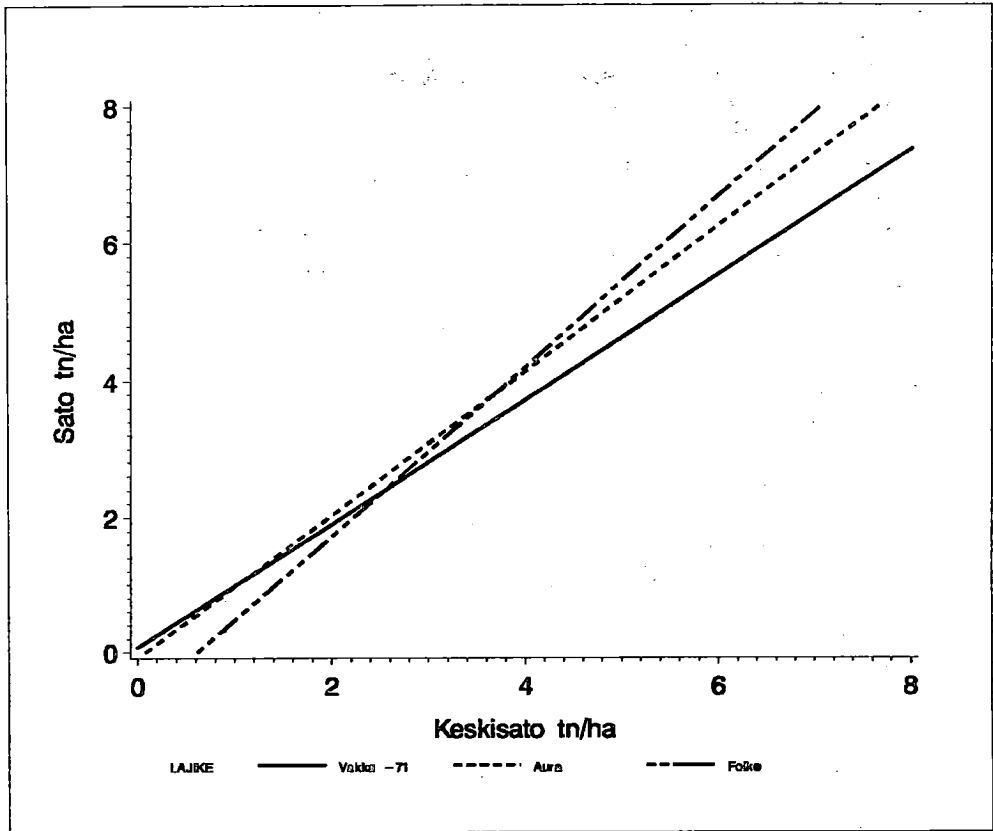
Finlay & Wilkinsonin (1963) laskivat kullekin lajikkeelle erilliset regressiokerroimien ja niitä vastaavien vakiotermien estimaatit suorittamalla regressioanalyysin erikseen kutakin lajiketta vastaavalle lajikekoaineiston osa-aineistolle. Lajikekohtai-

set herkkyuden tunnusluvut saadaan kuitenkin tarkemmin estimoiduiksi käyttämällä samalla kertaa koko aineistoa, mutta estimoimalla silti kullekin lajikkeelle erilliset regressiokertoimet ja lajikevaikutukset (vakiotermit). Tämä on helposti tehtävissä yleisten lineaaristen mallien periaatteiden mukaisesti. Käytettävää mallia valittaessa on kuitenkin otettava huomioon, että malliin selittäjiksi tulevat kokeiden keskisadot selittävät samaa vaihtelua kuin edellä olevissa sekamalleissa selittäjinä käytetyt satunnaistekijät, joten näitä satunnaistekijöitä ei mallissa tarvita. SAS MIXED -proseduurin avulla tarvittava malli voidaan kirjoittaa seuraavasti:

```
PROC MIXED;  
CLASS lajike vuosi paikka koe;  
MODEL sato = lajike lajike*ksato  
/ NOINT SOLUTION;  
RUN;
```

Liitteellä 6 on ajomääritystä vastaava tulostus. Käytetty SAS -proseduuri antaa myös tässä yhteydessä tarpeettomat testit sille, eroavatko herkkyyttä kuvaavat regressiokertoimet nollasta. Todellisuudessa tarvitaan testit sille, eroavatko ne ykköses-tä. Tällaisiin testeihin tarvittavat uudet t-arvot on helppo laskea SAS-tulostuksen avulla vähentämällä estimoidusta herkkyyskerroimesta ykkönen ja jakamalla saatu luku SAS-tulostuksesta löytyvällä herkkyyskerroimen hajonnalla. Moniver-tailutilannetta varten tarvitaan myös näitä t-arvoja vastaavat *p* -arvot, jotka on yksinkertaisinta katsoa jostain painetusta taulukosta. Taulukkoon 2 ne on laskettu SAS IML -ohjelmiston avulla.

Tunnuslukujen avulla voidaan kuvata lajikkeiden satoisuutta viljelyominaisuuk-siltaan erilaisissa kasvuympäristöissä. Hyvä yleiskäsitys saadaan, kun piirretään lajikkeen antaman sadon ja keskimääräisen sadon riippuvuutta kuvaavat suorat samaan koordinaatistoon. Kuvaan 1 on piirretty kolmen jäljempänä tarkasteltavan lajikkeen herkkyyttä kuvaavat regressioviivat. Tätäkin tarkemmin voidaan kasvuympäristöl-



Kuva1. Syysvehnäajikkeiden herkkyys kasvuypäristölle.

Taulukko 2. Syysvehnäajikkeiden herkkyyttä kuvaavia tunnuslukuja.

Lajike	Lajike-vaikutukset	Lajikevaik. hajonta	Herkkyyskerroin (k)	Herk.kert. hajonta	Vapausasteet	t-arvo $H_0:k=1$	p - arvo
Nisu	-63,5	234,1	1,009	0,048	576	0,195	0,846
Vakka -71	74,1	132,5	0,910	0,028	576	-3,152	0,002
Aura	-83,2	124,2	1,054	0,026	576	2,032	0,042
Pitko	146,5	154,1	0,967	0,033	576	-0,997	0,319
Jo 3077	742,9	746,3	0,926	0,148	576	-0,501	0,617
Jo 3079	-118,8	352,3	1,026	0,074	576	0,348	0,728
Linna P	290,0	187,8	0,958	0,040	576	-1,043	0,297
Folke	-882,0	260,6	1,272	0,060	576	4,556	0,000

le herkät lajikkeet tunnistaa yksinkertaisen tilastollisten testien avulla. Esimerkiksi valitsemalla aikaisempaan tapaan koko testaus tilanteen kriittiseksi erehtymistodennäköisyydeksi 5 % (experimentwise error rate) tulee yksittäisten testien kriittiseksi arvoksi $0,05/8 = 0,0063$. Tämän mukaisesti syysvehnälaajikkeeseen Vakka -71 (101851) herkkyyskerroin ($p = 0,002$) on myös tilastollisesti ykköstä pienempi, ts. se ei käytä kasvuympäristönsä mahdollisuuksia hyväkseen yhtä hyvin kuin muut koesarjan lajikkeet. Koska tämä sama lajike on jo edellä olevassa vertailussa (Taulukko 1) todettu myös keskisadoltaan mittarilajiketta Aura heikommaksi, ei sitä kannata suosittelaa ainakaan hyvin kasvuolosuhteisiin. Syysvehnälaajikkeeseen Folke herkkyyskerroin ($p=0,000$) taas on selkeästi ykköstä suurempi. Folke on myös keskisadoltaan suurempi kuin mittarilajike Aura. Siksi se näyttäisi olevan erityisen sopiva juuri hyvin kasvuolosuhteisiin. Taulukosta 2 voidaan myös nähdä, että esimerkkiaineiston syysvehnälaajikkeissa ei ole yhtään lajiketta, joka olisi tasaisen varma sekä hyvissä että huonoissa viljelyolosuhteissa, ts. lajiketta, jonka lajikevaikutusparametri olisi positiivinen ja herkkyyskerroin ykköstä suurempi.

Esimerkkiaineisto 2: ohra

Usein virallisissa lajikevertailuissa käytetään vain kahdeksan viimeisimmän vuoden aineistoa. Seuraavassa tarkastellaan tämän käytännön mukaisesti valittua osa-aineistoa ohran lajikekoeaineistosta vuosilta 1987–1994. Mukana on kymmenen lajiketta kuudeltatoista eri koepaikalta. Tämän osa-aineiston analysointi voidaan aloittaa lineaarista sekamalla soveltaen käyttämällä seuraavaa jo syysvehnäaineiston analysoinnista tuttua SAS MIXED -proseduurin ajomääritystä:

```
PROC MIXED;
CLASS lajike vuosi paikka;
MODEL sato = lajike;
```

```
RANDOM vuosi paikka vuosi*paikka
lajike*vuosi lajike*paikka;
RUN;
```

Liitteellä 7 on tätä SAS -proseduuria vastaava ajotulos. Tästä tulostuksesta voidaan päätellä mm seuraavaa:

- Ohralajikkeiden satoisuudessa on eroja, sillä lajikkeiden satoeroja testaava F-testi on merkitsevä ($p = 0,0000$). Samoin kuin syysvehnälaajikkeiden vertailussa tämä tulos on odotusten mukainen ja antaa perustan lajike-erojen tarkemmalle selvittämislle.
- Ohralajikkeiden satoisuuksien vaihtelua ei voida selittää pelkästään koepaikkojen eroilla. Tämä näkyy siitä, että koepaikatekijään liittyvä varianssikomponentti on tilastollisesti merkityksetön ($p = 0,0999$). Vastaava testitulos saatiin myös syysvehnäaineistolle.
- Ohralajikkeiden satotulosten vaihtelua ei voida selittää sääolojen erilaisuudellaan. Tämä näkyy siitä, että lajike x vuosi -yhdysoikutustekijään liittyvä varianssikomponentti ($p = 0,2121$) sekä vuositekijään liittyvä varianssikomponentti ($p = 0,1003$) ovat tilastollisesti merkityksettömiä. Vastaava testitulos saatiin syysvehnäaineistolle.
- Ohralajikkeiden satotulosten erot vaihtelevat eri koepaikoilla, sillä lajike x koepaikka -tekijään liittyvä varianssikomponentti on tilastollisesti merkitsevä ($p = 0,0310$). Tämän tuloksen mukaan lajikkeet ovat herkkiä koepaikkojen eroille ts. jotkut lajikkeet menestyvät tietyillä koepaikoilla toisia paremmin. Tämä tulos antaa perustan koepaikka-kohtaisten satoennusteiden tekemiselle. Syysvehnäaineistossa lajikkeiden välisen erojen ei voitu väittää riippuvan koepaikoista.
- Suurin ohralajikkeiden vaihtelun lähde on vuosi x koepaikka -tekijällä ($p = 0,0000$). Tämä viittaa siihen, että samantapaisetkin sääolot yhdistettynä kunkin koepaikan kasvutekijöihin muodostavat

erilaiset kasvuolosuhteet. Vastaava testi-
tulos saatiin myös syysvehnäaineistolle.

- Syysvehnäajikkeista poiketen ohralajikkeiden satotulosten erot eivät ole herkkiä säätekijöille, ts. lajikkeiden väliset erot eivät vaihtelee vuosittain. Tämä näkyy siitä, että lajike x vuosi -yhdysvaikutukseen liittyvä varianssikomponentti ei ole tilastollisesti nollasta eroava ($p = 0,2121$).

Koska tässä esimerkissä vuosi-, koepaikka- ja lajike x vuosi -tekijöillä ei varianssikomponentteihin liittyvien testien mukaan näyttänyt olevan merkitystä, voidaan mallia tarkentaa jättämällä nämä turhiksi todetut tekijät pois. Kun samalla lisätään malliin uusia mahdollisesti vaihtelun lähteinä toimivia koetekijään liittyviä yhdysvaikutustekijöitä päädytään seuraavaan kokeiltavaan malliin:

```
PROC MIXED;  
CLASS lajike vuosi paikka koe;  
MODEL sato = lajike;  
RANDOM lajike*paikka vuosi*paikka  
paikka*koe vuosi*paikka*koe;  
RUN;
```

Yksittäiset koevaikutukset ovat kuten aikaisemmin syysvehnäaineistoa analysoitaessa koepaikkojen sisäisiä satunnaisvaikutuksia. Siksi niiden itsenäisiä päävaikutuksia ei ole järkevää sisällyttää malliin. SAS MIXED -proseduurilla saadaan liitteen 8 mukainen tulostus. Tähän

tulokseen sisältyvien testiarvojen mukaan malli on yhteensopiva tarkasteltavana olevan ohra-aineiston kanssa, joten sitä voidaan käyttää jatkoanalyysihin.

Ohralajikkeiden satotulosten ennustaminen

Edellä määriteltyä mallia voidaan käyttää ohralajikkeiden satotulosten ennustamiseen ja lajikkeiden satoisuuden vertaamiseen. Edellä todettiin lajikkeiden satotulosten erojen riippuvan myös koepaikoista, mistä johtuen on perusteltua tehdä analyysi myös koepaikkakohtaisesti. Ennustamisessa käytetään mallin avulla estimoituja kiinteitä lajikevaikutuksia ja estimoituja koepaikka x lajike -tekijään liittyviä satunnaisvaikutuksia. Käytännössä ennusteita voidaan tehdä SAS MIXED -proseduuriin liittyvien LSMEANS- ja ESTIMATE -komentojen avulla. Taulukossa 3 on esitetty satotulosten havaitut ja ennustetut keskiarvot lajikkeittain eri koepaikoille. Tilan säästämiseksi tähän taulukkoon on valittu edellä esitetystä ohra-aineistosta vain kymmenen lajiketta ja kuusi koepaikkaa.

Taulukko 3. Ohran koepaikkakohtaiset satoennusteet (t/ha).

Lajike	Koepaikka						Yht.
	Jokioinen	Mietoinen	Häme	Etelä-Savo	Pohj.-Savo	Etelä-Pohj.	
Arra	4,22	4,35	4,44	4,90	4,34	4,33	4,35
	4,34	4,40	4,31	4,46	4,34	4,39	4,17
	0,250	0,251	0,252	0,253	0,253	0,249	0,247
	33	19	11	8	8	42	121
Arturi	4,85	4,94	5,57	5,27	4,96	3,70	4,66
	4,35	4,40	4,36	4,39	4,38	4,37	4,54
	0,252	0,252	0,254	0,253	0,253	0,251	0,251
	14	12	4	6	6	16	58
Hja Pokko	4,55	4,41	4,63	3,99	3,18	4,43	4,45
	4,38	4,45	4,35	4,32	4,33	4,41	4,21
	0,250	0,251	0,252	0,254	0,255	0,250	0,248
	31	16	11	4	1	34	97
Hja Eero	5,20	4,20	5,04	4,13	5,22	4,91	4,83
	4,39	4,37	4,37	4,35	4,42	4,35	4,42
	0,252	0,252	0,253	0,254	0,254	0,250	0,251
	9	11	8	3	4	31	66
Pohto	4,93	4,64	4,92	4,97	5,03	4,94	4,89
	4,35	4,41	4,33	4,38	4,38	4,39	4,59
	0,251	0,251	0,252	0,253	0,253	0,249	0,248
	17	17	11	8	7	39	99
Loviisa	5,17	4,64	5,08	4,89	5,21	5,06	4,98
	4,35	4,38	4,36	4,39	4,38	4,37	4,63
	0,252	0,251	0,252	0,253	0,254	0,250	0,249
	12	17	10	7	3	33	82
Kustaa	4,54	3,88	4,92	4,51	3,35	2,72	4,02
	4,40	4,28	4,48	4,38	4,36	4,34	4,04
	0,250	0,251	0,252	0,253	0,255	0,250	0,248
	38	21	11	7	1	23	101
Kymppi	4,76	3,95	4,90	4,82	4,30	4,12	4,40
	4,42	4,28	4,43	4,43	4,38	4,30	4,22
	0,250	0,251	0,252	0,253	0,255	0,249	0,247
	38	22	11	8	1	43	123
Kalle	4,24	4,25	4,64	4,23	4,84	4,23	4,29
	4,40	4,41	4,36	4,33	4,38	4,37	4,19
	0,250	0,252	0,252	0,253	0,255	0,250	0,249
	26	13	11	6	1	33	90
Arve	5,29	4,58	5,49	4,78	5,11	5,23	5,11
	4,36	4,37	4,39	4,31	4,38	4,43	4,72
	0,250	0,251	0,253	0,253	0,254	0,249	0,248
	21	16	8	8	5	37	95
Yht.	4,67	4,35	4,90	4,72	4,79	4,46	4,57
	4,37	4,37	4,37	4,37	4,37	4,37	4,37
	0,241	0,242	0,242	0,242	0,242	0,241	0,241
	239	164	96	65	37	331	932

Jokaista lajiketta kohti on ensimmäisellä rivillä kokeiden tuloksena saadut mitatut satokeskiarvot (t/ha), toisella rivillä on satotulosten ennustearvot, kolmannella rivillä on ennustearvojen estimoidut hajonnat ja neljännellä rivillä on analyysin perustana olleiden havaintojen lukumäärät.

Taulukko 4. Yksittäisten ohralajikkeiden vertaaminen tunnettuihin mittarilajikkeisiin.

Lajike	Satoero (t/ha)	Hajonta	Vapausasteet	t-arvo	p-arvo
Arturi	-0,299	0,136	1735	-2,20	0,0282
Hjan Pokko	0,176	0,126	1735	1,39	0,1639
Hjan Eero	-0,178	0,147	1735	-1,21	0,2270
Loviisa	-0,420	0,136	1735	-3,09	0,0020
Kymppi	-0,123	0,125	1735	-0,99	0,3227
Kalle	0,139	0,130	1735	1,07	0,2826
Arve	-0,536	0,122	1735	-4,38	0,0000

Ohralajikkeiden vertaaminen

Edellä on usein viitattu siihen, että testaus-tilanteissa kannattaa pyrkiä tekemään mahdollisimman vähän samanaikaisia testejä. Yksi tapa selviytyä suhteellisen pienellä yksittäisten testien määrällä on verrata kiinnostavia (uusia) lajikkeita vanhojen jo pitkään koeteltujen mittarilajikkeiden keskiarvoon. Tämän ajattelutavan mukaisesti on taulukossa 4 verrattu seitsemän vähemmän tunnetun ohralajikkeen satotasoa kolmen paremmin tunnetun mittarilajikkeen (Arra, Pohto ja Kustaa) keskimääräisiin satotasoihin.

Käyttämällä koko testaustilanteen kriittisenä tasona 5 % voidaan testattavina olevien yksittäisten lajikkeiden satojen todeta poikkeavan merkitsevästi vanhojen mittarilajikkeiden keskiarvosta silloin, kun taulukossa oleva p - arvo on pienempi kuin Bonferronin testin mukaisesti laskettu uusi kriittinen arvo $0,05/7 = 0,007$. Tällaisen testauksen perusteella lajikkeet Loviisa ja Arve näyttäisivät olevan keskimääräisiltä satotasoiltaan vanhoja mittarilajikkeita huonompia. Tämän testin perusteella yksittäisistä lajikkeista minkään ei voida väittää olevan vanhoja mittarilajikkeita parempi.

Mallintamiseen liittyvät oletukset

Tilastollinen malli on matemaattinen kuvaus käsiteltävänä olevan havaintoaineiston rakenteesta ja sen muuttujien välisistä riippuvuussuhteista. Tilastollisissa analyyseissä erityisen paljon käytetty tavanomainen kiinteävaikutteinen lineaarinen malli voidaan matriisimuodossa kirjoittaa $y = X\alpha + \epsilon$, missä y on selitettävän vastemuuttujan havaitut arvot sisältävä havaintovektori, α on mallin kiinteät mutta tuntemattomat lineaariset parametrit sisältävä vektori, X on tätä vektoria vastaava mallimatriisi ja ϵ on toisistaan riippumattomista tuntemattomista satunnaisvirheistä muodostuva jäännösvektori. Tällaisen mallin avulla voidaan mallintaa vektorin y keskiarvon lineaarista riippuvuutta mallimatriisiin X sisältyvistä selittävästä muuttujista ja parametrivektoriin α sisältyvistä kiinteistä parametreista.

Kiinteävaikutteinen lineaarinen malli voidaan laajentaa sekamalliksi, jos oletetaan, että osa malliin sisältyvistä parametreista on satunnaismuuttujia. Edellä suoritetuissa esimerkkianalyyseissä tällainen laajennus tehtiin käyttämällä yllä olevan mallin asemesta mallia

$y = X\alpha + \sum Z_i \beta_i + \epsilon$, missä jokainen β_i on yhteen satunnaistekijään liittyvät satunnaisvaikutukset sisältävä vektori ja Z_i on tätä samaista satunnaistekijää vastaava mallimatriisi. Tällainen mallin laajennus antaa

mahdollisuuden mallintaa myös vektorin y hajontaa. Edellä olevissa esimerkkiratkaisuissa kaikki lajikevaikutukset oletettiin kiinteiksi. Sen sijaan kaikki vuosiin, koepaikkoihin ja kokeisiin liittyvät tekijät sekä kaikki mahdolliset yhdysvaikutustekijät oletettiin satunnaisvaikutteisiksi. Malliin liittyviä parametreja estimoidessa oletettiin myös, että kaikki satunnaisvektorit β_i ja jäännösvektori ϵ ovat toisistaan riippumattomia ja kuhunkin näistä vektoreista sisältyvät satunnaisvaikutukset ovat vektorikohtaisesti samaa normaalijakaumaa noudattavia, toisistaan riippumattomia satunnaismuuttujia.

Lajikekoeaineistojen mallintamisessa tarvittavien satunnaistekijöiden riippumattomuus toisistaan ei yleisesti ottaen liene epäilyksenalaista. Sen sijaan ei ole itsestään selvää, että satunnaistekijöihin tai jäännösvirheisiin liitetyt muut jakaumaominaisuudet olisivat aina voimassa. Siksi lajikekoesarjoja analysoivan tutkijan tulee aina erikseen varmistua siitä, että lineaaristen sekamallien käyttöön liittyvät jakaumaominaisuudet ovat mallinnettavan aineiston kanssa yhteensopivia. Jakaumaominaisuuksien selvittämiseen verrattavaa satunnaisvaikutusten tutkimusta on tehty jo edellä, kun poikkeavia kokeita tai muita yksittäisiä poikkeuksellisia satunnaisvaikutuksia on etsitty tilastollisten testien avulla. Yksittäiset poikkeamat (outliers) eivät tietenkään vielä sinänsä estä mallien käyttöä, mutta niiden tulisi kyllä aina johtaa poikkeamien syiden lähempään tarkasteluun. Toinen yksinkertainen tapa tutkia satunnaistekijöihin tai jäännösvirheisiin liittyviä jakaumaominaisuuksia on käyttää tilastollisessa analyysissä tavanomaisia graafisia keinoja. Usein tällä tavalla löytyy itse havaintoaineistoista virheitä tai aineiston osia, jotka jatkotarkasteluissa on helppo karsia pois tai muuten korjata.

Kuvassa 2 on graafisesti esitetty ohra-aineiston analysointiin liittyvää kolmea satunnaistekijää (β_i) ja jäännösvektoria ϵ vastaavat käytetyn mallin mukaisesti estimoidut empiiriset kertymäfunktiot. Kunkin kertymäfunktion vierelle on koottu

myös kyseisiä satunnaisvaikutuksia kuvaavia SAS UNIVARIATE -proseduurin avulla laskettuja tavanomaisia tunnuslukuja. Silmämääräisesti katsoen minkään piirretystä neljästä kertymäfunktiosta ei voida väittää poikkeavan tutusta normaalijakauman kertymäfunktiosta. Myös lasketut tunnusluvut ovat pääosin normaalijakauman kanssa yhteensopivia. Tämä näennäinen yhteensopivuus ei kuitenkaan ole matemaattisen tarkkaa, sillä piirroksiin liittyvien Shapiro-Wilk -testisuureiden mukaisesti voidaan perustellusti väittää, että kolme neljästä satunnaisvaikutusjoukosta ei tarkkaan ottaen ole peräisin normaalijakaumasta. Vain lajike x paikka -tekijän ei voida väittää poikkeavan normaalisuusoletuksista.

Graafisen analyysin tarkentamiseksi voidaan siihen lisäksi liittää puhdasoppisempaa tilastollista päättelyä. Eräs usein käytetty tapa perustuu siihen, että otokseen kasvaessa empiirisen kertymäfunktion kvantiilit lähestyvät aina asymptoottisesti kyseessä olevan satunnaismuuttujan todellisen kertymäfunktion kvantileja (esim. Rao 1973, 420–425). Siksi hyvässä yhteensopivuustilanteessa empiiristen ja oletettujen eksaktien kertymäfunktioiden vastinpisteet osuvat graafisessa esityksessä samalle origon kautta kulkevalle suoralle (esim. Cook & Weisberg 1986, 53–55). Tämän keinon havainnollistamiseksi on kuvassa 3 esitetty edellä olevien satunnaisvaikutusten empiiristen kertymäfunktioiden ja näitä vastaavien eksaktien normaalijakaumien kertymäfunktioiden lineaarinen riippuvuus regressiotekniikkaa hyväksi käyttäen (ns. Q-Q plot). Vertaamalla saatua ennustepisteiden joukkoa kuvassa näkyviin 95 % -virherajoihin voidaan tämänkin analyysin avulla havaita, että satunnaisvaikutusten jakaumien yhteensopivuus normaalijakauman kanssa ei ole täydellistä.

Edellä havaitut poikkeamat satunnaisvaikutusten normaalisuusoletuksista eivät sellaisinaan ole kovin merkityksellisiä. Vähäiset poikkeamat ovat odotettavissa, koska suurehkoissa havaintoaineistoissa on

lähes aina mukana muutamia virheellisiä mittaustuloksia tai muunlaisia virheellisiä kirjauksia. Tämä näkyy yleensä sekä tilastollisissa testeissä että graafisissa esityksissä. Onneksi tilastollisen tarkastelun perustaksi riittää käytettyjen mallien ja oletusten likimääräinen voimassolo ja se, että yksittäiset mittavirheet eivät ole kovin suuria. Esimerkkitapauksessammekin havaitut poikkeamat normaalisuudesta ovat tilastollisen analyysin kannalta todennäköisesti merkityksettömiä. Tehtyjen graafisten tarkastelujen valossa ohra-aineiston rakenteessa on kuitenkin niin paljon selittämättömäksi jäävää vaihtelua, että se todellisessa aineiston analysointitilanteessa kannattaisi ottaa jottenkin huomioon. Tässä tapauksessa voitaisiin esim. karsia poikkeaviksi todettuja kokeita pois jatkoanalyyseistä ja/tai jakaa aineisto viljelyominaisuuksiltaan koko maata yhtenäisempiin osa-aineistoihin ja näitä vastaaviin osa-analyyseihin.

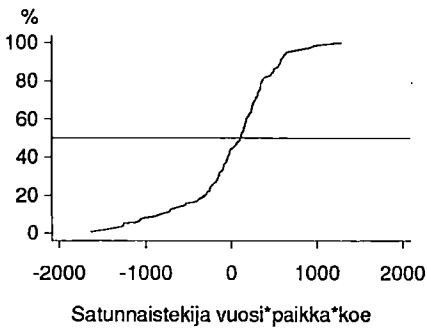
Edellä olevissa esimerkeissä on sovellettu viime vuosikymmeninä jokseenkin vähän käytettyä Bonferronin testiä. Aikaisempi vähäinen käyttö johtunee pääasiassa siitä, että tarvittavia taulukoita Studentin t-arvojen prosenttipisteiden määrittämiseksi ei ole helposti saatavilla. Tosin ainakin yksi vaatimaton taulukko on tarjolla (Miller 1981). Testin vaivaton käyttö on nykyisin kuitenkin mahdollista, koska prosenttipisteet on tarvittaessa helppo laskea valmiilla sovellusohjelmistoilla. Näin tehtiin edellä IML-ohjelmaa käyttämällä. Bonferronin testi on monessa mielessä sopiva lajikevertailujen tekemiseen, sillä jokaisessa testausstilanteessa se antaa käyttäjälleen mahdollisuuden valita samanaikaisesti tehtävien yksittäisten testien määrän vielä varsinaisten tietokoneajojen jälkeen. Jokaisessa käyttötilanteessa se myös näyttää käyttäjälleen, mitä testin kokonaisvoimakkuudessa menetetään tai saavutetaan, jos yksittäisten osatestien määrää lisätään tai vähennetään. Lisäksi Bonferronin testi voidaan mittateorian avulla perustella varsin lyhyesti, mistä johtuen myös sen eri ominaisuudet ovat helposti arvioitavissa. Perustelu on seuraava: jokaiselle tapahtu-

majoukolle A_i , $i = 1, \dots, k$ on voimassa $P(\cap A_i) = 1 - P(\cup \bar{A}_i) = 1 - P(\cup -A_i) \geq 1 - \sum P(-A_i)$; tulkitsemalla jokainen A_i oikein osuneeksi yksittäisen nollahypoteesin hylkäämistapahtumaksi ja valitsemalla jokaisen yksittäisen nollahypoteesin hylkäämiseen johtavan testin kriittiseksi tasoksi α/k saadaan tällöin edellisen mukaisesti koko testaus tapahtuman kriittiselle tasolle alaraja $P(\cap A_i) \geq 1 - \sum P(-A_i) = 1 - k\alpha/k = 1 - \alpha$. Tämän mittateoreettisen perustelun avulla nähdään myös mm. se tärkeä Bonferronin testin perusominaisuus, että sen avulla voidaan yhdistää mitä tahansa erillisiä testejä eikä vain esim. edellä käytettyjä t-testejä.

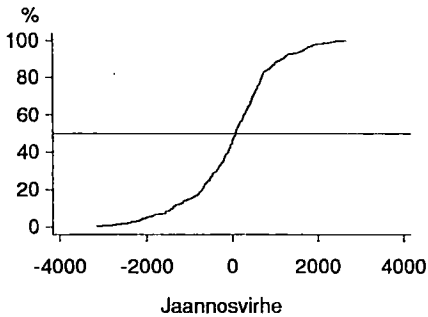
Lienee syytä todeta, että kaikki SAS MIXED -proseduurin käyttämät testit ovat asymptoottisesti tarkentuvia mutta tosiasiallisesti kuitenkin vain likimääräisiä testejä. Epätarkkuus liittyy erityisesti varianssikomponentteja koskeviin testeihin siten, että testausvirheen mahdollisuus on erityisen suuri, kun käytettävät aineistot ovat pieniä ja/tai varianssikomponentit ovat merkitykseltään vähäisiä. Lajikekoeaineistojen analysoinnissa likimääräisten testien käytöstä ei kuitenkaan liene suurta haittaa. Lajikekoeaineistot ovat yleensä niin suuria, että testisuureet noudattavat testeissä käytettyjä asymptoottisia jakaumia. Toisaalta edellä olevien malliesimerkkien mukaisessa lajikekokeiden analysoinnissa varianssikomponentteihin liittyviä testejä tarvitaan vain ja ainoastaan käyttökelpoisten mallien etsintään. Tässä etsinnässä vähämerkityksellisten varianssikomponenttien tunnistamisessa tehty virheet eivät lopputuloksen kannalta ole kovin tärkeitä. Varianssikomponentteja koskevia tilastollisesti tarkkoja testejäkin on toki olemassa (Öfversten 1993), mutta edellä mainituista syistä niiden käyttöön tuskin kannattaa ryhtyä.



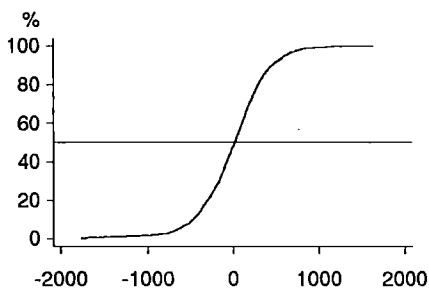
Satunnaistekijä:	lajike*paikka
Normaalisuustesti:	<i>p</i> -arvo 0.5389
Havaintojen lkm:	150
Keskiarvo:	0.9047
Hajonta:	112.51
Maksimi-arvo:	279.70
Yläkvartiili	70.00
Mediaani	8.57
Alakvartiili	-62.09
Minimi-arvo	-367.44



Satunnaistekijä:	vuosi*paikka
Normaalisuustesti:	<i>p</i> -arvo 0.0001
Havaintojen lkm:	121
Keskiarvo:	0.0000
Hajonta:	540.90
Maksimi-arvo:	1278.52
Yläkvartiili:	335.36
Mediaani:	113.78
Alakvartiili:	-220.04
Minimi-arvo:	-1629.95

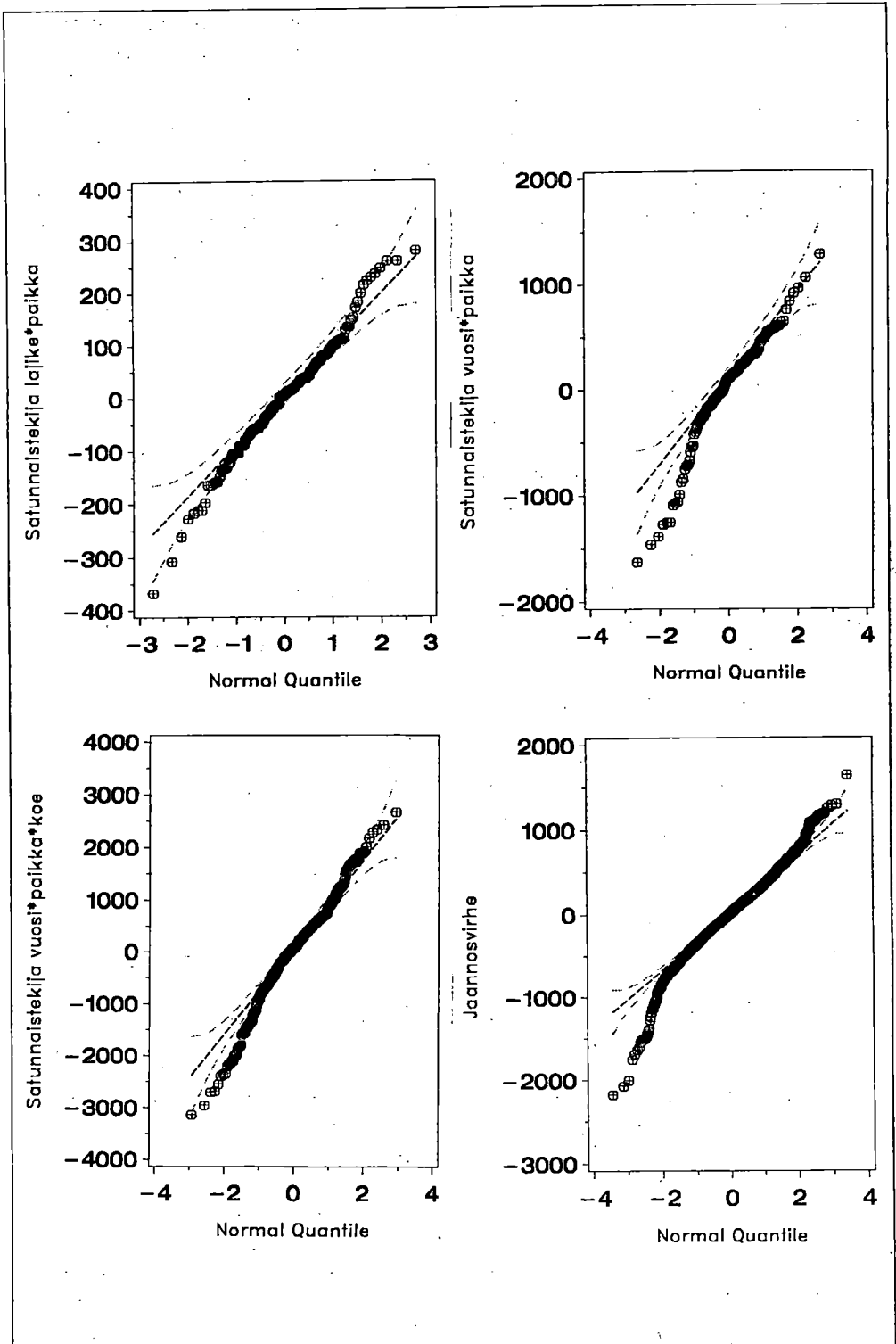


Satunnaistekijä:	vuosi*paikka*koe
Normaalisuustesti:	<i>p</i> -arvo 0.0037
Havaintojen lkm:	297
Keskiarvo:	0.0000
Hajonta:	1011.03
Maksimi-arvo:	2662.06
Yläkvartiili:	600.78
Mediaani:	81.36
Alakvartiili	-527.61
Minimi-arvo:	-3134.64



Normaalisuustesti:	Jäännösvirhe
Havaintojen lkm:	<i>p</i> -arvo 0.0001
Keskiarvo:	1745
Hajonta:	0.0000
Maksimi-arvo:	402.03
Yläkvartiili:	1635.02
Mediaani:	240.83
Alakvartiili	17.10
Minimi-arvo:	-227.58
	-2173.70

Kuva 2. Ohra-aineistoon liittyvien satunnaisvaikutusten empiriset kertymäfunktiot.



Kuva 3. Ohra-aineistoon liittyviä satunnaisvaikutuksia kuvaavat Q – Q -plotit.

Yhteenveto

Parivertailun puutteet ovat jo pitkään olleet tiedossa. Ensimmäiset yritykset epätäydellisten lajikekoeaineistojen analysoimiseksi tehtiin 1970-luvulla (vrt. Patterson & Silvey 1980). Teknisesti tällöin oli mahdollista käyttää vain kiinteävaikutteisia malleja. Johtopäätösten yleistettävyyden tai tarkan kohdistamisen kannalta lineaaristen sekamallien käyttö on kuitenkin paljon joustavampaa. Tarvittaessa näiden mallien avulla tehtävät johtopäätökset voidaan ulottaa koskemaan koko viljelyaluetta tai vain yhtä tai muutamaa koepaikkaa. Lisäksi edellä on osoitettu, että mallien antamia tuloksia voidaan käyttää myös poikkeuksellisesti käytettyjen kokeiden tunnistamiseen, poikkeuksellisten koepaikkojen tunnistamiseen ja lajikkeiden herkkyyserojen vertaamiseen. Lajikkeiden samanaikaisella vertailulla voidaan myös vähentää koetoinnin kustannuksia, sillä lineaaristen sekamallien avulla tehtävä vertailu ei edellytä kaikkia vertailtavina olevia lajikkeita viljelyiksi kaikissa vertailuun käytettävissä kokeissa. Jo koeteltuja vanhoja lajikkeita ei tarvitse sijoittaa kokeisiin vain siksi, että saataisiin kokoon perinteisen parivertailun edellyttämiä yleensä melko suuria täydellisiä osa-aineistoja.

Kiinteävaikutteisten lineaaristen sekamallien avulla mallinnetaan vastemuuttujavektoriin liittyvien luokkakeskisarvojen riippuvuutta selittävästä muuttujasta ja tekijöistä. Lineaarisia sekamalleja käyttämällä voidaan keskisarvojen lisäksi mallintaa myös vastemuuttujavektorin vaihtelua (hajontaa). Juuri hajonnan mallintamisen avulla testitulokset ja ennusteet voidaan laajentaa varsinaisten koepaikkojen lisäksi myös koepaikkojen välisiä viljelyalueita koskeviksi. Parivertailua tai kiinteävaikutteisia malleja käyttämällä ei saada tällä tavalla laajennuskelpoisia tuloksia. Hajonnan mallintaminen edellyttää kuitenkin sitä, että koepaikat ja koepaikkojen sisäiset kasvupaikat muodostavat edustavan otoksen koko viljelyalueen koe- ja kasvupaikoista.

Todellisuudessa koepaikkojen valintaa ei useinkaan ole tehty tilastollisessa mielessä satunnaisesti. Käsitellyissä esimerkkitapauksissa tämä näkyy mm. koepaikkojen välisenä vähäisenä vaihteluna. On myös huomattava, että kaikki koepaikat sisältävä aineisto ei aina sellaisenaan välttämättä edusta kasvuolosuhteiltaan yhtenäistä viljelyaluetta. Edellä ohra-aineisto oli esimerkiksi maantieteellisesti laajalta alueelta kerätystä mutta kuitenkin melkoisen yhtenäisestä tutkimusaineistosta. Toisaalta syysvehnäaineiston analyysi osoitti, että alueellisesti pienikin viljelyalue saattaa sisältää paljon poikkeavia yksittäisiä koepaikka- tai koepaikka x vuosi -vaikutuksia. Koepaikkatekijöihin liittyvät poikkeukselliset satunnaisvaikutukset tai oikukkaat jakaumat tuovat aina esille koepaikkojen valintaan liittyviä puutteita. Tällaiset puutteet hankaloittavat koesarjojen analysointia. Toiminnallisesti suurin haitta seuraa kuitenkin siitä, että alkuperäistä viljelyaluetta huonosti edustava otos aiheuttaa aina harhaa ja epävarmuutta johtopäätöksiin analysointitavasta riippumatta. Siksi tulevassa lajikekoetoinnassa koepaikkojen valintaan kannattaa kiinnittää nykyistä enemmän huomiota.

Lopuksi haluamme vielä huomauttaa, että edellä on esitetty vain esimerkinomaisia näytteitä siitä, mitä lajikekoeaineistojen analysointi lineaarisia sekamalleja käyttäen saattaa pitää sisällään. Siksi esitettyä tarkastelua ei tule pitää suoraan seurattavaksi tarkoitettuna ohjeena tai suosituksena, vaan todellakin vain johdatuksena uudentyyppiseen ajattelutapaan. Lineaarisia sekamalleja tosi tarkoituksella käyttävän tutkijan on jokaisessa eteen tulevassa uudessa koesarjojen analysointitilanteessa aina uudelleen nähtävä vaivaa käyttökelpoisten lineaaristen sekamallien ja poikkeavien kokeiden tunnistamiseksi, välttämättömien testaushypoteesien muodostamiseksi, tarkoituksenmukaisen estimointi- ja ennustustarpeiden määrittämiseksi ja lopulta vielä käyttämiensä oletusten todenperäisyyden varmistamiseksi.

Kirjallisuus

Cochran, W.G. & Cox, G.M. 1950. *Experimental Design, Second Edition*. New York: John Wiley & Sons. 611 p.

Cook, R.D. & Weisberg, S. 1986. *Residuals and Influence in Regression*. New York: Chapman and Hall. 230 p.

Finlay, K.W. & Wilkinson, G.N. 1963. The analysis of adaptation in a plant breeding programme. *Australian Journal of Agricultural Research* 14: 742–754.

Järvi, A., Kangas, A., Mustonen, L., Salo, Y., Talvitie, H., Vuorinen, M. & Mäkelä, L. 1995. Virallisten lajikekokeiden tuloksia 1987–1994. Jokioinen: Maatalouden tutkimuskeskus. Tiedote 2/95. 126 p. (ISSN 0359-7652)

Mattila, I., Mäkelä, L. & Nikander, H. 1995. Ruutusalaskennan ohjeet. Jokioinen: Maatalouden tutkimuskeskus. 15 p.

— & **Vähämäki, S.** 1994. Ruutukohtaisista tiedoista koetuloslomakkeiksi. Jokioinen: Maatalouden tutkimuskeskus. 18 p.

Miller, R.G., Jr. 1981. *Simultaneous Statistical Inference*. New York: Springer-Verlag. 299 p.

Patterson, H.D. & Silvey, V. 1980. *Statutory and Recommended List of Crop varieties in the United*

Kingdom. *Journal of the Royal Statistical Society, Series A* 143: 219–252.

— & **Nabugoomu, F.** 1992. REML and the analysis of crop variety trials. In: *Proceedings of Invited Papers, 1992 International Biometric Conference*, Hamilton, New Zealand.

Rao, C.R. 1973. *Linear Statistical Inference and Its Applications, Second Edition*. New York: John Wiley & Sons. 625 p.

SAS Institute. 1992. *SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07*. Cary, NC: SAS Institute. 620 p.

Talbot, M. 1984. Yield variability of crop varieties in the U.K. *Journal of Agricultural Science* 102: 315–321.

Yates, F. & Cochran, W.G. 1938. The analysis of groups of experiments. *Journal of Agricultural Science* 28: 556–580.

Öfversten, J. 1993. Exact Tests for Variance Components in Unbalanced Mixed Linear Models. *Biometrics* 49: 45–57.

— 1995. Estimation in mixed models via layer triangular transformation. *Computational Statistics and Data Analysis* 20: 657–667.


```

LIBNAME vehna '<lst1.lajike>';
* OPTIONS LS=108 PS=42;
PROC MIXED DATA=vehna.svehna2;
CLASS lajike vuosi paikka;
MODEL sato=lajike;
RANDOM vuosi paikka vuosi*paikka lajike*vuosi lajike*paikka;
RUN;

```

The MIXED Procedure

Class Level Information

Class	Levels	Values
LAJIKE	8	100835 101851 103045 103063 103077 103079 222585 523977
VUOSI	14	81 82 83 84 85 86 87 88 89 90 91 92 93 94
PAIKKA	8	3 13 14 15 16 19 21 25

The MIXED Procedure

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	9007.19	
1	4	8374.86	
2	2	8371.38	0.0002
3	1	8370.25	0.0000
4	1	8370.16	0.0000
5	1	8370.16	0.0000

Convergence criteria met.

The MIXED Procedure

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > Z
VUOSI	0.6465	236051.99	186318.69	1.27	0.2052
PAIKKA	1.2096	441618.59	312947.52	1.41	0.1582
VUOSI*PAIKKA	2.7126	990315.61	202326.05	4.89	0.0000
LAJIKE*VUOSI	0.1110	40527.60	20013.59	2.03	0.0429
LAJIKE*PAIKKA	0.0316	11568.09	11725.00	0.99	0.3238
Residual	1.0000	365070.25	24903.36	14.66	0.0000

The MIXED Procedure

Model Fitting Information for SATO

Description	Value
Observations	592
Variance Estimate	365070.3
Standard Deviation Estimate	604.2104
REML Log Likelihood	4721.74
Akaike's Information Criterion	4727.74
Schwarz's Bayesian Criterion	4740.85
2 REML Log Likelihood	9443.488

The MIXED Procedure

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	7	584	4.36	0.0001

```

LIBNAME vehna '<lst1.lajike>';
* OPTIONS LS=108 PS=42;
PROC MIXED DATA=vehna.svehna2;
CLASS lajike vuosi paikka koe;
MODEL sato=lajike;
RANDOM lajike*vuosi vuosi*paikka paikka*koe vuosi*paikka*koe;
RUN;

```

The MIXED Procedure

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > Z
LAJIKE*VUOSI	0.1522	47968.54	21034.43	2.28	0.0226
VUOSI*PAIKKA	3.9525	1245233.11	255363.26	4.88	0.0000
PAIKKA*KOE	0.3998	125972.02	113823.99	1.11	0.2684
VUOSI*PAIKKA*KOE	0.1943	61231.92	25849.18	2.37	0.0178
Residual	1.0000	315043.61	22927.53	13.74	0.0000

The MIXED Procedure

Model Fitting Information for SATO

Description	Value
Observations	592
Variance Estimate	315043.6
Standard Deviation Estimate	561.2875
REML Log Likelihood	-4713.41
Akaike's Information Criterion	-4718.41
Schwarz's Bayesian Criterion	-4729.34
2 REML Log Likelihood	9426.821

The MIXED Procedure

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	7	584	5.05	0.0000


```

LIBNAME vehna '<lst1.lajike>';
* OPTIONS LS=108 PS=42;
PROC MIXED DATA=vehna.svehna2;
CLASS lajike vuosi paikka koe;
MODEL sato=lajike;
RANDOM lajike*vuosi vuosi*paikka vuosi*paikka*koe;
RUN;

```

The MIXED Procedure

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > Z
LAJIKE*VUOSI	0.1446	45748.28	20434.86	2.24	0.0252
VUOSI*PAIKKA	4.7200	1492482.22	268052.20	5.57	0.0000
VUOSI*PAIKKA*KOE	0.2932	92736.94	29944.96	3.10	0.0020
Residual	1.0000	316199.36	23040.83	13.72	0.0000

The MIXED Procedure

Model Fitting Information for SATO

Description	Value
Observations	592
Variance Estimate	316199.4
Standard Deviation Estimate	562.3161
REML Log Likelihood	-4717.36
Akaike's Information Criterion	-4721.36
Schwarz's Bayesian Criterion	-4730.10
2 REML Log Likelihood	9434.711

The MIXED Procedure

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	7	584	5.09	0.0000

The MIXED Procedure

Solution for Fixed Effects

Parameter	Estimate	Std Error	DDF	T	Pr > T
LAJIKE 100835	4434.67	192.90	584	22.99	0.0000
LAJIKE 101851	4046.67	163.78	584	24.71	0.0000
LAJIKE 103045	4564.82	161.22	584	28.31	0.0000
LAJIKE 103063	4398.28	168.03	584	26.17	0.0000
LAJIKE 103077	4855.48	251.93	584	19.27	0.0000
LAJIKE 103079	4395.50	218.44	584	20.12	0.0000
LAJIKE 222585	4579.31	179.56	584	25.50	0.0000
LAJIKE 523977	4614.72	197.89	584	23.32	0.0000

The MIXED Procedure

Solution for Random Effects

Parameter	Estimate	Std Error	DDF	T	Pr > T
LAJIKE*VUOSI 100835 81	40.96	178.72	584	0.23	0.8188
LAJIKE*VUOSI 100835 82	113.18	160.95	584	0.70	0.4822
LAJIKE*VUOSI 100835 83	-228.33	162.08	584	-1.41	0.1594
LAJIKE*VUOSI 100835 84	-9.01	182.00	584	-0.05	0.9605
LAJIKE*VUOSI 100835 85	22.87	202.45	584	0.11	0.9101
LAJIKE*VUOSI 100835 86	80.56	193.50	584	0.42	0.6773
LAJIKE*VUOSI 100835 87	-5.15	202.45	584	-0.03	0.9797
LAJIKE*VUOSI 100835 88	50.73	202.46	584	0.25	0.8022
LAJIKE*VUOSI 100835 89	-22.70	202.76	584	-0.11	0.9109
LAJIKE*VUOSI 100835 90	17.57	202.77	584	0.09	0.9310
LAJIKE*VUOSI 100835 91	-17.64	202.81	584	-0.09	0.9307
LAJIKE*VUOSI 100835 93	-43.04	202.89	584	-0.21	0.8320
LAJIKE*VUOSI 101851 81	52.24	172.46	584	0.30	0.7620
LAJIKE*VUOSI 101851 82	-138.69	151.32	584	-0.92	0.3598
LAJIKE*VUOSI 101851 83	104.34	151.29	584	0.69	0.4907
LAJIKE*VUOSI 101851 84	40.69	154.18	584	0.26	0.7919
LAJIKE*VUOSI 101851 85	161.66	152.94	584	1.06	0.2910
LAJIKE*VUOSI 101851 86	-103.69	156.01	584	-0.66	0.5065
LAJIKE*VUOSI 101851 87	-10.13	177.13	584	-0.06	0.9544
LAJIKE*VUOSI 101851 88	-5.30	156.37	584	-0.03	0.9730
LAJIKE*VUOSI 101851 89	-78.05	158.11	584	-0.49	0.6217
LAJIKE*VUOSI 101851 90	102.58	161.43	584	0.64	0.5254
LAJIKE*VUOSI 101851 91	-101.92	164.84	584	-0.62	0.5366
LAJIKE*VUOSI 101851 92	3.32	84.31	584	0.02	0.9856
LAJIKE*VUOSI 101851 93	-133.91	195.20	584	-0.69	0.4930
LAJIKE*VUOSI 101851 94	106.87	196.05	584	0.55	0.5859
LAJIKE*VUOSI 103045 81	-168.39	172.31	584	-0.98	0.3289
LAJIKE*VUOSI 103045 82	-71.94	150.93	584	-0.48	0.6338
LAJIKE*VUOSI 103045 83	161.53	150.89	584	1.07	0.2848
LAJIKE*VUOSI 103045 84	-75.11	153.81	584	-0.49	0.6255
LAJIKE*VUOSI 103045 85	158.12	152.48	584	1.04	0.3002
LAJIKE*VUOSI 103045 86	-44.11	155.57	584	-0.28	0.776
LAJIKE*VUOSI 103045 87	-58.89	176.97	584	-0.33	0.7394

Parameter	Estimate	Std Error	DDF	T	Pr > T
LAIJKE*VUOSI 103045 88	-181.56	155.87	584	-1.16	0.2446
LAIJKE*VUOSI 103045 89	80.20	157.65	584	0.51	0.6111
LAIJKE*VUOSI 103045 90	-4.02	159.61	584	-0.03	0.9799
LAIJKE*VUOSI 103045 91	176.15	160.36	584	1.10	0.2725
LAIJKE*VUOSI 103045 92	-28.85	169.72	584	-0.17	0.8651
LAIJKE*VUOSI 103045 93	16.13	172.70	584	0.09	0.9256
LAIJKE*VUOSI 103045 94	40.75	191.14	584	0.21	0.8312
LAIJKE*VUOSI 103063 81	73.19	193.40	584	0.38	0.7052
LAIJKE*VUOSI 103063 82	-63.34	175.21	584	-0.36	0.7178
LAIJKE*VUOSI 103063 83	-45.25	186.30	584	-0.24	0.8082
LAIJKE*VUOSI 103063 84	35.42	201.353	584	0.18	0.8604
LAIJKE*VUOSI 103063 85	-14.33	168.56	584	-0.09	0.9323
LAIJKE*VUOSI 103063 86	-89.29	168.626	584	-0.53	0.5966
LAIJKE*VUOSI 103063 87	66.41	177.67	584	0.37	0.7087
LAIJKE*VUOSI 103063 88	-99.35	160.48	584	-0.62	0.5361
LAIJKE*VUOSI 103063 89	131.63	159.99	584	0.82	0.4110
LAIJKE*VUOSI 103063 90	-67.87	161.49	584	-0.42	0.6744
LAIJKE*VUOSI 103063 91	34.39	166.53	584	0.21	0.8364
LAIJKE*VUOSI 103063 92	75.24	173.79	584	0.43	0.6652
LAIJKE*VUOSI 103063 93	72.55	179.97	584	0.40	0.6870
LAIJKE*VUOSI 103063 94	-109.41	191.36	584	-0.57	0.5677
LAIJKE*VUOSI 103077 81	-135.31	196.21	584	-0.69	0.4907
LAIJKE*VUOSI 103077 82	171.52	186.58	584	0.92	0.3583
LAIJKE*VUOSI 103077 83	26.63	196.16	584	0.14	0.8920
LAIJKE*VUOSI 103077 84	-5.00	203.43	584	-0.02	0.9804
LAIJKE*VUOSI 103077 85	-49.86	203.44	584	-0.25	0.8065
LAIJKE*VUOSI 103077 87	-7.98	203.21	584	-0.04	0.9687
LAIJKE*VUOSI 103079 81	-197.02	194.75	584	-1.01	0.3121
LAIJKE*VUOSI 103079 82	152.61	176.99	584	0.86	0.3889
LAIJKE*VUOSI 103079 83	-30.69	176.86	584	-0.17	0.8623
LAIJKE*VUOSI 103079 84	-63.35	176.79	584	-0.36	0.7202
LAIJKE*VUOSI 103079 85	70.60	203.01	584	0.35	0.7281
LAIJKE*VUOSI 103079 87	67.84	202.77	584	0.33	0.7380
LAIJKE*VUOSI 222585 81	256.26	174.12	584	1.47	0.1416
LAIJKE*VUOSI 222585 82	-134.42	156.65	584	-0.86	0.3912
LAIJKE*VUOSI 222585 83	-56.18	155.41	584	-0.36	0.7178
LAIJKE*VUOSI 222585 84	38.47	157.96	584	0.24	0.8077
LAIJKE*VUOSI 222585 85	18.68	160.23	584	0.12	0.9072
LAIJKE*VUOSI 222585 86	-168.72	176.65	584	-0.96	0.3399
LAIJKE*VUOSI 222585 87	142.13	202.31	584	0.70	0.4826
LAIJKE*VUOSI 222585 88	-110.36	202.32	584	-0.55	0.5856
LAIJKE*VUOSI 222585 89	-63.73	202.63	584	-0.31	0.7532
LAIJKE*VUOSI 222585 90	77.65	202.64	584	0.38	0.7017
LAIJKE*VUOSI 222585 91	32.32	202.68	584	0.16	0.8733
LAIJKE*VUOSI 222585 92	-53.97	203.28	584	-0.27	0.7907
LAIJKE*VUOSI 222585 93	33.46	202.76	584	0.17	0.8690
LAIJKE*VUOSI 222585 94	-11.58	203.60	584	-0.06	0.9546
LAIJKE*VUOSI 523977 83	38.96	194.22	584	0.20	0.8411
LAIJKE*VUOSI 523977 84	-26.30	173.00	584	-0.15	0.8792
LAIJKE*VUOSI 523977 85	-366.91	173.51	584	-2.11	0.0349
LAIJKE*VUOSI 523977 86	401.30	173.58	584	2.31	0.0211
LAIJKE*VUOSI 523977 87	-305.48	188.30	584	-1.62	0.1053
LAIJKE*VUOSI 523977 88	140.69	177.36	584	0.79	0.4279

Parameter	Estimate	Std Error	DDF	T	Pr > T
LAIKE*VUOSI 523977 89	76.19	202.32	584	0.38	0.7066
LAIKE*VUOSI 523977 93	41.53	197.40	584	0.21	0.8334
VUOSI*PAIKKA 81 3	340.07	263.54	584	1.29	0.1974
VUOSI*PAIKKA 81 16	-227.07	397.71	584	-0.57	0.5683
VUOSI*PAIKKA 81 19	-2859.65	407.59	584	-6.53	0.0000
VUOSI*PAIKKA 82 3	1505.33	254.05	584	5.93	0.0000
VUOSI*PAIKKA 82 13	565.40	272.05	584	2.08	0.0381
VUOSI*PAIKKA 82 14	2062.92	312.53	584	6.60	0.0000
VUOSI*PAIKKA 82 15	-701.98	330.04	584	-2.13	0.0338
VUOSI*PAIKKA 82 16	-496.46	407.21	584	-1.22	0.2233
VUOSI*PAIKKA 82 19	-1991.57	406.36	584	-4.90	0.0000
VUOSI*PAIKKA 83 3	958.57	255.59	584	3.75	0.0002
VUOSI*PAIKKA 83 13	186.63	273.26	584	0.68	0.4949
VUOSI*PAIKKA 83 14	747.42	323.00	584	2.31	0.0210
VUOSI*PAIKKA 83 15	-16.90	321.58	584	-0.05	0.9581
VUOSI*PAIKKA 83 16	349.17	407.18	584	0.86	0.3915
VUOSI*PAIKKA 83 19	-3170.61	406.24	584	-7.80	0.0000
VUOSI*PAIKKA 84 3	-242.71	259.42	584	-0.94	0.3499
VUOSI*PAIKKA 84 13	-1955.00	285.96	584	-6.84	0.0000
VUOSI*PAIKKA 84 14	923.19	396.01	584	2.33	0.0201
VUOSI*PAIKKA 84 15	1094.26	330.65	584	3.31	0.0010
VUOSI*PAIKKA 84 16	523.77	396.01	584	1.32	0.1865
VUOSI*PAIKKA 84 19	-2437.96	447.06	584	-5.45	0.0000
VUOSI*PAIKKA 85 3	472.62	259.03	584	1.82	0.0686
VUOSI*PAIKKA 85 13	-646.30	285.36	584	-2.26	0.0239
VUOSI*PAIKKA 85 14	626.60	324.49	584	1.93	0.0540
VUOSI*PAIKKA 85 15	801.39	330.27	584	2.43	0.0155
VUOSI*PAIKKA 85 16	-435.61	422.69	584	-1.03	0.3032
VUOSI*PAIKKA 85 19	-791.36	447.05	584	-1.77	0.0772
VUOSI*PAIKKA 86 3	1361.86	278.42	584	4.89	0.0000
VUOSI*PAIKKA 86 13	19.45	291.13	584	0.07	0.9468
VUOSI*PAIKKA 86 14	1346.22	335.68	584	4.01	0.0001
VUOSI*PAIKKA 86 15	498.13	342.05	584	1.46	0.1459
VUOSI*PAIKKA 86 16	-358.42	422.98	584	-0.85	0.3971
VUOSI*PAIKKA 86 19	-386.82	490.45	584	-0.79	0.4306
VUOSI*PAIKKA 87 14	-2158.86	325.71	584	-6.63	0.0000
VUOSI*PAIKKA 87 15	-1018.63	326.36	584	-3.12	0.0019
VUOSI*PAIKKA 87 16	-452.06	424.28	584	-1.07	0.2871
VUOSI*PAIKKA 88 3	-1041.79	265.61	584	-3.92	0.0001
VUOSI*PAIKKA 88 13	-96.23	294.60	584	-0.33	0.7440
VUOSI*PAIKKA 88 14	-637.45	324.76	584	-1.96	0.0501
VUOSI*PAIKKA 88 15	-2846.64	338.79	584	-8.40	0.0000
VUOSI*PAIKKA 88 16	163.74	447.84	584	0.37	0.7148
VUOSI*PAIKKA 88 19	-2234.53	447.84	584	-4.99	0.0000
VUOSI*PAIKKA 89 3	807.02	261.18	584	3.09	0.0021
VUOSI*PAIKKA 89 13	526.61	306.74	584	1.72	0.0865
VUOSI*PAIKKA 89 14	763.55	330.53	584	2.31	0.0212
VUOSI*PAIKKA 89 15	652.18	349.97	584	1.86	0.0629
VUOSI*PAIKKA 89 16	1267.24	448.47	584	2.83	0.0049
VUOSI*PAIKKA 89 19	13.66	448.47	584	0.03	0.9757
VUOSI*PAIKKA 90 3	-351.39	268.62	584	-1.31	0.1913
VUOSI*PAIKKA 90 13	1563.12	314.72	584	4.97	0.0000
VUOSI*PAIKKA 90 14	527.83	330.90	584	1.60	0.1112

Parameter	Estimate	Std Error	DDF	T	Pr > T
VUOSI*PAIKKA 90 15	1883.17	350.54	584	5.37	0.0000
VUOSI*PAIKKA 90 16	1212.24	448.86	584	2.70	0.0071
VUOSI*PAIKKA 90 19	-726.98	448.86	584	-1.62	0.1059
VUOSI*PAIKKA 91 3	606.41	279.30	584	2.17	0.0303
VUOSI*PAIKKA 91 13	837.56	331.93	584	2.52	0.0119
VUOSI*PAIKKA 91 14	291.14	342.07	584	0.85	0.3951
VUOSI*PAIKKA 91 15	1739.01	380.55	584	4.57	0.0000
VUOSI*PAIKKA 91 16	850.90	448.54	584	1.90	0.0583
VUOSI*PAIKKA 91 19	-897.09	491.42	584	-1.83	0.0684
VUOSI*PAIKKA 91 25	594.88	448.54	584	1.33	0.1853
VUOSI*PAIKKA 92 3	-1008.96	319.66	584	-3.16	0.0017
VUOSI*PAIKKA 92 13	616.97	353.22	584	1.75	0.0812
VUOSI*PAIKKA 92 14	-630.59	374.01	584	-1.69	0.0923
VUOSI*PAIKKA 92 15	-1297.09	416.56	584	-3.11	0.0019
VUOSI*PAIKKA 92 16	1344.94	449.20	584	2.99	0.0029
VUOSI*PAIKKA 92 25	835.86	449.20	584	1.86	0.0633
VUOSI*PAIKKA 93 3	1085.04	374.48	584	2.90	0.0039
VUOSI*PAIKKA 93 13	764.25	327.72	584	2.33	0.0200
VUOSI*PAIKKA 93 14	-2160.78	362.40	584	-5.96	0.0000
VUOSI*PAIKKA 93 19	-432.46	592.55	584	-0.73	0.4658
VUOSI*PAIKKA 93 21	674.29	592.55	584	1.14	0.2556
VUOSI*PAIKKA 93 25	-363.29	448.69	584	-0.81	0.4185
VUOSI*PAIKKA 94 3	416.97	450.55	584	0.93	0.3551
VUOSI*PAIKKA 94 13	1042.55	495.75	584	2.10	0.0359
VUOSI*PAIKKA 94 14	-590.91	424.87	584	-1.39	0.1648
VUOSI*PAIKKA*KOE 81 3 101	55.98	223.14	584	0.25	0.8020
VUOSI*PAIKKA*KOE 81 3 102	266.60	242.24	584	1.10	0.2715
VUOSI*PAIKKA*KOE 81 3 103	-223.60	233.87	584	-0.96	0.3394
VUOSI*PAIKKA*KOE 81 3 104	-77.85	233.87	584	-0.33	0.7393
VUOSI*PAIKKA*KOE 81 16 8	-14.10	295.94	584	-0.05	0.9620
VUOSI*PAIKKA*KOE 81 19 203	-165.26	296.00	584	-0.56	0.5768
VUOSI*PAIKKA*KOE 82 3 101	-18.40	210.79	584	-0.09	0.9305
VUOSI*PAIKKA*KOE 82 3 102	-113.45	221.60	584	-0.51	0.6089
VUOSI*PAIKKA*KOE 82 3 103	127.54	231.27	584	0.55	0.5815
VUOSI*PAIKKA*KOE 82 3 104	97.85	231.27	584	0.42	0.6724
VUOSI*PAIKKA*KOE 82 13 103	-165.13	233.86	584	-0.71	0.4804
VUOSI*PAIKKA*KOE 82 13 104	-103.89	227.56	584	-0.46	0.6482
VUOSI*PAIKKA*KOE 82 13 403	304.15	230.34	584	1.32	0.1872
VUOSI*PAIKKA*KOE 82 14 321	265.60	251.56	584	1.06	0.2915
VUOSI*PAIKKA*KOE 82 14 421	-137.42	251.16	584	-0.55	0.5845
VUOSI*PAIKKA*KOE 82 15 133	-106.56	260.60	584	-0.41	0.6828
VUOSI*PAIKKA*KOE 82 15 474	62.94	259.33	584	0.24	0.8083
VUOSI*PAIKKA*KOE 82 16 8	-30.84	295.99	584	-0.10	0.9170
VUOSI*PAIKKA*KOE 82 19 203	-123.74	295.99	584	-0.42	0.6760
VUOSI*PAIKKA*KOE 83 3 101	-149.07	212.34	584	-0.70	0.4829
VUOSI*PAIKKA*KOE 83 3 102	104.84	223.01	584	0.47	0.6385
VUOSI*PAIKKA*KOE 83 3 103	67.32	232.04	584	0.29	0.7718
VUOSI*PAIKKA*KOE 83 3 104	36.47	240.67	584	0.15	0.8796
VUOSI*PAIKKA*KOE 83 13 103	74.90	230.93	584	0.32	0.7458
VUOSI*PAIKKA*KOE 83 13 104	-344.87	230.93	584	-1.49	0.1359
VUOSI*PAIKKA*KOE 83 13 403	281.56	235.30	584	1.20	0.2320
VUOSI*PAIKKA*KOE 83 14 321	210.62	255.21	584	0.83	0.4095
VUOSI*PAIKKA*KOE 83 14 421	-164.18	255.73	584	-0.64	0.5211

Parameter	Estimate	Std Error	DDF	T	Pr > T
VUOSI*PAIKKA*KOE 83 15 133	154.16	256.22	584	0.60	0.5476
VUOSI*PAIKKA*KOE 83 15 474	-155.21	255.66	584	-0.61	0.5440
VUOSI*PAIKKA*KOE 83 16 8	21.69	295.99	584	0.07	0.9416
VUOSI*PAIKKA*KOE 83-19 203	-197.00	295.99	584	-0.67	0.5059
VUOSI*PAIKKA*KOE 84 3 101	-455.86	213.16	584	-2.14	0.0329
VUOSI*PAIKKA*KOE 84 3 102	286.86	234.87	584	1.22	0.2225
VUOSI*PAIKKA*KOE 84 3 103	33.27	241.42	584	0.14	0.8904
VUOSI*PAIKKA*KOE 84 3 104	120.64	241.42	584	0.50	0.6175
VUOSI*PAIKKA*KOE 84 13 103	7.08	238.27	584	0.03	0.9763
VUOSI*PAIKKA*KOE 84 13 104	-258.08	238.27	584	-1.08	0.2792
VUOSI*PAIKKA*KOE 84 13 403	129.51	248.07	584	0.52	0.6018
VUOSI*PAIKKA*KOE 84 14 321	57.36	295.93	584	0.19	0.8464
VUOSI*PAIKKA*KOE 84 15 133	9.90	258.84	584	0.04	0.9695
VUOSI*PAIKKA*KOE 84 15 474	58.09	260.11	584	0.22	0.8234
VUOSI*PAIKKA*KOE 84 16 8	32.54	295.93	584	0.11	0.9125
VUOSI*PAIKKA*KOE 84 19 203	-151.48	296.22	584	-0.51	0.6093
VUOSI*PAIKKA*KOE 85 3 101	-438.98	217.93	584	-2.01	0.0444
VUOSI*PAIKKA*KOE 85 3 102	146.37	234.93	584	0.62	0.5335
VUOSI*PAIKKA*KOE 85 3 103	223.23	234.67	584	0.95	0.3419
VUOSI*PAIKKA*KOE 85 3 104	98.73	241.50	584	0.41	0.6828
VUOSI*PAIKKA*KOE 85 13 103	55.94	238.26	584	0.23	0.8144
VUOSI*PAIKKA*KOE 85 13 104	-351.91	238.26	584	-1.48	0.1402
VUOSI*PAIKKA*KOE 85 13 403	255.80	248.05	584	1.03	0.3028
VUOSI*PAIKKA*KOE 85 14 321	127.72	257.39	584	0.50	0.6199
VUOSI*PAIKKA*KOE 85 14 421	-88.78	259.06	584	-0.34	0.7319
VUOSI*PAIKKA*KOE 85 15 133	59.69	258.81	584	0.23	0.8177
VUOSI*PAIKKA*KOE 85 15 474	-9.89	260.1	584	-0.04	0.9697
VUOSI*PAIKKA*KOE 85 16 8	-27.06	296.08	584	-0.09	0.9272
VUOSI*PAIKKA*KOE 85 19 203	-49.17	296.21	584	-0.17	0.8682
VUOSI*PAIKKA*KOE 86 3 101	-70.24	227.42	584	-0.31	0.7575
VUOSI*PAIKKA*KOE 86 3 102	92.90	255.39	584	0.36	0.7162
VUOSI*PAIKKA*KOE 86 3 103	18.96	255.39	584	0.07	0.9408
VUOSI*PAIKKA*KOE 86 3 104	42.99	255.39	584	0.17	0.8664
VUOSI*PAIKKA*KOE 86 13 103	91.97	240.77	584	0.38	0.7026
VUOSI*PAIKKA*KOE 86 13 104	-182.71	240.77	584	-0.76	0.4482
VUOSI*PAIKKA*KOE 86 13 403	91.94	258.53	584	0.36	0.7222
VUOSI*PAIKKA*KOE 86 14 321	34.87	263.27	584	0.13	0.8947
VUOSI*PAIKKA*KOE 86 14 421	48.77	266.05	584	0.18	0.8546
VUOSI*PAIKKA*KOE 86 15 133	64.21	264.40	584	0.24	0.8082
VUOSI*PAIKKA*KOE 86 15 474	-33.26	266.82	584	-0.12	0.9008
VUOSI*PAIKKA*KOE 86 16 8	-22.27	296.08	584	-0.08	0.9401
VUOSI*PAIKKA*KOE 86 19 203	-24.03	296.48	584	-0.08	0.9354
VUOSI*PAIKKA*KOE 87 14 321	75.00	257.66	584	0.29	0.7711
VUOSI*PAIKKA*KOE 87 14 421	-209.15	259.37	584	-0.81	0.4204
VUOSI*PAIKKA*KOE 87 15 133	-103.66	257.82	584	-0.40	0.6878
VUOSI*PAIKKA*KOE 87 15 474	40.36	259.48	584	0.16	0.8764
VUOSI*PAIKKA*KOE 87 16 8	-28.08	296.09	584	-0.09	0.9245
VUOSI*PAIKKA*KOE 88 3 101	200.39	219.18	584	0.91	0.3610
VUOSI*PAIKKA*KOE 88 3 102	-120.34	253.00	584	-0.48	0.6345
VUOSI*PAIKKA*KOE 88 3 103	-185.04	253.00	584	-0.73	0.4648
VUOSI*PAIKKA*KOE 88 3 104	-42.70	253.00	584	-0.17	0.8660
VUOSI*PAIKKA*KOE 88 3 105	82.97	241.27	584	0.34	0.7310
VUOSI*PAIKKA*KOE 88 13 103	415.12	244.00	584	1.70	0.0894

Liite 4 (6/7)

Parameter	Estimate	Std Error	DDF	T	Pr > T
VUOSI*PAIKKA*KOE 88 13 104	-379.78	244.00	584	-1.56	0.1201
VUOSI*PAIKKA*KOE 88 13 403	-41.32	249.42	584	-0.17	0.8685
VUOSI*PAIKKA*KOE 88 14 321	71.35	257.56	584	0.28	0.7819
VUOSI*PAIKKA*KOE 88 14 421	-110.95	259.25	584	-0.43	0.6688
VUOSI*PAIKKA*KOE 88 15 133	-332.32	260.96	584	-1.27	0.2034
VUOSI*PAIKKA*KOE 88 15 474	155.44	261.71	584	0.59	0.5528
VUOSI*PAIKKA*KOE 88 16 8	10.17	296.22	584	0.03	0.9726
VUOSI*PAIKKA*KOE 88 19 203	-138.84	296.22	584	-0.47	0.6394
VUOSI*PAIKKA*KOE 89 3 101	45.41	217.77	584	0.21	0.8349
VUOSI*PAIKKA*KOE 89 3 102	-198.86	239.73	584	-0.83	0.4072
VUOSI*PAIKKA*KOE 89 3 103	-18.05	252.55	584	-0.07	0.9430
VUOSI*PAIKKA*KOE 89 3 104	11.756	239.73	584	0.05	0.9609
VUOSI*PAIKKA*KOE 89 3 105	209.89	239.73	584	0.88	0.3816
VUOSI*PAIKKA*KOE 89 13 103	114.91	251.39	584	0.46	0.6478
VUOSI*PAIKKA*KOE 89 13 104	15.06	251.39	584	0.06	0.9522
VUOSI*PAIKKA*KOE 89 13 403	-97.26	251.39	584	-0.39	0.6990
VUOSI*PAIKKA*KOE 89 14 321	-48.62	259.19	584	-0.19	0.8513
VUOSI*PAIKKA*KOE 89 14 421	96.06	260.50	584	0.37	0.7124
VUOSI*PAIKKA*KOE 89 15 133	312.01	264.03	584	1.18	0.2378
VUOSI*PAIKKA*KOE 89 15 474	-271.48	264.03	584	-1.03	0.3043
VUOSI*PAIKKA*KOE 89 16 8	78.74	296.22	584	0.27	0.7905
VUOSI*PAIKKA*KOE 89 19 203	0.84	296.22	584	0.00	0.9977
VUOSI*PAIKKA*KOE 90 3 101	162.32	223.23	584	0.73	0.4674
VUOSI*PAIKKA*KOE 90 3 102	-233.44	240.98	584	-0.97	0.3331
VUOSI*PAIKKA*KOE 90 3 103	-139.83	240.98	584	-0.58	0.5620
VUOSI*PAIKKA*KOE 90 3 104	55.18	240.98	584	0.23	0.8190
VUOSI*PAIKKA*KOE 90 3 105	133.93	272.37	584	0.49	0.6231
VUOSI*PAIKKA*KOE 90 13 103	9.71	260.62	584	0.04	0.9703
VUOSI*PAIKKA*KOE 90 13 104	-84.22	253.46	584	-0.33	0.7398
VUOSI*PAIKKA*KOE 90 13 403	171.63	253.46	584	0.68	0.4986
VUOSI*PAIKKA*KOE 90 14 321	8.45	259.19	584	0.03	0.9740
VUOSI*PAIKKA*KOE 90 14 421	24.34	260.51	584	0.09	0.9256
VUOSI*PAIKKA*KOE 90 15 133	-17.16	264.03	584	-0.06	0.9482
VUOSI*PAIKKA*KOE 90 15 474	134.17	264.03	584	0.51	0.6115
VUOSI*PAIKKA*KOE 90 16 8	75.32	296.23	584	0.25	0.7994
VUOSI*PAIKKA*KOE 90 19 203	-45.17	296.23	584	-0.15	0.8789
VUOSI*PAIKKA*KOE 91 3 101	-10.97	229.04	584	-0.05	0.9618
VUOSI*PAIKKA*KOE 91 3 102	-36.32	255.07	584	-0.14	0.8868
VUOSI*PAIKKA*KOE 91 3 103	3.48	243.90	584	0.01	0.9886
VUOSI*PAIKKA*KOE 91 3 104	81.49	243.90	584	0.33	0.7384
VUOSI*PAIKKA*KOE 91 13 103	15.01	263.52	584	0.06	0.9546
VUOSI*PAIKKA*KOE 91 13 104	68.61	263.52	584	0.26	0.7947
VUOSI*PAIKKA*KOE 91 13 403	-31.58	264.33	584	-0.12	0.9049
VUOSI*PAIKKA*KOE 91 14 321	163.76	264.62	584	0.62	0.5363
VUOSI*PAIKKA*KOE 91 14 421	-145.67	267.06	584	-0.55	0.5856
VUOSI*PAIKKA*KOE 91 15 133	134.63	273.19	584	0.49	0.6223
VUOSI*PAIKKA*KOE 91 15 474	-26.57	273.18	584	-0.10	0.9225
VUOSI*PAIKKA*KOE 91 16 8	52.87	296.22	584	0.18	0.8584
VUOSI*PAIKKA*KOE 91 19 203	-55.74	296.49	584	-0.19	0.8509
VUOSI*PAIKKA*KOE 91 25 120	36.96	296.22	584	0.12	0.9007
VUOSI*PAIKKA*KOE 92 3 102	-498.09	254.96	584	-1.95	0.0512
VUOSI*PAIKKA*KOE 92 3 103	454.45	260.96	584	1.74	0.0821
VUOSI*PAIKKA*KOE 92 3 104	318.16	274.77	584	1.16	0.2474

Liite 4 (7/7)

VUOSI*PAIKKA*KOE 92 3 110	-337.21	274.77	584	-1.23	0.2202
VUOSI*PAIKKA*KOE 92 13 103	86.24	266.21	584	0.32	0.7461
VUOSI*PAIKKA*KOE 92 13 104	-165.16	266.21	584	-0.62	0.5352
VUOSI*PAIKKA*KOE 92 13 403	117.25	276.89	584	0.42	0.6721
VUOSI*PAIKKA*KOE 92 14 321	-1.68	275.23	584	-0.01	0.9951
VUOSI*PAIKKA*KOE 92 14 421	-37.50	279.22	584	-0.13	0.8932
VUOSI*PAIKKA*KOE 92 15 133	-198.77	279.13	584	-0.71	0.4767
VUOSI*PAIKKA*KOE 92 15 474	118.17	281.40	584	0.42	0.6747
VUOSI*PAIKKA*KOE 92 16 8	83.56	296.23	584	0.28	0.7780
VUOSI*PAIKKA*KOE 92 25 121	51.93	296.23	584	0.18	0.8609
VUOSI*PAIKKA*KOE 93 3 101	58.80	270.69	584	0.22	0.8281
VUOSI*PAIKKA*KOE 93 3 102	64.40	277.95	584	0.23	0.8168
VUOSI*PAIKKA*KOE 93 3 103	-55.78	277.95	584	-0.20	0.8410
VUOSI*PAIKKA*KOE 93 13 103	70.07	257.08	584	0.27	0.7853
VUOSI*PAIKKA*KOE 93 13 104	-28.21	257.08	584	-0.11	0.9127
VUOSI*PAIKKA*KOE 93 13 111	5.62	276.13	584	0.02	0.9838
VUOSI*PAIKKA*KOE 93 14 113	9.03	278.88	584	0.03	0.9742
VUOSI*PAIKKA*KOE 93 14 321	-143.29	274.49	584	-0.52	0.6018
VUOSI*PAIKKA*KOE 93 19 114	-26.87	297.20	584	-0.09	0.9280
VUOSI*PAIKKA*KOE 93 21 108	41.89	297.20	584	0.14	0.8879
VUOSI*PAIKKA*KOE 93 25 122	-22.57	296.22	584	-0.08	0.9393
VUOSI*PAIKKA*KOE 94 3 101	25.90	296.24	584	0.09	0.9303
VUOSI*PAIKKA*KOE 94 13 103	64.78	296.51	584	0.22	0.8271
VUOSI*PAIKKA*KOE 94 14 321	-36.71	296.09	584	-0.12	0.9014

The MIXED Procedure

Solution for Fixed Effects

Parameter	Estimate	Std Error	DDF	T	Pr > T
LAJIKE 100835	4434.67	192.90	584	22.99	0.0000
LAJIKE 101851	4046.67	163.78	584	24.71	0.0000
LAJIKE 103045	4564.82	161.22	584	28.31	0.0000
LAJIKE 103063	4398.28	168.03	584	26.17	0.0000
LAJIKE 103077	4855.48	251.93	584	19.27	0.0000
LAJIKE 103079	4395.50	218.44	584	20.12	0.0000
LAJIKE 222585	4579.31	179.56	584	25.50	0.0000
LAJIKE 523977	4614.72	197.89	584	23.32	0.0000

The MIXED Procedure

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	8	584	114.47	0.0000

The MIXED Procedure

ESTIMATE Statement Results

Parameter	Estimate	Std Error	DDF	T	Pr > T
100835-103045	-130.15	147.93	584	-0.88	0.3793
101851-103045	-518.14	107.69	584	-4.81	0.0000
103063-103045	-166.53	114.03	584	-1.46	0.1447
103077-103045	290.66	218.88	584	1.33	0.1847
103079-103045	-169.32	179.83	584	-0.94	0.3468
222585-103045	14.49	130.03	584	0.11	0.9113
523977-103045	49.89	154.14	584	0.32	0.7463

```
LIBNAME ohra '<lst2.jukka>';
PROC MIXED DATA=ohra.svehna2;
  CLASS lajike vuosi paikka koe;
  MODEL sato= lajike ksato*lajike/noint s;
RUN;
```

The MIXED Procedure
Model Fitting Information for SATO

Description	Value
Observations	592
Variance Estimate	176564.0
Standard Deviation Estimate	420.1952
REML Log Likelihood	-4385.24
Akaike's Information Criterion	-4386.24
Schwarz's Bayesian Criterion	-4388.42
2 REML Log Likelihood	8770.482

The MIXED Procedure
Solution for Fixed Effects

Parameter	Estimate	Std Error	DDF	T	Pr > T
LAJIKE 100835	-63.52	234.10	576	-0.27	0.7862
LAJIKE 101851	74.16	132.54	576	0.56	0.5760
LAJIKE 103045	-83.28	124.28	576	-0.67	0.5031
LAJIKE 103063	146.50	154.19	576	0.95	0.3424
LAJIKE 103077	742.98	746.36	576	1.00	0.3199
LAJIKE 103079	-118.88	352.31	576	-0.34	0.7359
LAJIKE 222585	290.07	187.82	576	1.54	0.1230
LAJIKE 523977	-882.00	260.62	576	-3.38	0.0008
KSATO*LAJIKE 100835	1.0093	0.0480	576	21.02	0.0000
KSATO*LAJIKE 101851	0.9101	0.0285	576	31.92	0.0000
KSATO*LAJIKE 103045	1.0536	0.0264	576	39.90	0.0000
KSATO*LAJIKE 103063	0.9671	0.0329	576	29.31	0.0000
KSATO*LAJIKE 103077	0.9257	0.1483	576	6.24	0.0000
KSATO*LAJIKE 103079	1.0259	0.0744	576	13.78	0.0000
KSATO*LAJIKE 222585	0.9585	0.0397	576	24.12	0.0000
KSATO*LAJIKE 523977	1.2725	0.0598	576	21.27	0.0000

The MIXED Procedure
Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	8	576	2.09	0.0354
KSATO*LAJIKE	8	576	646.83	0.0000

```

LIBNAME ohra '<lst2.jukka>';
* OPTIONS LS=108 PS=42;
PROC MIXED DATA=ohra.ohra2;
CLASS lajike vuosi paikka;
MODEL sato=lajike;
RANDOM vuosi paikka vuosi*paikka lajike*vuosi lajike*paikka;
RUN;

```

The MIXED Procedure

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > Z
VUOSI	0.3420	370908.52	225719.26	1.64	0.1003
PAIKKA	0.1480	160563.49	97591.02	1.65	0.0999
VUOSI*PAIKKA	0.5831	632390.92	106192.69	5.96	0.0000
LAJIKE*VUOSI	0.0139	15139.59	12132.03	1.25	0.2121
LAJIKE*PAIKKA	0.0260	28277.50	13107.93	2.16	0.0310
Residual	1.0000	1084485.28	39766.07	27.27	0.0000

The MIXED Procedure

Model Fitting Information for SATO

Description	Value
Observations	1745
Variance Estimate	1084485
Standard Deviation Estimate	1041.386
REML Log Likelihood	-14703.8
Akaike's Information Criterion	-14709.8
Schwarz's Bayesian Criterion	-14726.1
2 REML Log Likelihood	29407.51

The MIXED Procedure

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	9	1735	4.84	0.0000

```

LIBNAME ohra '<lst2.jukka>';
* OPTIONS LS=108 PS=42;
PROC MIXED DATA=ohra.ohra2;
CLASS lajike vuosi paikka koe;
MODEL sato=lajike;
RANDOM lajike*paikka vuosi*paikka paikka*koe vuosi*paikka*koe;
RUN;

```

The MIXED Procedure

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > Z
LAJIKE*PAIKKA	0.1363	27815.35	6705.11	4.15	0.0000
VUOSI*PAIKKA	3.5899	732441.92	155033.90	4.72	0.0000
PAIKKA*KOE	2.6770	546190.28	148364.35	3.68	0.0002
VUOSI*PAIKKA*KOE	3.6844	751713.70	92263.85	8.15	0.0000
Residual	1.0000	204024.11	7941.87	25.69	0.0000

The MIXED Procedure

Model Fitting Information for SATO

Description	Value
Observations	1745
Variance Estimate	204024.1
Standard Deviation Estimate	451.6903
REML Log Likelihood	-13685.0
Akaike's Information Criterion	-13690.0

The MIXED Procedure

Model Fitting Information for SATO

Description	Value
Schwarz's Bayesian Criterion	-13703.7
2 REML Log Likelihood	27370.05

The MIXED Procedure
Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
LAJIKE	9	1735	11.91	0.0000

Julkaisun sarja ja numero
Maatalouden tutkimuskeskuksen julkaisuja.
Sarja B 2

Julkaisuaika (kk ja vuosi)
Maaliskuu 1996

Tekijä(t)

Jukka Öfversten
Hannele Nikander

Tutkimushankkeen nimi

Toimeksiantaja(t)
Maatalouden tutkimuskeskus

Nimike
Maatalouden tutkimuskeskuksen julkaisusarjat B
Lajikekoesarjojen analysointi

Tiivistelmä

Kasvinjalostajat ja siemenkauppiat testaavat uusien peltokasvilajikkeiden viljelyarvon puolueettomien tutkimuslaitosten järjestämissä kenttäkokeissa ennen lajikkeiden kaupallista viljelyä. Suomesa kokeita koordinoi maa- ja metsätalousministeriön toimeksiannosta Maatalouden tutkimuskeskus. Lajikekokeista koottuja tiivistettyjä lajikekoeaineistoja on Maatalouden tutkimuskeskuksen käytössä noin kolmen vuosikymmenen ajalta. Tiivistettyjen koeaineistojen tilastollinen analyysi on helppoa, jos jokainen vertailussa tarvittava lajike on ollut mukana kaikissa vertailuun käytettävissä kokeissa. Paljon lajikkeita sisältävät monivuotiset koesarjat eivät yleensä ole tällaisia. Suurista aineistoista voidaan kuitenkin aina valita osa-aineistoja, joissa tietyn kasvin kaksi eri lajiketta ovat olleet yhtä aikaa samoissa kokeissa. Tällaista aineistoa voidaan käyttää kätevästi kyseisten kahden lajikkeen vertaamiseen. Osa-aineistojen valintaa on viime aikoihin asti käytetty lähtökohtana lajikekoesarjojen analysoimiseen. Tämä parivertailuksi kutsuttu menettely on kuitenkin monessa mielessä kyseenalainen, sillä osa-aineistojen käyttö voi aiheuttaa ristiriitaisia johtopäätöksiä ja heikentää tilastollisten testien voimakkuutta. Sen avulla ei myöskään saada juuri mitään viitteitä lajikkeiden herkkyydestä erilaisille ulkoisille olosuhteille.

Tämän raportin tarkoituksena on kuvata lineaaristen sekamallien käyttöä lajikekoesarjojen analysoinnin lähtökohtana. Raportissa osoitetaan, että sekamalleilla saadaan päätöksenteon kannalta käyttökelpoisempia ja joustavampia tuloksia kuin perinteistä parivertailua käyttämällä. Johtopäätökset voidaan tehdä kaiken käytettävissä olevan aineiston perusteella. Keskiarvoparametrien lisäksi voidaan mallintaa myös vastemuuttujien hajontaa. Hajonnan mallintamisen avulla tehtävät johtopäätökset voidaan ulottaa koskemaan joko koko viljelyaluetta tai vain yhtä tai muutamaa koepaikkaa. Raportin esimerkeissä käytetään virallisista syysvehnän ja ohran lajikekokeista koottuja aineistoja. Kaikissa malliratkaisuissa käytetään vastemuuttujana koetuloksina saatuja jyväsatoja.

Avainsanat
peltokasvitutkimus, lajikevertailu, mallintaminen, sekamalli,
sopeutuminen

Toimintayksikkö
Tietopalveluyksikkö, 31600 JOKIOINEN

ISSN

1238-9943

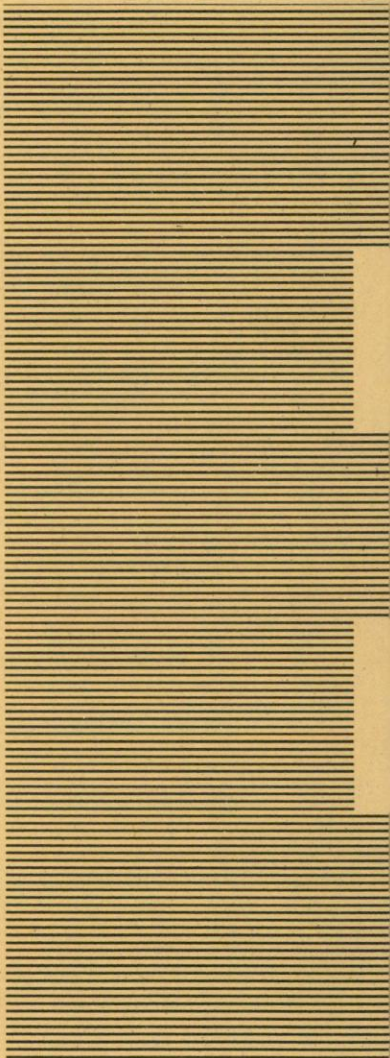
 Tuloksia voi soveltaa luomuviljelyssä

Myynti: MTT tietopalveluyksikkö, 31600 JOKIOINEN
Puh. (916) 41 881
Telekopio (916) 4188 339

Sivuja

27 s. + 8 liitettä

Hinta



Jokioinen 1996
ISSN 1238-9943