Package ‘agridat’

December 20, 2021

Title  Agricultural Datasets
Version  1.20
Type  Package
Description  Datasets from books, papers, and websites related to agriculture. Example graphics and analyses are included. Data come from small-plot trials, multi-environment trials, uniformity trials, yield monitors, and more.
Suggests  AER, agricolae, betareg, broom, car, coin, corrgram, desplot, dplyr, effects, equivalence, emmeans, FrF2, gam, gge, ggplot2, gnm, gstat, HH, knitr, lattice, latticeExtra, lme4, lucid, mapproj, maps, MASS, MCMCglmm, metafor, mgcv, NADA, nlme, nullabor, ordinal, pbkrtest, pls, pscl, reshape2, rgdal, rmarkdown, qicharts, qtl, sp, SpATS, survival, vcd, testthat
License  CC BY-SA 4.0
LazyData  yes
URL  https://kwstat.github.io/agridat/
BugReports  https://github.com/kwstat/agridat/issues
VignetteBuilder  knitr
Language  en-US
Encoding  UTF-8
RoxygenNote  7.1.1
NeedsCompilation  no
Author  Kevin Wright [aut, cre] (<https://orcid.org/0000-0002-0617-8673>)
Maintainer  Kevin Wright <kw.stat@gmail.com>
Repository  CRAN
Date/Publication  2021-12-20 08:50:02 UTC
R topics documented:

aastveit.barley ......................................................... 9
acorsi.grayleafspot .................................................. 11
adugna.sorghum ...................................................... 14
agridat ................................................................. 15
allcroft.lodging ....................................................... 24
alwan.lamb ........................................................... 25
ansari.wheat.uniformity ........................................... 28
arankacami.groundnut.uniformity ............................... 29
archbold.apple ......................................................... 30
ars.earlywhitecorn96 .............................................. 32
australia.soybean .................................................... 33
baker.barley.uniformity ............................................. 35
baker.strawberry.uniformity ...................................... 37
baker.wheat.uniformity ............................................. 38
bancroft.peanut.uniformity ....................................... 39
barrero.maize ........................................................ 41
batchelor.uniformity ............................................... 43
battese.survey ....................................................... 45
beall.webworms ...................................................... 48
beaven.barley ........................................................ 50
becker.chicken ....................................................... 51
besag.bayesian ....................................................... 52
besag.beans .......................................................... 54
besag.checks .......................................................... 56
besag.elbatan ........................................................ 57
besag.endive .......................................................... 59
besag.met .............................................................. 60
besag.triticale ........................................................ 64
blackman.wheat ...................................................... 66
bliss.borers ............................................................ 67
bond.diallel ............................................................ 69
bose.multi.uniformity .............................................. 71
box.cork ................................................................. 72
bradley.multi.uniformity .......................................... 74
brandle.rape ........................................................... 76
brandt.switchback .................................................. 78
bridges.cucumber ................................................... 79
broadbalk.wheat ..................................................... 81
bryan.corn.uniformity ............................................. 83
buntaran.wheat ....................................................... 84
burgueno.alpha ....................................................... 85
burgueno.rowcol ..................................................... 88
burgueno.unreplicated ............................................ 90
butron.maize ......................................................... 91
byers.apple ........................................................... 94
caribbean.maize ..................................................... 95
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>carlson.germination</td>
<td>97</td>
</tr>
<tr>
<td>carmer.density</td>
<td>98</td>
</tr>
<tr>
<td>cate.potassium</td>
<td>100</td>
</tr>
<tr>
<td>chinloy.fractionalfactorial</td>
<td>101</td>
</tr>
<tr>
<td>christidis.competition</td>
<td>103</td>
</tr>
<tr>
<td>christidis.cotton.uniformity</td>
<td>105</td>
</tr>
<tr>
<td>christidis.wheat.uniformity</td>
<td>106</td>
</tr>
<tr>
<td>cleveland.soil</td>
<td>107</td>
</tr>
<tr>
<td>cochran.beets</td>
<td>109</td>
</tr>
<tr>
<td>cochran.bib</td>
<td>110</td>
</tr>
<tr>
<td>cochran.crd</td>
<td>111</td>
</tr>
<tr>
<td>cochran.eelworms</td>
<td>112</td>
</tr>
<tr>
<td>cochran.factorial</td>
<td>114</td>
</tr>
<tr>
<td>cochran.latin</td>
<td>115</td>
</tr>
<tr>
<td>cochran.lattice</td>
<td>116</td>
</tr>
<tr>
<td>cochran.wireworms</td>
<td>117</td>
</tr>
<tr>
<td>connolly.potato</td>
<td>119</td>
</tr>
<tr>
<td>cornelius.maize</td>
<td>121</td>
</tr>
<tr>
<td>corsten.interaction</td>
<td>123</td>
</tr>
<tr>
<td>cox.strip.split</td>
<td>124</td>
</tr>
<tr>
<td>cramer.cucumber</td>
<td>126</td>
</tr>
<tr>
<td>crampton.pig</td>
<td>129</td>
</tr>
<tr>
<td>crossa.wheat</td>
<td>130</td>
</tr>
<tr>
<td>crowder.seeds</td>
<td>132</td>
</tr>
<tr>
<td>cullis.earlygen</td>
<td>137</td>
</tr>
<tr>
<td>damesa.maize</td>
<td>140</td>
</tr>
<tr>
<td>darwin.maize</td>
<td>141</td>
</tr>
<tr>
<td>dasilva.maize</td>
<td>143</td>
</tr>
<tr>
<td>dasilva.soybean.uniformity</td>
<td>145</td>
</tr>
<tr>
<td>davidian.soybean</td>
<td>146</td>
</tr>
<tr>
<td>davies.pasture.uniformity</td>
<td>148</td>
</tr>
<tr>
<td>day.wheat.uniformity</td>
<td>149</td>
</tr>
<tr>
<td>denis.missing</td>
<td>151</td>
</tr>
<tr>
<td>denis.ryegrass</td>
<td>152</td>
</tr>
<tr>
<td>depalluel.sheep</td>
<td>153</td>
</tr>
<tr>
<td>devries.pine</td>
<td>154</td>
</tr>
<tr>
<td>digby.jointregression</td>
<td>156</td>
</tr>
<tr>
<td>diggle.cow</td>
<td>158</td>
</tr>
<tr>
<td>draper.safflower.uniformity</td>
<td>160</td>
</tr>
<tr>
<td>durban.competition</td>
<td>162</td>
</tr>
<tr>
<td>durban.rowcol</td>
<td>163</td>
</tr>
<tr>
<td>durban.splitplot</td>
<td>165</td>
</tr>
<tr>
<td>eden.nonnormal</td>
<td>167</td>
</tr>
<tr>
<td>eden.potato</td>
<td>170</td>
</tr>
<tr>
<td>eden.tea.uniformity</td>
<td>172</td>
</tr>
<tr>
<td>edwards.oats</td>
<td>173</td>
</tr>
<tr>
<td>engelstad.nitro</td>
<td>175</td>
</tr>
<tr>
<td>fan.stability</td>
<td>177</td>
</tr>
</tbody>
</table>
federer.diagcheck .................................................. 178
federer.tobacco ................................................... 181
fisher.barley ...................................................... 183
fisher.latin ....................................................... 185
foulley.calving .................................................... 187
fox.wheat .......................................................... 189
garber.multi.uniformity ........................................ 190
gartner.corn ....................................................... 191
gathmann.bt ......................................................... 194
gauch.soy .......................................................... 196
george.wheat ....................................................... 197
giles.wheat ........................................................ 199
gilmour.serpentine ............................................... 201
gilmour.slatehall ................................................ 203
gomez.fractionalfactorial .................................... 204
gomez.groupsplot ............................................... 206
gomez.heterogeneity .......................................... 207
gomez.heteroskedastic ....................................... 208
gomez.multilocsplitplot ..................................... 209
gomez.nitrogen ................................................. 211
gomez.nonnormal1 .............................................. 212
gomez.nonnormal2 .............................................. 214
gomez.nonnormal3 .............................................. 215
gomez.rice.uniformity ....................................... 217
gomez.seedrate ................................................. 218
gomez.splitplot.subsample .................................. 219
gomez.splitplot ................................................ 221
gomez.stripplot ................................................ 222
gomez.stripssplitplot ....................................... 225
gomez.wetdry ..................................................... 226
gotway.hessianfly ............................................... 227
goulden.barley.uniformity ..................................... 229
goulden.eggs ....................................................... 230
goulden.latin ..................................................... 231
goulden.splitsplit .............................................. 232
graybill.heteroskedastic .................................... 234
gregory.cotton ................................................... 236
grover.diallel ..................................................... 238
grover.rcb.subsample ......................................... 239
gumpertz.pepper ................................................ 240
hadusch.lettuce ................................................ 242
hanks.sprinkler ................................................ 244
hanover.whitepine ............................................. 247
harris.multi.uniformity ....................................... 248
harris.wateruse ................................................ 251
harrison.priors .................................................. 255
hartman.tomato.uniformity .................................. 257
harvey.lsmeans .................................................. 259
R topics documented:

lambert.soiltemp .................................................. 340
lander.multi.uniformity ........................................ 341
lasrosas.corn .................................................. 343
lavoranti.eucalyptus ............................................ 345
laycock.tea.uniformity ........................................... 346
lee.potatoblight ................................................ 348
lehmann.millet.uniformity .................................... 350
lehner.soybeanmold ............................................ 352
lessman.sorghum.uniformity .................................. 354
li.millet.uniformity ............................................ 356
libs ................................................................. 357
lillemo.wheat .................................................. 358
lin.superiority .................................................. 360
lin.unbalanced .................................................. 362
linder.wheat ..................................................... 364
little.splitblock ................................................ 365
loesell.bean.uniformity ....................................... 366
lonnquist.maize ................................................ 368
lord.rice.uniformity ........................................... 370
love.cotton.uniformity ....................................... 371
lu.stability ..................................................... 373
lucas.switchback .............................................. 374
lyon.potato.uniformity ........................................ 376
lyons.wheat ..................................................... 377
magistad.pineapple.uniformity ............................... 378
masood.rice.uniformity ....................................... 380
mcclelland.corn.uniformity .................................. 381
mcconway.turnip .............................................. 383
mcleod.barley .................................................. 384
mead.cauliflower ............................................. 386
mead.cowpea.maize .......................................... 387
mead.germination ............................................. 389
mead.lamb ....................................................... 390
mead.strawberry ............................................... 393
mead.turnip ..................................................... 394
mercer.mangold.uniformity ................................. 395
mercer.wheat.uniformity ..................................... 397
miguez.biomass ................................................ 398
minnesota.barley.weather .................................... 400
minnesota.barley.yield ....................................... 403
montgomery.wheat.uniformity ............................... 405
moore.uniformity .............................................. 407
nagai.strawberry.uniformity ............................... 409
nair.turmeric.uniformity ..................................... 411
narain.sorghum.uniformity ................................. 412
nass.corn ....................................................... 413
nebraska.farmincome ......................................... 415
nonnecke.peas.uniformity .................................... 416
nonnecke.sweetcorn.uniformity ........................................ 418
obsi.potato.uniformity ........................................... 419
odland.soybean.uniformity ........................................ 420
omer.sorghum ......................................................... 422
onofri.winterwheat ............................................... 423
ortiz.tomato .............................................................. 424
pacheco.soybean ....................................................... 427
paez.coffee.uniformity ............................................ 428
panse.cotton.uniformity ........................................... 429
parker.orange.uniformity ........................................... 431
patterson.switchback ............................................. 433
pearce.apple .............................................................. 434
pearl.kernels ............................................................. 436
pederson.lettuce.repeated ...................................... 437
perry.springwheat ................................................... 439
petersen.sorghum.cowpea ......................................... 441
piepho.barley.uniformity ......................................... 442
piepho.cocksfoot ................................................... 444
polson.safflower.uniformity ..................................... 445
ratkowsky.onions .................................................... 447
reid.grasses ............................................................. 448
ridout.appleshoots .................................................. 449
robinson.peanut.uniformity ....................................... 452
rothamsted.brussels ............................................... 453
rothamsted.oats ....................................................... 454
ryder.groundnut ...................................................... 456
salmon.bunt ............................................................. 458
sawyer.multi.uniformity ......................................... 459
sayer.sugarcane.uniformity ..................................... 462
senshu.rice .............................................................. 463
shafi.tomato.uniformity ........................................... 465
shafii.rapeseed .......................................................... 466
shaw.oats ................................................................. 468
silva.cotton .............................................................. 469
sinclair.clover .......................................................... 472
smith.beans.uniformity ........................................... 474
smith.corn.uniformity ............................................. 476
snedecor.asparagus .................................................. 478
snijders.fusarium ..................................................... 481
stephens.sorghum.uniformity ................................... 483
steptoe.morex.pheno ................................................ 484
stickler.sorghum.uniformity .................................... 487
stirret.borers .......................................................... 489
streibig.competition ............................................... 490
strickland.apple.uniformity ..................................... 492
strickland.grape.uniformity ..................................... 493
strickland.peach.uniformity .................................... 494
strickland.tomato.uniformity ................................... 495
<table>
<thead>
<tr>
<th>R topics documented:</th>
</tr>
</thead>
<tbody>
<tr>
<td>stroup.nin</td>
</tr>
<tr>
<td>stroup.splitplot</td>
</tr>
<tr>
<td>student.barley</td>
</tr>
<tr>
<td>tai.potato</td>
</tr>
<tr>
<td>talbot.potato</td>
</tr>
<tr>
<td>theobald.barley</td>
</tr>
<tr>
<td>theobald.covariate</td>
</tr>
<tr>
<td>thompson.cornsoy</td>
</tr>
<tr>
<td>turner.herbicide</td>
</tr>
<tr>
<td>urquhart.feedlot</td>
</tr>
<tr>
<td>usgs.herbicides</td>
</tr>
<tr>
<td>vaneeuwijk.drymatter</td>
</tr>
<tr>
<td>vaneeuwijk.fusarium</td>
</tr>
<tr>
<td>vaneeuwijk.nematodes</td>
</tr>
<tr>
<td>vargas.txe</td>
</tr>
<tr>
<td>vargas.wheat1</td>
</tr>
<tr>
<td>vargas.wheat2</td>
</tr>
<tr>
<td>verbyla.lupin</td>
</tr>
<tr>
<td>vold.longterm</td>
</tr>
<tr>
<td>vsn.lupin3</td>
</tr>
<tr>
<td>wallace.iowaland</td>
</tr>
<tr>
<td>walsh.cottonprice</td>
</tr>
<tr>
<td>wassom.brome.uniformity</td>
</tr>
<tr>
<td>waynick.soil</td>
</tr>
<tr>
<td>wedderburn.barley</td>
</tr>
<tr>
<td>weiss.incblock</td>
</tr>
<tr>
<td>weiss.lattice</td>
</tr>
<tr>
<td>welch.bermudagrass</td>
</tr>
<tr>
<td>wheately.carrot</td>
</tr>
<tr>
<td>wiebe.wheat.uniformity</td>
</tr>
<tr>
<td>wiedemann.safflower.uniformity</td>
</tr>
<tr>
<td>williams.barley.uniformity</td>
</tr>
<tr>
<td>williams.cotton.uniformity</td>
</tr>
<tr>
<td>williams.trees</td>
</tr>
<tr>
<td>woodman.pig</td>
</tr>
<tr>
<td>wyatt.multi.uniformity</td>
</tr>
<tr>
<td>yan.winterwheat</td>
</tr>
<tr>
<td>yang.barley</td>
</tr>
<tr>
<td>yates.missing</td>
</tr>
<tr>
<td>yates.oats</td>
</tr>
<tr>
<td>zuidhof.broiler</td>
</tr>
</tbody>
</table>

**Index**  575
Description

Average height for 15 genotypes of barley in each of 9 years. Also 19 covariates in each of the 9 years.

Usage

data("aastveit.barley.covs")
data("aastveit.barley.height")

Format

The 'aastveit.barley.covs' dataframe has 9 observations on the following 20 variables.

year  year
R1  avg rainfall (mm/day) in period 1
R2  avg rainfall (mm/day) in period 2
R3  avg rainfall (mm/day) in period 3
R4  avg rainfall (mm/day) in period 4
R5  avg rainfall (mm/day) in period 5
R6  avg rainfall (mm/day) in period 6
S1  daily solar radiation (ca/cm^2) in period 1
S2  daily solar radiation (ca/cm^2) in period 2
S3  daily solar radiation (ca/cm^2) in period 3
S4  daily solar radiation (ca/cm^2) in period 4
S5  daily solar radiation (ca/cm^2) in period 5
S6  daily solar radiation (ca/cm^2) in period 6
ST  sowing date
T1  avg temp (deg Celsius) in period 1
T2  avg temp (deg Celsius) in period 2
T3  avg temp (deg Celsius) in period 3
T4  avg temp (deg Celsius) in period 4
T5  avg temp (deg Celsius) in period 5
T6  avg temp (deg Celsius) in period 6
value value of the covariate

The 'aastveit.barley.height' dataframe has 135 observations on the following 3 variables.

year  year, 9
gen  genotype, 15 levels
height height (cm)
Details

Experiments were conducted at As, Norway. The height dataframe contains average plant height (cm) of 15 varieties of barley in each of 9 years. The growth season of each year was divided into eight periods from sowing to harvest. Because the plant stop growing about 20 days after ear emergence, only the first 6 periods are included here. Used with permission of Harald Martens.

Source


References


Examples

## Not run:

```r
library(agridat)
data("aastveit.barley.covs")
data("aastveit.barley.height")

# First, PCA of each matrix separately

Z <- acast(aastveit.barley.height, year ~ gen, value.var="height")
Z <- sweep(Z, 1, rowMeans(Z))
Z <- sweep(Z, 2, colMeans(Z)) # Double-centered
sum(Z^2)*4 # Total SS = 10165
sv <- svd(Z)$d
round(100 * sv^2/sum(sv^2),1) # Prop of variance each axis
# Aastveit Figure 1. PCA of height
biplot(prcomp(Z),
    main="aastveit.barley - height", cex=0.5)

U <- aastveit.barley.covs
rownames(U) <- U$year
U$year <- NULL
U <- scale(U) # Standardized covariates
sv <- svd(U)$d
# Proportion of variance on each axis
round(100 * sv^2/sum(sv^2),1)

# Now, PLS relating the two matrices
m1 <- plsr(Z-U)
```
acorsi.grayleafspot

Multi-environment trial evaluating 36 maize genotypes in 9 locations

Description

Multi-environment trial evaluating 36 maize genotypes in 9 locations

Usage

data(acorsi.grayleafspot)
Format

A data frame with 324 observations on the following 3 variables.

- gen: genotype, 36 levels
- env: environment, 9 levels
- rep: replicate, 2 levels
- y: grey leaf spot severity

Details

Experiments conducted in 9 environments in Brazil in 2010-11. Each location had an RCB with 2 reps.

The response variable is the percentage of leaf area affected by gray leaf spot within each experimental unit (plot).

Acorsi et al. use this data to illustrate the fitting of a generalized AMMI model with non-normal data.

Source


Electronic data and R code kindly provided by Marlon Coan.

References

None

Examples

```r
## Not run:
library(agridat)
data(acorsi.grayleafspot)
dat <- acorsi.grayleafspot

# Acorsi figure 2. Note: Acorsi used cell means
op <- par(mfrow=c(2,1), mar=c(5,4,3,2))
labs(lattice)
boxplot(y ~ env, dat, las=2,
       xlab="environment", ylab="GLS severity")
title("acorsi.grayleafspot")
boxplot(y ~ gen, dat, las=2,
       xlab="genotype", ylab="GLS severity")
par(op)

# GLM models
# glm main-effects model with logit u(1-u) and wedderburn u^2(1-u)^2
# variance functions
# glm1 <- glm(y~ env/rep + gen + env, data=dat, family=quasibinomial)
# glm2 <- glm(y~ env/rep + gen + env, data=dat, family=wedderburn)
# plot(glm2, which=1); plot(glm2, which=2)

# GAMM models of Acorsi. See also section 7.4 of Turner
# "Generalized nonlinear models in R: An overview of the gnm package"

# full gnm model with wedderburn, seems to work
libs(gnm)
set.seed(1)

gnm1 <- gnm(y ~ env/rep + env + gen + instances(Mult(env,gen),2),
            data=dat,
            family=wedderburn, iterMax =800)
device(gnm1) # 433.8548
# summary(gnm1)
# anova(gnm1, test ="F") # anodev, Acorsi table 4
## Df Deviance Resid. Df Resid. Dev F Pr(>F)
## NULL 647 3355.5
## env 8 1045.09 639 2310.4 68.4696 < 2.2e-16 ***
## env:rep 9 12.33 630 2298.1 0.7183 0.6923
## gen 35 1176.23 595 1121.9 17.6142 < 2.2e-16 ***
## Mult(env, gen, inst = 1) 42 375.94 553 745.9 4.6915 < 2.2e-16 ***
## Mult(env, gen, inst = 2) 40 312.06 513 433.9 4.0889 3.712e-14 ***

# maybe better, start simple and build up the model

gnm2a <- gnm(y ~ env/rep + env + gen,
             data=dat,
             family=wedderburn, iterMax =800)

# add first interaction term
res2a <- residSVD(gnm2a, env, gen, 2)
gnm2b <- update(gnm2a, . ~ . + Mult(env,gen,inst=1),
                start = c(coef(gnm2a), res2a[, 1]))
device(gnm2b) # 692.19

# add second interaction term
res2b <- residSVD(gnm2b, env, gen, 2)
gnm2c <- update(gnm2b, . ~ . + Mult(env,gen,inst=1) + Mult(env,gen,inst=2),
                start = c(coef(gnm2a), res2a[, 1], res2b[,1]))
device(gnm2c) # 433.8548
# anova(gnm2c) # weird error message

# note, to build the ammi biplot, use the first column of res2a to get
# axis 1, and the FIRST column of res2b to get axis 2. Slightly confusing
emat <- cbind(res2a[1:9, 1], res2b[1:9, 1])
rownames(emat) <- gsub("fac1:","", rownames(emat))

emat <- cbind(res2a[10:45, 1], res2b[10:45, 1])
rownames(gmat) <- gsub("fac2:","", rownames(gmat))
adugna.sorghum

Multi-environment trial of sorghum at 3 locations across 5 years

Description

Multi-environment trial of sorghum at 3 locations across 5 years

Format

A data frame with 289 observations on the following 6 variables.

- gen: genotype, 28 levels
- trial: trial, 2 levels
- env: environment, 13 levels
- yield: yield kg/ha
- year: year, 2001-2005
- loc: location, 3 levels

Details

Sorghum yields at 3 locations across 5 years. The trials were carried out at three locations in dry, hot lowlands of Ethiopia:
- Melkassa (39 deg 21 min E, 8 deg 24 min N)
- Mieso (39 deg 22 min E, 8 deg 41 min N)
- Kobo (39 deg 37 min E, 12 deg 09 min N)

Trial 1 was 14 hybrids and one open-pollinated variety.
Trial 2 was 12 experimental lines.

Used with permission of Asfaw Adugna.

Source

Examples

## Not run:

```r
library(agridat)
data(adugna.sorghum)
dat <- adugna.sorghum

libs(lattice)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(yield ~ env*gen, data=dat, main="adugna.sorghum gxe heatmap",
         col.regions=redblue)

# Genotype means match Adugna
 tapply(dat$yield, dat$gen, mean)

# CV for each genotype. G1..G15 match, except for G2.
# The table in Adugna scrambles the means for G16..G28
libs(reshape2)
mat <- acast(dat, gen~env, value.var='yield')
round(sqrt(apply(mat, 1, var, na.rm=TRUE)) / apply(mat, 1, mean, na.rm=TRUE) * 100,2)

 dat1 <- droplevels(subset(dat, trial=="T1"))
 mat1 <- acast(dat1, gen~env, value.var='yield')
 w <- mat1; k=15; n=8  # k=p gen, n=q env
 w <- sweep(w, 1, rowMeans(mat1, na.rm=TRUE))
 w <- sweep(w, 2, colMeans(mat1, na.rm=TRUE))
 w <- w + mean(mat1, na.rm=TRUE)
 w <- rowSums(w^2, na.rm=TRUE)
 sig2 <- k*w/((k-2)*(n-1)) - sum(w)/((k-1)*(k-2)*(n-1))
 round(sig2/10000,1)  # Genotypes in T1 are divided by 10000
```

## End(Not run)

agridat

Datasets from agricultural experiments

Description

This package contains datasets from publications relating to agriculture, including field crops, tree crops, animal studies, and a few others.

Details

If you use these data, please cite both the agridat package and the original source of the data.

Abbreviations in the 'other' column include: xy = coordinates, pls = partial least squares, rsm = response surface methodology, row-col = row-column design, ts = time series.

Uniformity trials with a single genotype
<table>
<thead>
<tr>
<th>name</th>
<th>dimensions</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>ansari.wheat.uniformity</td>
<td>96 x 8</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>baker.barley.uniformity</td>
<td>3 x 19</td>
<td>xy, 10 years</td>
<td></td>
</tr>
<tr>
<td>baker.strawberry.uniformity</td>
<td>10 x 50</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>baker.wheat.uniformity</td>
<td>12 x 12</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>bancroft.peanut.uniformity</td>
<td>6 x 18</td>
<td>xy, 2 blocks</td>
<td></td>
</tr>
<tr>
<td>batchelor.apple.uniformity</td>
<td>8 x 28</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>batchelor.lemon.uniformity</td>
<td>14 x 16</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>batchelor.navel1.uniformity</td>
<td>20 x 50</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>batchelor.navel2.uniformity</td>
<td>15 x 33</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>batchelor.valencia.uniformity</td>
<td>12 x 20</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>batchelor.walnut.uniformity</td>
<td>10 x 28</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>bose.multi.uniformity</td>
<td>15 x 26</td>
<td>xy, 3 years</td>
<td></td>
</tr>
<tr>
<td>bradley.multi.uniformity</td>
<td>10 x 11</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>bryan.corn.uniformity</td>
<td>6 x 48</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>christidis.cotton.uniformity</td>
<td>16 x 16</td>
<td>xy, 4 blocks</td>
<td></td>
</tr>
<tr>
<td>christidis.wheat.uniformity</td>
<td>12 x 24</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>dasilva.soybean.uniformity</td>
<td>24 x 48</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>davies.pasture.uniformity</td>
<td>19 x 40</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>day.wheat.uniformity</td>
<td>31 x 100</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>draper.safflower.uniformity</td>
<td>18 x 16</td>
<td>xy, 2 expt</td>
<td>smith</td>
</tr>
<tr>
<td>eden.tea.uniformity</td>
<td>12 x 12</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>garber.multi.uniformity</td>
<td>45 x 6</td>
<td>xy, 2 years/crops</td>
<td></td>
</tr>
<tr>
<td>gomez.rice.uniformity</td>
<td>18 x 36</td>
<td>xy</td>
<td>aov,smith</td>
</tr>
<tr>
<td>goulden.barley.uniformity</td>
<td>20 x 20</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>harris.multi.uniformity</td>
<td>2 x 23</td>
<td>xy, 23 crops</td>
<td>corrgram</td>
</tr>
<tr>
<td>hartman.tomato.uniformity</td>
<td>12 x 32</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>heath.raddish.uniformity</td>
<td>10 x 10</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>holtsmark.timothy.uniformity</td>
<td>6 x 40</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>iigue.sugarcane.uniformity</td>
<td>36 x 42</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>immer.sugarcane.uniformity</td>
<td>10 x 60</td>
<td>xy, 3 traits</td>
<td></td>
</tr>
<tr>
<td>iyer.wheat.uniformity</td>
<td>25 x 80</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>kadam.millet.uniformity</td>
<td>8 x 20</td>
<td>xy, 2 expts</td>
<td></td>
</tr>
<tr>
<td>kalamkar.potato.uniformity</td>
<td>6 x 96</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>kalamkar.wheat.uniformity</td>
<td>16 x 80</td>
<td>xy, 2 traits</td>
<td></td>
</tr>
<tr>
<td>kempton.barley.uniformity</td>
<td>7 x 28</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>kerr.sugarcane.uniformity</td>
<td>12 x 12</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>khin.rice.uniformity</td>
<td>30 x 36</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>kiesselbach.oats.uniformity</td>
<td>3 x 69</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>kristensen.barley.uniformity</td>
<td>22 x 11</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>kulkarni.sorghum.uniformity</td>
<td>4 x 40</td>
<td>xy, 3 years</td>
<td></td>
</tr>
<tr>
<td>lander.multi.uniformity</td>
<td>5 x 39</td>
<td>xy, 4 years</td>
<td></td>
</tr>
<tr>
<td>laycock.tea.uniformity</td>
<td>6 x 9</td>
<td>xy, 2 locations</td>
<td></td>
</tr>
<tr>
<td>lehmann.millet.uniformity</td>
<td>6 x 22</td>
<td>xy, 3 years</td>
<td></td>
</tr>
<tr>
<td>lessman.sorghum.uniformity</td>
<td>46 x 60</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>li.millet.uniformity</td>
<td>6 x 100</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>loesell.bean.uniformity</td>
<td>21 x 90</td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>lord.rice.uniformity</td>
<td>5 x 14</td>
<td>xy, 8 fields</td>
<td></td>
</tr>
</tbody>
</table>
love.cotton.uniformity  16 x 10  xy
lyon.potato.uniformity  34 x 6  xy
magistad.pineapple.uniformity  5 x 5  xy
masood.rice.uniformity  12 x 24  xy
mcclelland.corn.uniformity  2 x 219  xy
mercer.mangold.uniformity  10 x 20  xy
mercer.wheat.uniformity  25 x 20  xy, 2 traits  spplot
montgomery.wheat.uniformity  14 x 16  xy, 2 years  lm
moore.polebean.uniformity  12 x 12  xy
moore.bushbean.uniformity  24 x 24  xy
moore.sweetcorn.uniformity  24 x 12  xy
moore.carrot.uniformity  24 x 12  xy
moore.springcauliflower.uniformity  12 x 20  xy
moore.fallcauliflower.uniformity  12 x 20  xy
nagai.strawberry.uniformity  18 x 24  xy
nagai.strawberry.uniformity  18 x 24  xy
nair.turmeric.uniformity  72 x 12  xy
narain.sorghum.uniformity  10 x 16  xy
nongeke.peas.uniformity  15 x 18  xy, 2 traits
nongeke.sweetcorn.uniformity  32 x 18  xy, 3 loc
obsi.potato.uniformity  63 x 26, 49 x 19  xy
odland.soybean.uniformity  25 x 42  xy
odland.soyhay.uniformity  28 x 55  xy
panse.cotton.uniformity  32 x 40  xy
parker.orange.uniformity  10 x 27  xy, 6 yr
paez.coffee.uniformity  32 x 36  xy, 5 yr
piepho.barley.uniformity  30 x 36  xy
polson.safflower.uniformity  52 x 33  xy
robinson.peanut.uniformity  16 x 36  xy
sawyer.multi.uniformity  8 x 6  xy, 3 year
sayer.sugarcane.uniformity  8 x 136, 8 x 121  xy, 2 year
shafi.tomato.uniformity  10 x 20  xy
smith.beans.uniformity  18 x 12, 16 x 15  xy, 2 yr, 2 crops
smith.corn.uniformity  6 x 20  xy, 3 years  rgl
stephens.sorghum.uniformity  100 x 20  xy
stickler.sorghum.uniformity  20 x 20  xy, 4 expts, 2 years
strickland.apple.uniformity  11 x 18  xy
strickland.grape.uniformity  5 x 31  xy
strickland.peach.uniformity  8 x 18  xy
strickland.tomato.uniformity  30 x 6  xy
wassom.brome.uniformity  36 x 36  xy, 3 expts
wiebe.wheat.uniformity  12 x 125  xy  medianpolish, loess
wiedemann.safflower.uniformity  54 x 33  xy
smith.barley.uniformity  48 x 15  xy
williams.cotton.uniformity  24 x 12  xy
wyatt.multi.uniformity  43 x 3  xy
### Yield monitor

<table>
<thead>
<tr>
<th>name</th>
<th>reps</th>
<th>years</th>
<th>trt</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>gartner.corn</td>
<td>3</td>
<td>2</td>
<td></td>
<td>xy,ym</td>
<td>lm</td>
</tr>
<tr>
<td>lasrosas.corn</td>
<td>3</td>
<td>6</td>
<td></td>
<td>xy,ym</td>
<td>lm</td>
</tr>
<tr>
<td>kayad.alfalfa</td>
<td>4</td>
<td></td>
<td></td>
<td>xy,ym</td>
<td></td>
</tr>
</tbody>
</table>

### Animals

<table>
<thead>
<tr>
<th>name</th>
<th>gen</th>
<th>years</th>
<th>trt</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>alwan.lamb</td>
<td>34</td>
<td>2</td>
<td></td>
<td>ordinal</td>
<td>clmm</td>
</tr>
<tr>
<td>becker.chicken</td>
<td>5,12</td>
<td>5</td>
<td></td>
<td>heritability</td>
<td>lmer</td>
</tr>
<tr>
<td>crampton.pig</td>
<td>10</td>
<td>2</td>
<td></td>
<td>2</td>
<td>aov</td>
</tr>
<tr>
<td>depalluel.sheep</td>
<td>4</td>
<td>4</td>
<td></td>
<td>latin</td>
<td></td>
</tr>
<tr>
<td>diggle.cow</td>
<td>4</td>
<td></td>
<td></td>
<td>ts</td>
<td></td>
</tr>
<tr>
<td>foulley.calving</td>
<td></td>
<td></td>
<td></td>
<td>ordinal</td>
<td>polr</td>
</tr>
<tr>
<td>goulden.eggs</td>
<td></td>
<td></td>
<td></td>
<td>controlchart</td>
<td></td>
</tr>
<tr>
<td>harvey.lsmmeans</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
<td>lm</td>
</tr>
<tr>
<td>harville.lamb</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>lmer</td>
</tr>
<tr>
<td>henderson.milkfat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nls,lm,glm,gam</td>
</tr>
<tr>
<td>holland.arthropods</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ilri.sheep</td>
<td>4</td>
<td>6</td>
<td></td>
<td>diallel</td>
<td>lmer, asreml</td>
</tr>
<tr>
<td>kenward.cattle</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>asreml</td>
</tr>
<tr>
<td>lucas.switchback</td>
<td>12</td>
<td>3</td>
<td></td>
<td></td>
<td>aov</td>
</tr>
<tr>
<td>mead.lamb</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td>glm</td>
</tr>
<tr>
<td>patterson.switchback</td>
<td>12</td>
<td>4</td>
<td></td>
<td></td>
<td>aov</td>
</tr>
<tr>
<td>urquhart.feedlot</td>
<td>11</td>
<td>3</td>
<td></td>
<td></td>
<td>lm</td>
</tr>
<tr>
<td>woodman.pig</td>
<td></td>
<td>3</td>
<td></td>
<td>cov</td>
<td>lm</td>
</tr>
<tr>
<td>zuidhof.broiler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ts</td>
</tr>
</tbody>
</table>

### Trees

<table>
<thead>
<tr>
<th>name</th>
<th>gen</th>
<th>loc</th>
<th>reps</th>
<th>years</th>
<th>trt</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>box.cork</td>
<td></td>
<td>4</td>
<td>3,3</td>
<td></td>
<td></td>
<td>xy,graeco</td>
<td>aov</td>
</tr>
<tr>
<td>devries.pine</td>
<td></td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td>repeated</td>
<td>asreml</td>
</tr>
<tr>
<td>harris.wateruse</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>hanover.whitepine</td>
<td>7*4</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td>repeated</td>
<td>asreml,lme</td>
</tr>
<tr>
<td>johnson.douglasfir</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>heritability</td>
<td></td>
</tr>
<tr>
<td>lavoranti.eucalyptus</td>
<td>70</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>xy</td>
<td></td>
</tr>
<tr>
<td>pearce.apple</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>cov</td>
<td></td>
</tr>
<tr>
<td>williams.trees</td>
<td>37</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Field and horticulture crops

<table>
<thead>
<tr>
<th>name</th>
<th>gen</th>
<th>loc</th>
<th>reps</th>
<th>years</th>
<th>trt</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>acorsi.grayleafspot</td>
<td>36</td>
<td>9</td>
<td>2</td>
<td>5</td>
<td></td>
<td>nonnormal</td>
<td>gnm,anmi</td>
</tr>
<tr>
<td>Dataset</td>
<td>Subjects</td>
<td>Gen*Gen</td>
<td>Gen*Trt</td>
<td>Gen*Year</td>
<td>Year*Year</td>
<td>Year*Gen</td>
<td>Year*Trt</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>adugna.sorghum</td>
<td>28</td>
<td>13</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aastveit.barley</td>
<td>15</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>allcroft.lodging</td>
<td>32</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>archbold.apple</td>
<td>2</td>
<td>5</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ars.earlywhitecorn96</td>
<td>60</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>australia.soybean</td>
<td>58</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>barrero.maize</td>
<td>847</td>
<td>16</td>
<td>4</td>
<td>11</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>battese.survey</td>
<td>12</td>
<td>1-5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>beall.webworms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>beaven.barley</td>
<td>8</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.bayesian</td>
<td>75</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.beans</td>
<td>6</td>
<td>4*6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.checks</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.elbatan</td>
<td>50</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.endive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.met</td>
<td>64</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>besag.triticale</td>
<td>3</td>
<td></td>
<td></td>
<td>2,2,3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bliss.borers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>blackman.wheat</td>
<td>12</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond.diallel</td>
<td>6*6</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bridges.cucumber</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>brandle.rape</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>buntaran.wheat</td>
<td>30</td>
<td>18</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>burgueno.alpha</td>
<td>15</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>burgueno.rowcol</td>
<td>64</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>burgueno.unreplicated</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>butron.maize</td>
<td>49</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>caribbean.maize</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>carmer.density</td>
<td>8</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>carlson.germination</td>
<td>15</td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chinloy.fractionalfactorial</td>
<td></td>
<td></td>
<td></td>
<td>1/3 3^5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>christidis.competition</td>
<td>9</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.beets</td>
<td>6</td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.bib</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.crd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.factorial</td>
<td>2</td>
<td></td>
<td>4*2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.latin</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.lattice</td>
<td>5</td>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.wireworms</td>
<td>5</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cochrane.eelworms</td>
<td>4</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>connolly.potato</td>
<td>20</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cornelius.maize</td>
<td>9</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>corsten.interaction</td>
<td>20</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cramer.cucumber</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>crossa.wheat</td>
<td>18</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>crowder.seeds</td>
<td>2</td>
<td>21</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cox.stripssplit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3,4,2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cullis.earlygen</td>
<td>532</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dataset</td>
<td>Columns</td>
<td>Rows</td>
<td>Equation</td>
<td>Package</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>------</td>
<td>----------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>damaesa.maize</td>
<td>22</td>
<td>4</td>
<td>xy,incblock,two-stage</td>
<td>asreml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dasilva.maize</td>
<td>55</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>darwin.maize</td>
<td>12</td>
<td>2</td>
<td></td>
<td>t.test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>davidian.soaybean</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>denis.missing</td>
<td>5</td>
<td>26</td>
<td></td>
<td>lme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>denis.ryegrass</td>
<td>21</td>
<td>7</td>
<td></td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>digby.jointregression</td>
<td>10</td>
<td>17</td>
<td></td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>durban.competition</td>
<td>36</td>
<td>3</td>
<td>xy, competition</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>durban.rowcol</td>
<td>272</td>
<td>2</td>
<td>xy</td>
<td>lm, gam, asreml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>durban.splitplot</td>
<td>70</td>
<td>4</td>
<td>xy</td>
<td>lm, gam, asreml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eden.potato</td>
<td>4</td>
<td>3</td>
<td>4-12</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eden.nonnormal</td>
<td>4</td>
<td>4</td>
<td></td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edwards.oats</td>
<td>80</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>engelstad.nitro</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>rsm1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fan.stability</td>
<td>13</td>
<td>10</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>federer.diagcheck</td>
<td>122</td>
<td></td>
<td>xy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>federer.tobacco</td>
<td>8</td>
<td>7</td>
<td>xy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fisher.barley</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fisher.latin</td>
<td>5</td>
<td>5</td>
<td>xy,latin</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fox.wheat</td>
<td>22</td>
<td>14</td>
<td></td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gathmann.bt</td>
<td>2</td>
<td>8</td>
<td></td>
<td>tost</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gauch.soy</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>ammi</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>george.wheat</td>
<td>211</td>
<td>9</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>giles.wheat</td>
<td>19</td>
<td>13</td>
<td>2 traits</td>
<td>gnm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gilmour.serpentine</td>
<td>108</td>
<td>3</td>
<td>xy, serpentine</td>
<td>asreml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gilmour.slatehall</td>
<td>25</td>
<td>6</td>
<td>xy</td>
<td>asreml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.fractionalfactorial</td>
<td>2</td>
<td>1/2</td>
<td>2^6</td>
<td>xy</td>
<td>lm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.groupsplit</td>
<td>45</td>
<td>3</td>
<td>xy, 3 gen groups</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.heteroskedastic</td>
<td>35</td>
<td>3</td>
<td>hetero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.multilocsplitplot</td>
<td>2</td>
<td>3</td>
<td>rsm1,nitro</td>
<td>aov, lmer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.nitrogen</td>
<td>4</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.nonnormal1</td>
<td>4</td>
<td>9</td>
<td>log10</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.nonnormal2</td>
<td>14</td>
<td>3</td>
<td>sqrt</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.nonnormal3</td>
<td>12</td>
<td>3</td>
<td>arcsin</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.seedrate</td>
<td>4</td>
<td>6</td>
<td>rate</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.splitplot.subsample</td>
<td>3</td>
<td>8,4</td>
<td>subsample</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.splitplot</td>
<td>3</td>
<td>3</td>
<td>xy, nitro, mgmt</td>
<td>aov, lmer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.stripplot</td>
<td>6</td>
<td>3</td>
<td>xy, nitro</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.stripsplitplot</td>
<td>6</td>
<td>3</td>
<td>xy, nitro</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gomez.wetdry</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>lmer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gotway.hessianfly</td>
<td>16</td>
<td>4</td>
<td>xy</td>
<td>lmer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>goulden.latin</td>
<td>5</td>
<td>5</td>
<td>xy, latin</td>
<td>lm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>goulden.splitplot</td>
<td>2</td>
<td>4</td>
<td>2*5</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>graybill.heteroskedastic</td>
<td>4</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gregory.cotton</td>
<td>2</td>
<td>4<em>3</em>2*2</td>
<td>polar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grover.diallel</td>
<td>4</td>
<td>6*6</td>
<td>diallel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grover.rcb.subsample</td>
<td>4</td>
<td>2</td>
<td>subsample</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gumpertz.pepper</td>
<td>2</td>
<td>9</td>
<td>xy</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment</td>
<td>N</td>
<td>Trials</td>
<td>Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hadasch.lettuce</td>
<td>89</td>
<td>3</td>
<td>markers asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hanks.sprinkler</td>
<td>3</td>
<td>3</td>
<td>xy asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hayman.tobacco</td>
<td>8</td>
<td>2</td>
<td>diallel asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hazell.vegetables</td>
<td>4</td>
<td>6</td>
<td>linprog</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heady.fertilizer</td>
<td>2</td>
<td>9*9</td>
<td>rsm2 lm, rgl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hernandez.nitrogen</td>
<td>5</td>
<td>4</td>
<td>rsm1 lm, nls asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hildebrand.systems</td>
<td>14</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>holshouser.splitstrip</td>
<td>4</td>
<td>4</td>
<td>2*4 rsm1.pop lmer huehn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hughes.wheat</td>
<td>20</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hunter.corn</td>
<td>12</td>
<td>3</td>
<td>6 binomial lmer, aod, glmm xyplot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ivins.herbs</td>
<td>13</td>
<td>6</td>
<td>2 traits lm, friedman</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jansen.apple</td>
<td>3</td>
<td>4</td>
<td>3 binomial glmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jansen.carrot</td>
<td>16</td>
<td>3</td>
<td>2 binomial glmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jansen.strawberry</td>
<td>12</td>
<td>4</td>
<td>ordinal mosaicplot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jayaraman.bamboo</td>
<td>6</td>
<td>2</td>
<td>3 heritability lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jenkyn.mildew</td>
<td>9</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>john.alpha</td>
<td>24</td>
<td>3</td>
<td>xy, alpha lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>johnson.blight</td>
<td>17</td>
<td>4</td>
<td>2 logistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kang.peanut</td>
<td>10</td>
<td>15</td>
<td>4 gge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>karcher.turfgrass</td>
<td>4</td>
<td>2,4</td>
<td>ordinal polr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>keen.potatodamage</td>
<td>6</td>
<td>4</td>
<td>2,3,8 ordinal mosaicplot, clmm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kempton.competition</td>
<td>36</td>
<td>3</td>
<td>xy, competition lme AR1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kempton.rowcol</td>
<td>35</td>
<td>2</td>
<td>xy, row-col lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kempton.slatehall</td>
<td>25</td>
<td>6</td>
<td>xy asreml, lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lee.potatoblight</td>
<td>337</td>
<td>4</td>
<td>11 xy, ordinal, repeated metafor, lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lehner.soybeanmold</td>
<td>35</td>
<td>4</td>
<td>11 qq medpolish, huehn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lillemo.wheat</td>
<td>24</td>
<td>13</td>
<td>7 superiority</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lin.superiority</td>
<td>33</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lin.unbalanced</td>
<td>33</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>linder.wheat</td>
<td>9</td>
<td>7</td>
<td>4 gge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>little.splitblock</td>
<td>4</td>
<td>4,5</td>
<td>xy, split-block aov</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lonnquist.maize</td>
<td>11</td>
<td>4</td>
<td>diallel asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lyons.wheat</td>
<td>12</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lu.stability</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>meconway.turnip</td>
<td>2</td>
<td>4</td>
<td>2,4 hetero aov, lme</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mcleod.barley</td>
<td>8</td>
<td>6</td>
<td>aggregate glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mead.cauliflower</td>
<td>3</td>
<td>4</td>
<td>2 poisson</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mead.cowpea.maize</td>
<td>3,2</td>
<td>4</td>
<td>3 intercrop</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mead.germination</td>
<td>4</td>
<td>4,4</td>
<td>binomial glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mead.strawberry</td>
<td>8</td>
<td>4</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mead.turnip</td>
<td>3</td>
<td>5,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miguez.biomass</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minnesota.barley.weather</td>
<td>6</td>
<td>10</td>
<td>dotplot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minnesota.barley.yield</td>
<td>22</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>omer.sorghum</td>
<td>18</td>
<td>6</td>
<td>4 jags</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>onofri.winterwheat</td>
<td>8</td>
<td>3</td>
<td>7 ammi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Treatments</td>
<td>Software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ortiz.tomato</td>
<td>15 18 16 env<em>gen-env</em>gen*cov</td>
<td>pls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pacheco.soybean</td>
<td>18 11</td>
<td>ammi, nlm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pederson.lettuce.repeated</td>
<td>18 3</td>
<td>nlme, asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>perry.springwheat</td>
<td>28 5 4 gain</td>
<td>lm, lmer, asreml</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pedersen.sorghum.cowpea</td>
<td>2 4 7 4 intercrop</td>
<td>mumm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>piepho.cocksfoot</td>
<td>25 7</td>
<td>nlme SSfpl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ratkowsky.onions</td>
<td>4 3 21</td>
<td>zeroinfl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reid.grasses</td>
<td>30 2,4 zip</td>
<td>betareg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ridout.appleshoots</td>
<td>rothamsted.brussels 4 6</td>
<td>biplot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rothamsted.oats</td>
<td>8 9 rcb</td>
<td>aov, lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ryder.groundnut</td>
<td>5 4 xy, rcb</td>
<td>glm, poisson</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>salmon.bunt</td>
<td>10 2 20</td>
<td>glm, mitzerlich</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>senshu.rice</td>
<td>40</td>
<td>glm/gnm, gammi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shafii.rapeseed</td>
<td>6 14 3 3</td>
<td>lmer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shaw.oats</td>
<td>13 2 5 3</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>silva.cotton</td>
<td>5 5 5 traits</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sinclair.clover</td>
<td>4 4 4</td>
<td>glm/gnm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>snijders.asparagus</td>
<td>17 3 4 percent</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>steptoe.morex.pheno</td>
<td>152 16 10 traits</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>steptoe.morex.gen</td>
<td>150 223 markers, qtl</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>streibig.competition</td>
<td>2 3</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stroup.nin</td>
<td>56 4</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stroup.splitplot</td>
<td>4</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>student.barley</td>
<td>2 51 6</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tai.potato</td>
<td>8 3 2</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>talbot.barley</td>
<td>9 12 gen<em>env-gen</em>gen*trt</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>theobald.barley</td>
<td>3 5 2 5</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>theobald.covariate</td>
<td>10 7 5</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>thompson.comsonoy</td>
<td>5 33 repeated measures</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vaneeuwijk.fusarium</td>
<td>20 4 7 3-way</td>
<td>jags</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vaneeuwijk.drymatter</td>
<td>6 4 7</td>
<td>aov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vaneeuwijk.nematodes</td>
<td>11 nonnormal, poisson</td>
<td>glm, gammi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vargas.wheat1</td>
<td>7 6 gen<em>yr-gen</em>trt, yr<em>gen-yr</em>cov</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vargas.wheat2</td>
<td>8 7</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vargas.txr</td>
<td>10 24 yr<em>trt-yr</em>cov</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>verbyla.lupin</td>
<td>9 8 3 2 7 rsm1, xy, density</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vold.longterm</td>
<td>19 4 rsm1</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vsn.lupin3</td>
<td>336 3</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wedderburn.barley</td>
<td>10 9 percent</td>
<td>glm/gnm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weiss.incblock</td>
<td>31 6 xy, incblock</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weiss.lattice</td>
<td>49 4</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>welch.bermudagrass</td>
<td>4,4,4 rsm3, factorial</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wheatley.carrot</td>
<td>3 11</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yan.winterwheat</td>
<td>18 9</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yang.barley</td>
<td>6 18</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yates.missing</td>
<td>10 3^2 factorial</td>
<td>glm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>pca</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Time series**

<table>
<thead>
<tr>
<th>name</th>
<th>years</th>
<th>trt</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>byers.apple</td>
<td>74</td>
<td>17</td>
<td></td>
<td>lme</td>
</tr>
<tr>
<td>broadbalk.wheat</td>
<td>30</td>
<td></td>
<td>temp,precip</td>
<td></td>
</tr>
<tr>
<td>hessling.argentina</td>
<td>4</td>
<td>5</td>
<td>plant growth</td>
<td></td>
</tr>
<tr>
<td>kreusler.maize</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lambert.soiltemp</td>
<td>146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.barley</td>
<td>146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.corn</td>
<td>146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.cotton</td>
<td>146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.hay</td>
<td>104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.sorghum</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.wheat</td>
<td>146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.rice</td>
<td>117</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nass.soybean</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>walsh.cottonprice</td>
<td>34</td>
<td></td>
<td>cor</td>
<td></td>
</tr>
</tbody>
</table>

**Other**

<table>
<thead>
<tr>
<th>name</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>cate.potassium</td>
<td>cate-nelson</td>
</tr>
<tr>
<td>cleveland.soil</td>
<td>loess 2D</td>
</tr>
<tr>
<td>harrison.priors</td>
<td>nls, prior</td>
</tr>
<tr>
<td>nebraska.farmincome</td>
<td>choropleth</td>
</tr>
<tr>
<td>pearl.kernels</td>
<td>chisq</td>
</tr>
<tr>
<td>stirret.borers</td>
<td>lm, 4 trt</td>
</tr>
<tr>
<td>turner.herbicide</td>
<td>glm, 4 trt</td>
</tr>
<tr>
<td>usgs.herbicides</td>
<td>non-detect</td>
</tr>
<tr>
<td>wallace.iowaland</td>
<td>lm, choropleth</td>
</tr>
<tr>
<td>waynick.soil</td>
<td>spatial, nitro/carbon</td>
</tr>
</tbody>
</table>

**Summaries:**

Multi-environment trials with multi-genotype,loc,rep,year:

<table>
<thead>
<tr>
<th>name</th>
<th>gen</th>
<th>loc</th>
<th>reps</th>
<th>years</th>
<th>trt</th>
<th>other</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>barrero.maize</td>
<td>847</td>
<td>16</td>
<td>4</td>
<td>11</td>
<td>6</td>
<td></td>
<td>asreml</td>
</tr>
<tr>
<td>edwards.oats</td>
<td>80</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gauch.soay</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>12</td>
<td></td>
<td></td>
<td>ammi</td>
</tr>
<tr>
<td>george.wheat</td>
<td>211</td>
<td>9</td>
<td>4</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shafii.rapeseed</td>
<td>6</td>
<td>14</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td>biplot</td>
</tr>
<tr>
<td>shaw.oats</td>
<td>13</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td>aov</td>
</tr>
<tr>
<td>verbyla.lupin</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td></td>
<td>asreml</td>
</tr>
</tbody>
</table>

Data with markers: hadasch.lettuce.markers, steptoe.morex.geno
Data with pedigree: butron.maize

Author(s)
Kevin Wright, with support from many people who granted permission to include their data in this package.

References

---

**allcroft.lodging**  
*Multi-environment trial of cereal with lodging data*

Description
Percent lodging is given for 32 genotypes at 7 environments.

Format
A data frame with 224 observations on the following 3 variables.

- `env` environment, 1-7
- `gen` genotype, 1-32
- `y` percent lodged

Details
This data is for the first year of a three-year study.  
Used with permission of Chris Glasbey.

Source

Examples
```r
## Not run:
library(agridat)
data(allcroft.lodging)
dat <- allcroft.lodging

# Transformation
dat$sy <- sqrt(dat$y)
# Variety 4 has no lodging anywhere, so add a small amount
```

dat[dat$env=='E5' & dat$gen=='G04',]$sy <- .01

libs(lattice)
dotplot(env~y|gen, dat, as.table=TRUE,
       xlab="Percent lodged (by genotype)", ylab="Variety",
       main="allcroft.lodging")

# Tobit model
libs(AER)
m3 <- tobit(sy ~ 1 + gen + env, left=0, right=100, data=dat)

# Table 2 trial/variety means
preds <- expand.grid(gen=levels(dat$gen), env=levels(dat$env))
preds$pred <- predict(m3, newdata=preds)
round(tapply(preds$pred, preds$gen, mean),2)
round(tapply(preds$pred, preds$env, mean),2)

## End(Not run)

---

### alwan.lamb

For the 34 sheep sires, the number of lambs in each of 5 foot shape classes.

#### Description

For the 34 sheep sires, the number of lambs in each of 5 foot shape classes.

#### Usage

data("alwan.lamb")

#### Format

A data frame with 340 observations on the following 11 variables.

- **year**: numeric 1980/1981
- **breed**: breed PP, BRP, BR
- **sex**: sex of lamb M/F
- **sire0**: sire ID according to Alwan
- **shape**: sire ID according to Gilmour
- **count**: number of lambs
- **sire**: shape of foot
- **yr**: numeric contrast for year
- **b1**: numeric contrast for breeds
- **b2**: numeric contrast for breeds
- **b3**: numeric contrast for breeds
Details

There were 2513 lambs classified on the presence of deformities in their feet. The lambs represent the offspring of 34 sires, 5 strains, 2 years.

The variables yr, b1, b2, b3 are numeric contrasts for the fixed effects as defined in the paper by Gilmour (1987) and used in the SAS example. Gilmour does not explain the reason for the particular contrasts. The counts for classes LF1, LF2, LF3 were combined.

Source


References


Examples

```r
## Not run:
library(agridat)
data(alwan.lamb)
dat <- alwan.lamb

# merge LF1 LF2 LF3 class counts, and combine M/F
dat$shape <- as.character(dat$shape)
dat$shape <- ifelse(dat$shape=="LF2", "LF3", dat$shape)
dat$shape <- ifelse(dat$shape=="LF1", "LF3", dat$shape)
dat <- aggregate(count ~ year+breed+sire0+sire+shape+yr+b1+b2+b3, dat, FUN=sum)

dat <- transform(dat, year=factor(year), breed=factor(breed), sire0=factor(sire0), sire=factor(sire))

# LF5 or LF3 first is a bit arbitrary...affects the sign of the coefficients
dat <- transform(dat, shape=ordered(shape, levels=c("LF5","LF4","LF3")))

# View counts by year and breed
library(latticeExtra)
dat2 <- aggregate(count ~ year+breed+shape, dat, FUN=sum)
useOuterStrips(barchart(count ~ shape|year*breed, data=dat2, main="alwan.lamb"))

# Model used by Gilmour and SAS
dat <- subset(dat, count > 0)
```
alwan.lamb

27

libs(ordinal)
m1 <- clmm(shape ~ yr + b1 + b2 + b3 + (1|sire), data=dat,
weights=count, link="probit", Hess=TRUE)
summary(m1) # Very similar to Gilmour results
ordinal::ranef(m1) # sign is opposite of SAS
##
##
##
##
##
##
##
##

SAS var of sires .04849
Effect Shape Estimate Standard Error DF t Value
Intercept 1 0.3781 0.04907 29 7.71 <.0001
Intercept 2 1.6435 0.05930 29 27.72 <.0001
yr
0.1422 0.04834 2478 2.94 0.0033
b1
0.3781 0.07154 2478 5.28 <.0001
b2
0.3157 0.09709 2478 3.25 0.0012
b3
-0.09887 0.06508 2478 -1.52 0.1289

##
##
##
##
##
##
##

Gilmour results
Int1
.370 +/Int2 1.603 +/Year -.139 +/B1
-.370 +/B2
-.304 +/B3
.098 +/-

Pr > |t|

for probit analysis
.052
.061
.052
.076
.103
.070

# Plot random sire effects with intervals, similar to SAS example
plot.random <- function(model, random.effect, ylim=NULL, xlab="", main="") {
tab <- ordinal::ranef(model)[[random.effect]]
tab <- data.frame(lab=rownames(tab), est=tab$"(Intercept)")
tab <- transform(tab,
lo = est - 1.96 * sqrt(model$condVar),
hi = est + 1.96 * sqrt(model$condVar))
# sort by est, and return index
ix <- order(tab$est)
tab <- tab[ix,]
if(is.null(ylim)) ylim <- range(c(tab$lo, tab$hi))
n <- nrow(tab)
plot(1:n, tab$est, axes=FALSE, ylim=ylim, xlab=xlab,
ylab="effect", main=main, type="n")
text(1:n, tab$est, labels=substring(tab$lab,2) , cex=.75)
axis(1)
axis(2)
segments(1:n, tab$lo, 1:n, tab$hi, col="gray30")
abline(h=c(-.5, -.25, 0, .25, .5), col="gray")
return(ix)

}
ix <- plot.random(m1, "sire")

# foot-shape proportions for each sire, sorted by estimated sire effects
# positive sire effects tend to have lower proportion of lambs in LF4 and LF5
tab <- prop.table(xtabs(count ~ sire+shape, dat), margin=1)
tab <- tab[ix,]
tab <- tab[nrow(tab):1,] # reverse the order
lattice::barchart(tab,


ansari.wheat.uniformity

Uniformity trial of wheat in India in 1940.

Usage

data("ansari.wheat.uniformity")

Format

A data frame with 768 observations on the following 3 variables.

row row
col column
yield yield of grain per plot, in half-ounces

Details

An experiment was conducted at the Government Research Farm, Raya (Muttra District), during the rainy season of 1939-40.

"Wheat was sown over an area of 180 ft. x 243 ft. with 324 rows on a field of average fertility. It had wheat during 1938-39 rabi and was fallow during 1939-40 kharif. The seed was sown behind desi plough in rows 9 in. apart, the length of each row being 180 ft".

"At the time of harvest, 18 rows on both sides and 10 ft. at the end of the field were discarded to eliminate border effects and an area of 160 ft. x 216 ft. with 288 rows was harvested in small units, each being 2 ft. 3 in. broad with three rows 20 ft. long. There were 96 units across the rows and eight units along the rows. The total number of unit plots thus obtained was 768. The yield of grain for each unit plot was weighed and recorded separately and is given in the appendix."

Field width: 96 plots * 2.25 feet = 216 feet.
Field length: 8 plots * 20 feet = 160 feet.

Comment: There seems to be a strong cyclical pattern to the fertility gradient. "History of the field reveals no explanation for this phenomenon, as an average field usually found on the farm was selected for the trial."
Source

References
None

Examples
```r
## Not run:

library(agridat)
data(ansari.wheat.uniformity)
dat <- ansari.wheat.uniformity

# match Ansari figure 3
libs(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE, aspect=216/160, # true aspect
       main="ansari.wheat.uniformity")

## End(Not run)
```

---

**arankacami.groundnut.uniformity**

*Uniformity trial of groundnut*

**Description**

Uniformity trial of groundnut

**Usage**

data("arankacami.groundnut.uniformity")

**Format**

A data frame with 96 observations on the following 3 variables.

- row row
- col column
- yield yield, kg/plot

**Details**

Basic plot size is 0.75 m (rows) x 4 m (columns).
Source


References

None

Examples

```r
## Not run:
library(agridat)
data(arankacami.groundnut.uniformity)
dat <- arankacami.groundnut.uniformity

require(desplot)
desplot(yield ~ col*row, dat,
       flip=TRUE, aspect=(12*.75)/(8*1),
       main="arankacami.groundnut.uniformity")

## End(Not run)
```

archbold.apple

*Split-split plot experiment of apple trees*

Description

Split-split plot experiment of apple trees with different spacing, root stock, and cultivars.

Format

A data frame with 120 observations on the following 10 variables.

rep  block, 5 levels
row  row
pos  position within each row
spacing  spacing between trees, 6,10,14 feet
stock  rootstock, 4 levels
gen  genotype, 2 levels
yield  yield total, kg/tree from 1975-1979
trt  treatment code
Details

In rep 1, the 10-foot-spacing main plot was split into two non-contiguous pieces. This also happened in rep 4. In the analysis of Cornelius and Archbold, they consider each row x within-row-spacing to be a distinct main plot. (Also true for the 14-foot row-spacing, even though the 14-foot spacing plots were contiguous.)

The treatment code is defined as 100 * spacing + 10 * stock + gen, where stock=0,1,6,7 for Seedling, MM111, MM106, M0007 and gen=1,2 for Redspur, Golden, respectively.

Source


References


Examples

```r
## Not run:
library(agridat)
data(archbold.apple)
dat <- archbold.apple

# Define main plot and subplot
dat <- transform(dat, rep=factor(rep), spacing=factor(spacing), trt=factor(trt),
                 mp = factor(paste(row,spacing,sep="")),
                 sp = factor(paste(row,spacing,stock,sep="")))

# Due to 'spacing', the plots are different sizes, but the following layout
# shows the relative position of the plots and treatments. Note that the
# 'spacing' treatments are not contiguous in some reps.
libs(desplot)
  desplot(dat, spacing~row*pos,
          col=stock, cex=1, num=gen, # aspect unknown
          main="archbold.apple")

libs(lme4, lucid)
m1 <- lmer(yield ~ -1 + trt + (1|rep/mp/sp), dat)
vc(m1) # Variances/means on Cornelius, page 59
### grp  var1  var2  vcov  sdcor
### sp:(mp:rep) (Intercept) <NA>  193.3 13.9
### mp:rep (Intercept) <NA>  203.8 14.28
### rep (Intercept) <NA>  197.3 14.85
### Residual <NA> <NA> 1015 31.86
```
Multi-environment trial of early white food corn

Description

Multi-environment trial of early white food corn for 60 white hybrids.

Format

A data frame with 540 observations on the following 9 variables.

- loc location, 9 levels
- gen gen, 60 levels
- yield yield, bu/ac
- stand stand, percent
- rootlodge root lodging, percent
- stalklodge stalk lodging, percent
- earht ear height, inches
- flower days to flower
- moisture moisture, percent

Details

Data are the average of 3 replications.

Yields were measured for each plot and converted to bushels / acre and adjusted to 15.5 percent moisture.

Stand is expressed as a percentage of the optimum plant stand.

Lodging is expressed as a percentage of the total plants for each hybrid.

Ear height was measured from soil level to the top ear leaf collar. Heights are expressed in inches.

Days to flowering is the number of days from planting to mid-tassel or mid-silk.

Moisture of the grain was measured at harvest.

Source

Examples

```r
## Not run:

library(agridat)

data(ars.earlywhitecorn96)
dat <- ars.earlywhitecorn96

libs(lattice)
# These views emphasize differences between locations
dotplot(gen~yield, dat, group=loc, auto.key=list(columns=3),
main="ars.earlywhitecorn96")
## dotplot(gen~stalklodge, dat, group=loc, auto.key=list(columns=3),
## main="ars.earlywhitecorn96")
splom(~dat[,3:9], group=dat$loc, auto.key=list(columns=3),
main="ars.earlywhitecorn96")

# MANOVA
m1 <- manova(cbind(yield,earht,moisture) ~ gen + loc, dat)
m1
summary(m1)

## End(Not run)
```

australia.soybean

Multi-environment trial of soybean in Australia

Description

Yield and other traits of 58 varieties of soybeans, grown in four locations across two years in Australia. This is four-way data of Year x Loc x Gen x Trait.

Format

A data frame with 464 observations on the following 10 variables.

- `env` environment, 8 levels, first character of location and last two characters of year
- `loc` location
- `year` year
- `gen` genotype of soybeans, 1-58
- `yield` yield, metric tons / hectare
- `height` height (meters)
- `lodging` lodging
- `size` seed size, (millimeters)
- `protein` protein (percentage)
- `oil` oil (percentage)
Details

Measurement are available from four locations in Queensland, Australia in two consecutive years 1970, 1971.

The 58 different genotypes of soybeans consisted of 43 lines (40 local Australian selections from a cross, their two parents, and one other which was used a parent in earlier trials) and 15 other lines of which 12 were from the US.

Lines 1-40 were local Australian selections from Mamloxi (CPI 172) and Avoyelles (CPI 15939).

<table>
<thead>
<tr>
<th>No.</th>
<th>Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-40</td>
<td>Local selections</td>
</tr>
<tr>
<td>41</td>
<td>Avoyelles (CPI 15939) Tanzania</td>
</tr>
<tr>
<td>42</td>
<td>Hernon 49 (CPI 15948) Tanzania</td>
</tr>
<tr>
<td>43</td>
<td>Mamloxi (CPI 172) Nigeria</td>
</tr>
<tr>
<td>44</td>
<td>Dorman USA</td>
</tr>
<tr>
<td>45</td>
<td>Hampton USA</td>
</tr>
<tr>
<td>46</td>
<td>Hill USA</td>
</tr>
<tr>
<td>47</td>
<td>Jackson USA</td>
</tr>
<tr>
<td>48</td>
<td>Leslie USA</td>
</tr>
<tr>
<td>49</td>
<td>Semstar Australia</td>
</tr>
<tr>
<td>50</td>
<td>Wills USA</td>
</tr>
<tr>
<td>51</td>
<td>C26673 Morocco</td>
</tr>
<tr>
<td>52</td>
<td>C26671 Morocco</td>
</tr>
<tr>
<td>53</td>
<td>Bragg USA</td>
</tr>
<tr>
<td>54</td>
<td>Delmar USA</td>
</tr>
<tr>
<td>55</td>
<td>Lee USA</td>
</tr>
<tr>
<td>56</td>
<td>Hood USA</td>
</tr>
<tr>
<td>57</td>
<td>Ogden USA</td>
</tr>
<tr>
<td>58</td>
<td>Wayne USA</td>
</tr>
</tbody>
</table>

Note on the data in Basford and Tukey book. The values for line 58 for Nambour 1970 and Redland Bay 1971 are incorrectly listed on page 477 as 20.490 and 15.070. They should be 17.350 and 13.000, respectively. In the data set made available here, these values have been corrected.

Used with permission of Kaye Basford, Pieter Kroonenberg.

Source


Retrieved from: https://three-mode.leidenuniv.nl/data/soybeaninf.htm

References


Examples

```r
## Not run:
library(agridat)
data(australia.soybean)
dat <- australia.soybean

libs(reshape2)
dm <- melt(dat, id.var=c('env', 'year','loc','gen'))

# Joint plot of genotypes & traits. Similar to Figure 1 of Kroonenberg 1989
dmat <- acast(dm, gen~variable, fun=mean)
dmat <- scale(dmat)
bipplot(princomp(dmat), main="australia.soybean trait x gen biplot", cex=.75)

# Figure 1 of Kozak 2010, lines 44-58
libs(reshape2, lattice, latticeExtra)
data(australia.soybean)
dat <- australia.soybean
dat <- melt(dat, id.var=c('env', 'year','loc','gen'))
dat <- acast(dat, gen~variable, fun=mean)
dat <- scale(dat)
dat <- as.data.frame(dat)[,c(2:6,1)]
dat$gen <- rownames(dat)
# data for the graphic by Kozak
dat2 <- dat[44:58,

dat3 <- subset(dat2, is.element(gen, c("G48","G49","G50","G51")))

parallelplot(~ dat3[,1:6]|dat3$gen, main="australia.soybean",
as.table=TRUE, horiz=FALSE) +
parallelplot(~ dat2[,1:6], horiz=FALSE, col="gray80") +
parallelplot(~ dat3[,1:6]|dat3$gen,
as.table=TRUE, horiz=FALSE, lwd=2)

## End(Not run)
```

baker.barley.uniformity

Uniformity trials of barley, 10 years on same ground

Description

Uniformity trials of barley at Davis, California, 1925-1935, 10 years on same ground.
**Format**

A data frame with 570 observations on the following 4 variables.

- `row` row
- `col` column
- `year` year
- `yield` yield, pounds/acre

**Details**

Ten years of uniformity trials were sown on the same ground. Baker (1952) shows a map of the field, which gravel subsoil extended from the upper right corner diagonally lower-left. This results in lower yields as seen on the 10-year average map.

- Field width: 19 plots = 827 ft
- Field length: 3 plots * 161 ft + 2 alleys * 15 feet = 513 ft

**Source**


**Examples**

```r
## Not run:

library(agridat)
data(baker.barley.uniformity)
dat <- baker.barley.uniformity

# Ten-year average
dat2 <- aggregate(yield ~ row*col, data=dat, FUN=mean, na.rm=TRUE)

lib(desplot)
desplot(dat, yield~col*row|year,
    aspect = 513/827, # true aspect
    main="baker.barley.uniformity - heatmaps by year")

desplot(dat2, yield~col*row,
    aspect = 513/827, # true aspect
    main="baker.barley.uniformity - heatmap of 10-year average")

# Note low yield in upper right, slanting to left a bit due to sandy soil
# as shown in Baker figure 1.

# Baker fig 2, stdev vs mean
dat3 <- aggregate(yield ~ row*col, data=dat, FUN=sd, na.rm=TRUE)
plot(dat2$yield, dat3$yield, xlab="Mean yield", ylab="Std Dev yield",
     main="baker.barley.uniformity")
```
baker.strawberry.uniformity

# Baker table 4, correlation of plots across years
# libs(reshape2)
# mat <- acast(dat, row+col~year)
# round(cor(mat, use='pair'),2)

## End(Not run)

baker.strawberry.uniformity

Uniformity trial of strawberry

Description

Uniformity trial of strawberry

Usage

data("baker.strawberry.uniformity")

Format

A data frame with 700 observations on the following 4 variables.

trial  a factor with levels T1 T2
row    a numeric vector
col    a numeric vector
yield  a numeric vector

Details

In trial T1, 200 plants were grown in two double-row beds at Davis, CA, in 1946. The rows were 1 foot apart. The beds were 42 inches apart. The plants were 10 inches apart within a row, each row consisting of 50 plants.
Field length: 50 plants * 10 inches = 500 inches.
Field width: 12 in + 42 in + 12 in =

In trial T2, 500 plants were grown in single beds. The beds were 30 inches apart. Each bed was 50 plants long with 10 inches between plants.
Field length: 50 plants * 10 in = 500 in.
Field width: 10 beds * 30 in = 300 in.

Source

baker.wheat.uniformity

*baker.wheat.uniformity*

*Uniformity trial of wheat*

**Description**

Uniformity trial of wheat

**Usage**

```r
data("baker.wheat.uniformity")
```

**Format**

A data frame with 225 observations on the following 3 variables.

- `row` row
- `col` col
- `yield` yield
Details

Data was collected in 1939-1940. The trial consists of sixteen 40 ft. x 40 ft. blocks subdivided into nine plots each. The data were secured in 1939-1940 from White Federation wheat. The design of the experiment was square with alleys 20 feet wide between blocks. The plots were 10 feet long with two guard rows on each side.

Morning glories infested the middle two columns of blocks, uniformly over the blocks affected. The data here include missing values for the alleys so that the field map is approximately correct.

Field width: 4 blocks of 40 feet + 3 alleys of 20 feet = 220 feet.
Field length: 4 blocks of 40 feet + 3 alleys of 20 feet = 220 feet.

Source


References

None

Examples

```r
## Not run:

library(agridat)
data(baker.wheat.uniformity)
dat <- baker.wheat.uniformity

libs(desplot)
desplot(dat, yield ~ col*row, 
   flip=TRUE, aspect=1, 
   main="baker.wheat.uniformity")

## End(Not run)
```

---

bancroft.peanut.uniformity

Uniformity trial of peanuts

Description

Uniformity trial of peanuts in Alabama, 1946.

Usage

data("bancroft.peanut.uniformity")
Format

A data frame with 216 observations on the following 5 variables.

row row
col column
yield yield, pounds
block block

Details

The data are obtained from two parts of the same field, located at Wiregrass Substation, Headland, Alabama, USA. Each part had 18 rows, 3 feet wide, 100 feet long. Plots were harvested in 1946. Green weights in pounds were recorded.

Each plot was 16.66 linear feet of row and 3 feet in width, 50 sq feet.

Field width: 6 plots * 16.66 feet = 100 feet
Field length: 18 plots * 3 feet = 54 feet

Conclusions: Based on the relative efficiencies, increasing the size of the plot along the row is better than across the row. Narrow, rectangular plots are more efficient.

Source


References

None

Examples

```r
## Not run:

library(agritag)
data(bancroft.peanut.uniformity)
dat <- bancroft.peanut.uniformity

# match means Bancroft page 3
# dat
# # A tibble: 2 x 2
# #   block  mn
# # <chr> <dbl>
# 1 B1  2.46
# 2 B2  2.05

libs(desplot)
desplot(dat, yield ~ col*row|block, flip=TRUE, aspect=(18*3)/(6*16.66), # true aspect
```
barrero.maize

Multi-environment trial of maize in Texas.

Description
Multi-environment trial of maize in Texas.

Usage
data("barrero.maize")

Format
A data frame with 14568 observations on the following 15 variables.

- **year**: year of testing, 2000-2010
- **yor**: year of release, 2000-2010
- **loc**: location, 16 places in Texas
- **env**: environment (year+loc), 107 levels
- **rep**: replicate, 1-4
- **gen**: genotype, 847 levels
- **dastoflower**: numeric
- **plantheight**: plant height, cm
- **earheight**: ear height, cm
- **population**: plants per hectare
- **lodged**: percent of plants lodged
- **moisture**: moisture percent
- **testweight**: test weight kg/ha
- **yield**: yield, kg/ha

Details
This is a large (14500 records), multi-year, multi-location, 10-trait dataset from the Texas AgriLife Corn Performance Trials.
These data are from 2-row plots approximately 36in wide by 25 feet long.
Barrero et al. used this data to estimate the genetic gain in maize hybrids over a 10-year period of time.
Used with permission of Seth Meyers.
Source


References

None.

Examples

```r
## Not run:
library(agridat)
dat <- barrero.maize
library(lattice)
bwplot(yield ~ factor(year)|loc, dat,
  main="barrero.maize - Yield trends by loc",
  scales=list(x=list(rot=90)))

# Table 6 of Barrero. Model equation 1.
pacman::p_load(dplyr, asreml, lucid)
dat <- arrange(dat, env)
dat <- mutate(dat,
  yearf=factor(year), env=factor(env),
  loc=factor(loc), gen=factor(gen), rep=factor(rep))

m1 <- asreml(yield ~ loc + yearf + loc:yearf, data=dat,
  random = ~ gen + rep:loc:yearf +
  gen:yearf + gen:loc +
  gen:loc:yearf,
  residual = ~ dsum( ~ units|env) )

# Variance components for yield match Barrero table 6.
lucid::vc(m1)[1:5,]
## effect component std.error z.ratio bound
## rep:loc:yearf 0.111 0.01092 10 P 0
## gen 0.505 0.03988 13 P 0
## gen:yearf 0.05157 0.01472 3.5 P 0
## gen:loc 0.02283 0.01520 1.5 P 0.2
## gen:loc:yearf 0.2068 0.01806 11 P 0

summary(vc(m1)[6:112,"component"])) # Means match last row of table 6
## Min. 1st Qu. Median       Mean 3rd Qu.    Max.
## 0.1286 0.3577 0.5571       0.8330 1.0322    2.9867

## End(Not run)
```
Uniformity trials of apples, lemons, oranges, and walnuts, in California & Utah, 1915-1918.

Description

Uniformity trials of apples, lemons, oranges, and walnuts, in California & Utah, 1915-1918.

Format

Each dataset has the following format

- row  row
- col  column
- yield  yield per tree (pounds)

Details

A few of the trees affected by disease were eliminated and the yield was replaced by the average of the eight surrounding trees.

The following details are from Batchelor (1918).

**Jonathan Apples**

"The apple records were obtained from a 10-year old Jonathan apple orchard located at Providence, Utah. The surface soil of this orchard is very uniform to all appearances except on the extreme eastern edge, where the percentage of gravel increases slightly. The trees are planted 16 feet apart, east and west, and 30 feet apart north and south."

Note: The orientation of the field is not given in the paper, but all other fields in the paper have north at the top, so that is assumed to be true for this field as well. Yields may be from 1916.

Field width: 8 trees * 16 feet = 128 feet
Field length: 28 rows * 30 feet = 840 feet

**Eureka Lemon**

The lemon (Citrus limonia) tree yields were obtained from a grove of 364 23-year-old trees, located at Upland, California. The records extend from October 1, 1915, to October 1, 1916. The grove consists of 14 rows of 23-year-old trees, extending north and south, with 26 trees in a row, planted 24 by 24 feet apart. This grove presents the most uniform appearance of any under consideration [in this paper]. The land is practically level, and the soil is apparently uniform in texture. The records show a grouping of several low-yielding trees; yet a field observation gives one the impression that the grove as a whole is remarkably uniform.

Field width: 14 trees * 24 feet = 336 feet
Field length: 26 trees * 24 feet = 624 feet

**Navel 1 at Arlington**

These records were of the 1915-16 yields of one thousand 24-year-old navel-orange trees near Arlington station, Riverside, California. The grove consists of 20 rows of trees from north to south, with 50 trees in a row, planted 22 by 22 feet. A study of the records shows certain distinct high-
and low-yielding areas. The northeast corner and the south end contain notably high-yielding trees. The north two-thirds of the west side contains a large number of low-yielding trees. These areas are apparently correlated with soil variation. Variations from tree to tree also occur, the cause of which is not evident. These variations, which are present in every orchard, bring uncertainty into the results of field experiments.

Field width: 20 trees * 22 feet = 440 feet
Field length: 50 trees * 22 feet = 1100 feet

**Navel 2 at Antelope**

The navel-orange grove later referred to as the Antelope Heights navels is a plantation of 480 ten-year-old trees planted 22 by 22 feet, located at Naranjo, California. The yields are from 1916. The general appearance of the trees gives a visual impression of uniformity greater than a comparison of the individual tree production substantiates.

Field width: 15 trees * 22 feet = 330 feet
Field length: 33 trees * 22 feet = 726 feet

**Valencia Orange**

The Valencia orange grove is composed of 240 15-year-old trees, planted 21 feet 6 inches by 22 feet 6 inches, located at Villa Park, California. The yields were obtained in 1916.

Field width: 12 rows * 22 feet = 264 feet
Field length: 20 rows * 22 feet = 440 feet

**Walnut**

The walnut (Juglans regia) yields were obtained during the seasons of 1915 and 1916 from a 24-year-old Santa Barbara softshell seedling grove, located at Whittier, California. [Note, The yields here appear to be the 1915 yields.] The planting is laid out 10 trees wide and 32 trees long, entirely surrounded by additional walnut plantings, except on a part of one side which is adjacent to an orange grove. The trees are planted on the square system, 50 feet apart.

Field width: 10 trees * 50 feet = 500 feet
Field length: 32 trees * 50 feet = 1600 feet

**Source**


**References**


**Examples**

```r
## Not run:
library(agridat)
libs(desplot)
```
battese.survey

Survey and satellite data for corn and soy areas in Iowa

Description

Survey and satellite data for corn and soy areas in Iowa

Usage

data("battese.survey")
Format

A data frame with 37 observations on the following 9 variables.

- county  county name
- segment  sample segment number (within county)
- countysegs  number of segments in county
- cornhect  hectares of corn in segment
- soyhect  hectares of soy
- cornpix  pixels of corn in segment
- soypix  pixels of soy
- cornmean  county mean of corn pixels per segment
- soymean  county mean of soy pixels per segment

Details

The data are for 12 counties in north-central Iowa in 1978.
The USDA determined the area of soybeans in 37 area sampling units (called 'segments'). Each segment is about one square mile (about 259 hectares). The number of pixels of that were classified as corn and soybeans came from Landsat images obtained in Aug/Sep 1978. Each pixel represents approximately 0.45 hectares.
Data originally compiled by USDA.
This data is also available in R packages: 'rsae::landsat' and 'JoSAE::landsat'.

Source


References


Examples

```r
# Not run:
library(agridat)
data(battese.survey)
dat <- battese.survey

# Battese fig 1 & 2. Corn plot shows outlier in Hardin county
libs(lattice)
dat <- dat[order(dat$cornpix),]
```
```r
require(lattice)
require(lme4, lucid)

# Predict for Humboldt county as in Battese 1982 table 2
5.4662+.3878*290.74
# 118.2152 # \( \mu_i^0 \)
5.4662+.3878*290.74+ -2.8744
# 115.3408 # \( \mu_i^\gamma \)
(185.35+116.43)/2
# 150.89 # \( \bar{y}_i \)

# Survey regression estimator of Battese 1988

# Delete the outlier
data2 <- subset(data, !(county=="Hardin" & soyhect < 30))

# Results match top-right of Battese 1988, p. 33
m3 <- lmer(cornhect ~ cornpix + soyhect + (1|county), data=data2)
fixef(m3)
```
beall.webworms

Counts of webworms in a beet field, with insecticide treatments.

Description
Counts of webworms in a beet field, with insecticide treatments.

Usage
data("beall.webworms")

Format
A data frame with 1300 observations on the following 7 variables.

row row
col column
y count of webworms
block block
trt treatment
spray spray treatment yes/no
lead lead treatment yes/no

Details
The beet webworm lays egg masses as small as 1 egg, seldom exceeding 5 eggs. The larvae can move freely, but usually mature on the plant on which they hatch.

Each plot contained 25 unit areas, each 1 row by 3 feet long. The row width is 22 inches. The arrangement of plots within the blocks seems certain, but the arrangement of the blocks/treatments is not certain, since the authors say "since the plots were 5 units long and 5 wide it is only practicable to combine them into groups of 5 in one direction or the other".

Treatment 1 = None. Treatment 2 = Contact spray. Treatment 3 = Lead arsenate. Treatment 4 = Both spray, lead arsenate.
Source


References


Examples

```r
## Not run:
library(agridat)
data(beall.webworms)
dat <- beall.webworms

# Match Beall table 1
# with(dat, table(y,trt))

libs(lattice)
histogram(~y|trt, data=dat, layout=c(1,4), as.table=TRUE,
main="beall.webworms")

# Visualize Beall table 6. Block effects may exist, but barely.
libs(desplot)
grays <- colorRampPalette(c("white","#252525"))
desplot(dat, y ~ col*row,
      col.regions=grays(10),
at=0:10-0.5,
      out1=block, out2=trt, num=trt, flip=TRUE, # aspect unknown
      main="beall.webworms (count of worms)"
)

# Following plot suggests interaction is needed
# with(dat, interaction.plot(spray, lead, y))

# Try the models of Kosma et al, Table 1.

# Poisson model
m1 <- glm(y ~ block + spray*lead, data=dat, family="poisson")
logLik(m1) # -1497.719 (df=16)

# Negative binomial model
# libs(MASS)
# m2 <- glm.nb(y ~ block + spray*lead, data=dat)
# logLik(m2) # -1478.341 (df=17)

# Conway=Maxwell-Poisson model (takes several minutes)
# libs(spam)
# # estimate nu parameter
# m3 <- fitme(y ~ block + spray*lead, data=dat, family = COMPoisson())
# logLik(m3) # -1475.999
```
beaven.barley

Yields of 8 barley varieties in 1913 as used by Student.

Description

Yields of 8 barley varieties in 1913.

Usage

data("beaven.barley")

Format

A data frame with 160 observations on the following 4 variables.

row row
col column
gen genotype
yield yield (grams)

Details

Eight races of barley were grown on a regular pattern of plots.
These data were prepared from Richey (1926) because the text was cleaner.
Each plot was planted 40 inches on a side, but only the middle square 36 inches on a side was harvested.
Field width: 32 plots * 3 feet = 96 feet
Field length: 5 plots * 3 feet = 15 feet

Source

https://doi.org/10.1093/biomet/15.3-4.271

References

Examples

```r
## Not run:

library(agridat)

data(beaven.barley)
dat <- beaven.barley

# Match the means shown in Richey table IV
tapply(dat$yield, dat$gen, mean)
##  a  b  c  d  e  f  g  h
## 298.080 300.710 318.685 295.260 306.410 276.475 304.605 271.820

# Compare to Student 1923, diagram I,II
libs(desplot)
desplot(dat, yield ~ col*row,
        aspect=15/96, # true aspect
        main="beaven.barley - variety trial", text=gen)

## End(Not run)
```

becker.chicken

**Mating crosses of chickens**

**Description**

Mating crosses of chickens

**Usage**

`data("becker.chicken")`

**Format**

A data frame with 45 observations on the following 3 variables.

- **male** male parent
- **female** female parent
- **weight** weight (g) at 8 weeks

**Details**

From a large flock White Rock chickens, five male sires were chosen and mated to each of three female dams, producing 3 female progeny. The data are body weights at eight weeks of age. Becker (1984) used these data to demonstrate the calculation of heritability.

**Source**

Examples

```r
## Not run:

library(agridat)
data(becker.chicken)
dat <- Becker.chicken
libs(lattice)
dotplot(weight ~ female, data=dat, group=male,
main="becker.chicken - progeny weight by M*F",
  xlab="female parent",ylab="progeny weight",
auto.key=list(columns=5))

# Sums match Becker
# sum(dat$weight)
# aggregate(weight ~ male + female, dat, FUN=sum)

# Variance components
libs(lme4,lucid)
m1 <- lmer(weight ~ (1|male) + (1|female), data=dat)
# vc(m1)
## grp var1 var2 vcov sdcor
## 1 female (Intercept) <NA> 1096 33.1
## 2 male (Intercept) <NA> 776.8 27.87
## 3 Residual <NA> <NA> 5524 74.32

# Calculate heritabilities
# s2m <- 776 # variability for males
# s2f <- 1095 # variability for females
# s2w <- 5524 # variability within crosses
# vp <- s2m + s2f + s2w # 7395
# 4*s2m/vp # .42 male heritability
# 4*s2f/vp # .59 female heritability

## End(Not run)
```

**Description**

RCB experiment of spring barley in United Kingdom
**Format**

A data frame with 225 observations on the following 4 variables.

- **col**: column (also blocking factor)
- **row**: row
- **yield**: yield
- **gen**: variety/genotype

**Details**

RCB design, each column is one rep.

Used with permission of David Higdon.

**Source**


**References**


**Examples**

```r
## Not run:

library(agridat)
data(besag.bayesian)
dat <- besag.bayesian

# Yield values were scaled to unit variance
# var(dat$yield, na.rm=TRUE)
# .999

# Besag Fig 2. Reverse row numbers to match Besag, Davison
dat$rrow <- 76 - dat$row
libs(lattice)
xyplot(yield ~ rrow|col, dat, layout=c(1,3), type='s',
       xlab="row", ylab="yield", main="besag.bayesian")

library(asreml)

# Use asreml to fit a model with AR1 gradient in rows
dat <- transform(dat, cf=factor(col), rf=factor(rrow))
m1 <- asreml(yield ~ -1 + gen, data=dat, random= ~ ar1v(rf))
m1 <- update(m1)
m1 <- update(m1)
m1 <- update(m1)

# Visualize trends, similar to Besag figure 2.
```
# Need 'as.vector' because asreml4 uses a named vector
dat$res <- unname(m1$resid)
dat$geneff <- coef(m1)$fixed[as.numeric(dat$gen)]
dat <- transform(dat, fert=yield-geneff-res)
libs(lattice)
xyplot(geneff ~ rrow|col, dat, layout=c(1,3), type='s',
main="besag.bayesian - Variety effects", ylim=c(5,15 ))
xyplot(fert ~ rrow|col, dat, layout=c(1,3), type='s',
main="besag.bayesian - Fertility", ylim=c(-2,2))
xyplot(res ~ rrow|col, dat, layout=c(1,3), type='s',
main="besag.bayesian - Residuals", ylim=c(-4,4))

## End(Not run)

### Description

Competition experiment in beans with height measurements

### Usage

data("besag.beans")

### Format

A data frame with 152 observations on the following 6 variables.

gen  genotype / variety
height  plot height, cm
yield  plot yield, g
row  row / block
rep  replicate factor
col  column

### Details

Field beans of regular height were grown beside shorter varieties. In each block, each variety occurred once as a left-side neighbor and once as a right-side neighbor of every variety (including itself). Border plots were placed at the ends of each block. Each block with 38 adjacent plots. Each plot was one row, 3 meters long with 50 cm spacing between rows. No gaps between plots. Spacing between plants was 6.7 cm. Four blocks (rows) were used, each with six replicates.

Plot yield and height was recorded.

Kempton and Lockwood used models that adjusted yield according to the difference in height of neighboring plots.

Field length: 4 plots * 3m = 12m
Field width: 38 plots * 0.5 m = 19m
Source

Julian Besag and Rob Kempton (1986). Statistical Analysis of Field Experiments Using Neigh-

References

Kempton, RA and Lockwood, G. (1984). Inter-plot competition in variety trials of field beans (Vicia

Examples

```
### Not run:

library(agridat)

data(besag.beans)
dat = besag.beans

libs(desplot)
desplot(dat, yield ~ col*row,
       aspect=12/19, out1=row, out2=rep, num=gen, cex=1, # true aspect
       main="besag.beans")

libs(reshape2)
# Add a covariate = excess height of neighbors
mat <- acast(dat, row~col, value.var='height')
mat2 <- matrix(NA, nrow=4, ncol=38)
dat2 <- melt(mat2)
colnames(dat2) <- c('row','col','cov')
dat <- merge(dat, dat2)

# Drop border plots
dat <- subset(dat, rep != 'R0')

libs(lattice)
# Plot yield vs neighbors height advantage
xyplot(yield~cov, data=dat, group=gen,
       main="besag.beans",
       xlab="Mean excess heights of neighbor plots",
       auto.key=list(columns=3))

# Trial mean.
mean(dat$yield) # 391 matches Kempton table 3

# Mean excess height of neighbors for each genotype
# tapply(dat$cov, dat$gen, mean)/2 # Matches Kempton table 4

# Variety means, matches Kempton table 4 mean yield
m1 <- lm(yield ~ -1 + gen, dat)
coef(m1)
```
# Full model used by Kempton, eqn 5. Not perfectly clear.
# Appears to include rep term, perhaps within block

```r
dat$blk <- factor(dat$row)
dat$blkrep <- factor(paste(dat$blk, dat$rep))
m2 <- lm(yield ~ -1 + gen + blkrep + cov, data=dat)
coef(m2) # slope 'cov' = -0.72, while Kempton says -0.79
```

## End(Not run)

---

**besag.checks**

*Check variety yields in winter wheat.*

**Description**

Check variety yields in winter wheat.

**Usage**

`data("besag.checks")`

**Format**

A data frame with 364 observations on the following 4 variables.

- `yield` yield, units of 10g
- `row` row
- `col` column
- `gen` genotype/variety

**Details**

This data was used by Besag to show the spatial variation in a field experiment, but Besag did not use the data for any analysis.

Yields of winter wheat varieties (Bounty and Huntsman) at the Plant Breeding Institute, Cambridge, in 1980. These data are the 'checks' genotypes in a larger variety trial.

There is a column of checks, then five columns of new varieties. Repeat.

Plot dimensions approx 1.5 by 4.5 metres

Field length: 52 rows * 4.5 m = 234 m

Field width: 31 columns * 1.5 m = 46.5

Electronic version of data supplied by David Clifford.

**Source**

besag.elbatan

References

Kempton, Statistical Methods for Plant Variety Evaluation, page 91–92

Examples

```r
library(agridat)
data(besag.checks)
dat <- besag.checks
## Not run:
libs(desplot)
desplot(dat, yield~col*row,
       num=gen, aspect=234/46.5, # true aspect
       main="besag.checks")
## End(Not run)
```

besag.elbatan

RCB experiment of wheat, 50 varieties in 3 blocks with strong spatial trend.

Description

RCB experiment of wheat, 50 varieties in 3 blocks with strong spatial trend.

Format

A data frame with 150 observations on the following 4 variables.

- `yield` yield of wheat
- `gen` genotype, factor with 50 levels
- `col` column/block
- `row` row

Details

RCB experiment on wheat at El Batan, Mexico. There are three single-column replicates with 50 varieties in each replicate.

Plot dimensions are not given by Besag.

Data retrieved from https://web.archive.org/web/19991008143232/www.stat.duke.edu/~higdon/trials/elbatan.dat

Used with permission of David Higdon.

Source

References


Examples

```r
## Not run:
library(agridat)
data(besag.elbatan)
dat <- besag.elbatan

libs(desplot)
desplot(dat, yield~col*row,
num=gen, # aspect unknown
main="besag.elbatan - wheat yields")

# Besag figure 1
library(lattice)
xyplot(yield~row|col, dat, type=c('l'),
layout=c(1,3), main="besag.elbatan wheat yields")

# RCB
m1 <- lm(yield ~ 0 + gen + factor(col), dat)
p1 <- coef(m1)[1:50]

# Formerly used gam package, but as of R 3.1, Rcmd check --as-cran
# is complaining
# Calls: plot.gam ... model.matrix.gam -> predict -> predict.gam -> array
# but it works perfectly in interactive mode !!!
# Remove the FALSE to run the code below
if(is.element("gam", search())) detach(package:gam)
libs(mgcv)
m2 <- mgcv::gam(yield ~ -1 + gen + factor(col) + s(row), data=dat)
plot(m2, residuals=TRUE, main="besag.elbatan")
pred <- cbind(dat, predict(m2, dat, type="terms"))
# Need to correct for the average loess effect, which is like
# an overall intercept term.
adjlo <- mean(pred$"s(row)")
p2 <- coef(m2)[1:50] + adjlo

# Compare estimates
lims <- range(c(p1,p2))
plot(p1, p2, xlab="RCB prediction",
 ylab="RCB with smooth trend (predicted)",
 type='n', xlim=lims, ylim=lims,
 main="besag.elbatan")
text(p1, p2, 1:50, cex=.5)
abline(0,1,col="gray")
```
Description

Presence of footroot disease in an endive field

Format

A data frame with 2506 observations on the following 3 variables.

col column
row row
disease plant is diseased, Y=yes,N=no

Details

In a field of endives, does each plant have footrot, or not? Data are binary on a lattice of 14 x 179 plants.
Modeled as an autologistic distribution.
We assume the endives are a single genotype.
Besag (1978) may have had data taken at 4 time points. This data was extracted from Friel and Pettitt. It is not clear what, if any, time point was used.
Friel does not give the dimensions. Besag is not available.

Source


References


Examples

## Not run:

```r
library(agridat)
data(besag.endive)
dat <- besag.endive

# Incidence map. Figure 2 of Friel and Pettitt
```
libs(desplot)
grays <- colorRampPalette(c("#d9d9d9","#252525"))
desplot(dat, disease~col*row,
    col.regions=grays(2),
    aspect = 0.5, # aspect unknown
    main="besag.endive - Disease incidence")

# Besag (2000) "An Introduction to Markov Chain Monte Carlo" suggested
# that the autologistic model is not a very good fit for this data.
# We try it anyway. No idea if this is correct or how to interpret...

libs(nsgspatial)
A = adjacency.matrix(179,14)
X = cbind(x=dat$col, y=dat$row)
Z = as.numeric(dat$disease=="Y")
m1 <- autologistic(Z ~ 0+X, A=A, control=list(confint="none"))

summary(m1)
## Coefficients:
## Estimate Lower Upper MCSE
## Xx -0.007824 NA NA NA
## Xy -0.144800 NA NA NA
## eta 0.806200 NA NA NA

libs(asreml)
# Now try an AR1xAR1 model.
dat2 <- transform(dat, xf=factor(col), yf=factor(row),
    pres=as.numeric(disease=="Y"))

m2 <- asreml(pres ~ 1, data=dat2,
    resid = ~ar1(xf):ar1(yf))
# The 0/1 response is arbitrary, but there is some suggestion
# of auto-correlation in the x (.17) and y (.10) directions,
# suggesting the pattern is more 'patchy' than just random noise,
# but is it meaningful?

libs(lucid)
vc(m2)
## effect component std.error z.ratio bound
## xf:yf(R) 0.1301 0.003798 34 P 0
## xf:yf!xf!cor 0.1699 0.01942 8.7 U 0
## xf:yf!yf!cor 0.09842 0.02038 4.8 U 0

## End(Not run)
Description

Multi-environment trial of corn, incomplete-block design location.

Format

A data frame with 1152 observations on the following 7 variables.

- county
- row
- col
- rep
- block
- yield
- gen genotype, 1-64

Details

Multi-environment trial of 64 corn hybrids in six counties in North Carolina. Each location had 3 replicates in an incomplete-block design with an 18x11 lattice of plots whose length-to-width ratio was about 2:1.

Note: In the original data, each county had 6 missing plots. This data has rows for each missing plot that uses the same county/block/rep to fill-out the row, sets the genotype to G01, and sets the yield to missing. These missing values were added to the data so that asreml could more easily do AR1xAR1 analysis using rectangular regions.

Each location/panel is:
- Field length: 18 rows * 2 units = 36 units.
- Field width: 11 plots * 1 unit = 11 units.


Used with permission of David Higdon.

Source


Examples

```r
## Not run:
library(agridat)
data(besag.met)
dat <- besag.met
libs(desplot)
desplot(dat, yield ~ col*row|county,
```
# Average reps
datm <- aggregate(yield ~ county + gen, data=dat, FUN=mean)

# Sections below fit heteroskedastic variance models (variance for each variety)
# asreml takes 1 second, lme 73 seconds, SAS PROC MIXED 30 minutes

# lme
# libs(nlme)
# m1l <- lme(yield ~ -1 + gen, data=datm, random=-1|county,
# # weights = varIdent(form=-1|gen))
# m1l$sigma^2 * c(1, coef(m1l$modelStruct$varStruct, unc = FALSE))^2
## G02  G03  G04  G05  G06  G07  G08
## 91.90 210.75 63.03 112.05 28.39 237.36 72.72 42.97
## ... etc ...

# Note, the FA biplots from asreml3 and asreml4 are surprisingly
# different from each other. The predicted-value biplots are
# almost identical.

libs(asreml)
# asreml4

# Average reps
datm <- aggregate(yield ~ county + gen, data=dat, FUN=mean)
# asreml Using 'rcov' ALWAYS requires sorting the data
datm <- datm[order(datm$gen),]

m1 <- asreml(yield ~ gen, data=datm, random=~1|county, residual = ~ dsum(~ units|gen))

libs(lucid)
vc(m1)[1:7,]
## effect component std.error z.ratio bound
## county 1324 836.1 1.6 P 0.2
## gen_G01!R 91.98 58.91 1.6 P 0.1
## gen_G02!R 210.6 133.6 1.6 P 0.1
## gen_G03!R 63.06 40.58 1.6 P 0.1
## gen_G04!R 112.1 71.59 1.6 P 0.1
## gen_G05!R 28.35 18.57 1.5 P 0.2
## gen_G06!R 237.4 150.8 1.6 P 0

# We get the same results from asreml & lme
# plot(m1$vparameters[-1],
# m1$sigma^2 * c(1, coef(m1$modelStruct$varStruct, unc = FALSE))^2)
# The following example shows how to construct a GxE biplot
# from the FA2 model.

dat <- besag.met
dataf <- transform(dat, xf=factor(col), yf=factor(row))
do[dat$county, dat$xf, dat$yf],]

# First, AR1xAR1
m1 <- asreml(yield ~ county, data=dat,
random = ~ gen:county,
residual = ~ dsum(~ ar1(xf):ar1(yf)~country))
# Add FA1
m2 <- update(m1, random=-gen:fa(county,1)) # rotate.FA=FALSE
# FA2
m3 <- update(m2, random=-gen:fa(county,2))
asreml.options(extra=50)
m3 <- update(m3, maxit=50)
asreml.options(extra=0)

# Use the loadings to make a biplot
vars <- vc(m3)
psi <- vars[grep("!var", vars$effect), "component"]
l1 <- vars[grep("!fa1", vars$effect), "component"]
l2 <- vars[grep("!fa2", vars$effect), "component"]
mat <- as.matrix(data.frame(psi, l1, l2))
# I tried using rotate.fa=FALSE, but it did not seem to
# give orthogonal vectors. Rotate by hand.
rot <- svd(mat[,,-1])$v # rotation matrix
lam <- mat[,,-1]
colnames(lam) <- c("load1", "load2")

co3 <- coef(m3)$random # Scores are the GxE coefficients
ix1 <- grep("_Comp1", rownames(co3))
ix2 <- grep("_Comp2", rownames(co3))
sco <- matrix(c(co3[ix1], co3[ix2]), ncol=2, byrow=FALSE)
sco <- sco
dimnames(sco) <- list(levels(dat$gen), c("load1", "load2"))

rownames(lam) <- levels(dat$county)
sco[,1:2] <- -1 * sco[,1:2]
lam[,1:2] <- -1 * lam[,1:2]
biplot(sco, lam, cex=.5, main="FA2 coefficient biplot (asreml4)"
# G variance matrix
gvar <- lam

# Now get predictions and make an ordinary biplot
p3 <- predict(m3, data=dat, classify="county:gen")
p3 <- p3$pvals
libs("gge")
b3 <- gge(p3, predicted.value ~ gen*county, scale=FALSE)
if(interactive()) dev.new()
# Very similar to the coefficient biplot
biplot(b3, stand=FALSE, main="SVD biplot of FA2 predictions")
### Description

Four-way factorial agronomic experiment in triticale

### Usage

```r
data("besag.triticale")
```

### Format

A data frame with 54 observations on the following 7 variables.

- `yield` yield, g/m^2
- `row` row
- `col` column
- `gen` genotype / variety, 3 levels
- `rate` seeding rate, kg/ha
- `nitro` nitrogen rate, kw/ha
- `regulator` growth regulator, 3 levels

### Details

Experiment conducted as a factorial on the yields of triticale. Fully randomized. Plots were 1.5m x 5.5m, but the orientation is not clear.

Besag and Kempton show how accounting for neighbors changes non-significant genotype differences into significant differences.

### Source


### References

None.
Examples

```r
## Not run:
library(agridat)
data(besag.triticale)
dat <- besag.triticale
dat <- transform(dat, rate=factor(rate), nitro=factor(nitro))
dat <- transform(dat, xf=factor(col), yf=factor(row))

libs(desplot)
desplot(dat, yield ~ col*row, 
  # aspect unknown
  main="besag.triticale")

# Besag & Kempton are not perfectly clear on the model, but
# indicate that there was no evidence of any two-way interactions.
# A reduced, main-effect model had genotype effects that were
# "close to significant" at the five percent level.
# The model below has p-value of gen at .04, so must be slightly
# different than their model.
m2 <- lm(yield ~ gen + rate + nitro + regulator + yf, data=dat)
anova(m2)

# Similar, but not exact, to Besag figure 5
dat$res <- resid(m2)
libs(lattice)
xyplot(res ~ col|as.character(row), data=dat, 
  as.table=TRUE, type="s", layout=c(1,3),
  main="besag.triticale")

libs(asreml) # asreml4

# Besag uses an adjustment based on neighboring plots.
# This analysis fits the standard AR1xAR1 residual model

dat <- dat[order(dat$xf, dat$yf), ]
m3 <- asreml(yield ~ gen + rate + nitro + regulator +
  gen:rate + gen:nitro + gen:regulator +
  rate:nitro + rate:regulator +
  nitro:regulator + yf, data=dat,
  resid = ~ ar1(xf):ar1(yf))
wald(m3) # Strongly significant gen, rate, regulator
```

```r
## Df Sum of Sq Wald statistic Pr(Chisq)
## (Intercept) 1 1288255 103.971 < 2.2e-16 ***
## gen 2 903262 72.899 < 2.2e-16 ***
## rate 1 104774 8.456 0.003638 **
## nitro 1 282 0.023 0.880139
## regulator 2 231403 18.676 8.802e-05 ***
## yf 2 3788 0.306 0.858263
## gen:rate 2 1364 0.110 0.946461
## gen:nitro 2 30822 2.488 0.288289
## gen:regulator 4 37269 3.008 0.556507
```
## rate:nitro 1 1488 0.120 0.728954
## rate:regulator 2 49296 3.979 0.136795
## nitro:regulator 2 41019 3.311 0.191042
## residual (MS) 12391

## End(Not run)

### blackman.wheat

#### Multi-environment trial of wheat, conventional and semi-dwarf varieties

**Description**

Multi-environment trial of wheat, conventional and semi-dwarf varieties, 7 locs with low/high fertilizer levels.

**Format**

A data frame with 168 observations on the following 5 variables.

- gen  genotype
- loc  location
- nitro  nitrogen fertilizer, low/high
- yield  yield (g/m^2)
- type  type factor, conventional/semi-dwarf

**Details**

Conducted in U.K. in 1975. Each loc had three reps, two nitrogen treatments.

Locations were Begbroke, Boxworth, Crafts Hill, Earith, Edinburgh, Fowlmere, Trumpington.

At the two highest-yielding locations, Earith and Edinburgh, yield was _lower_ for the high-nitrogen treatment. Blackman et al. say "it seems probable that effects on development and structure of the crop were responsible for the reductions in yield at high nitrogen".

**Source**


**References**

Examples

```r
## Not run:
library(agridat)
data(blackman.wheat)
dat <- blackman.wheat
libs(lattice)
# Semi-dwarf generally higher yielding than conventional
# bwplot(yield~type|loc, dat, main="blackman.wheat")

# Peculiar interaction--Ear/Edn locs have reverse nitro response
dotplot(gen~yield|loc, dat, group=nitro, auto.key=TRUE,
    main="blackman.wheat: yield for low/high nitrogen")

# Height data from table 6 of Blackman. Height at Trumpington loc.
# Shorter varieties have higher yields, greater response to nitro.
heights <- data.frame(gen=c("Cap", "Dur", "Fun", "Hob", "Hun", "Kin",
    "Ran", "Spo", "T64", "T68", "T95", "Tem"),
    ht=c(101,76,80,98,88,81,86,73,78,93))
dat$height <- heights$ht[match(dat$gen, heights$gen)]
xyplot(yield~height|loc, dat, group=nitro, type=c("p","r"),
    main="blackman.wheat",
    subset=loc=="Tru", auto.key=TRUE)

libs(reshape2)
# AMMI-style biplot Fig 6.4 of Gower 2011
dat$env <- factor(paste(dat$loc, dat$nitro, sep="-"))
datm <- acast(dat, gen~env, value.var='yield')
datm <- sweep(datm, 1, rowMeans(datm))
datm <- sweep(datm, 2, colMeans(datm))
bilplot(prcomp(datm), main="blackman.wheat AMMI-style biplot")
```

## End(Not run)

bliss.borers

**Corn borer infestation under four treatments**

**Description**

Corn borer infestation under four treatments

**Format**

A data frame with 48 observations on the following 3 variables.
borers  number of borers per hill
   treat  treatment factor
   freq  frequency of the borer count

Details

Four treatments to control corn borers. Treatment 1 is the control.
In 15 blocks, for each treatment, 8 hills of plants were examined, and the number of corn borers
present was recorded. The data here are aggregated across blocks.
Bliss mentions that the level of infestation varied significantly between the blocks.

Source

Geoffrey Beall. 1940. The Fit and Significance of Contagious Distributions when Applied to Ob-

Examples

## Not run:

library(agridat)
data(bliss.borers)
dat <- bliss.borers

# Add 0 frequencies
dat0 <- expand.grid(borers=0:26, treat=c('T1','T2','T3','T4'))
dat0 <- merge(dat0,dat, all=TRUE)
dat0$freq[is.na(dat0$freq)] <- 0

# Expand to individual (non-aggregated) counts for each hill
dd <- data.frame(borers = rep(dat0$borers, times=dat0$freq),
                 treat = rep(dat0$treat, times=dat0$freq))

libs(lattice)
histogram(~borers|treat, dd, type='count', breaks=0:27-.5,
          layout=c(1,4), main="bliss.borers", xlab="Borers per hill")

libs(MASS)
m1 <- glm.nb(borers~0+treat, data=dd)
# Bliss, table 3, presents treatment means, which are matched by:
exp(coef(m1)) # 4.033333 3.166667 1.483333 1.508333
# Bliss gives treatment values k = c(1.532,1.764,1.333,1.190).
# The mean of these is 1.45, similar to this across-treatment estimate
m1$theta # 1.47

# Plot observed and expected distributions for treatment 2
bond.diallel

Diallel cross of winter beans

Description

Diallel cross of winter beans

Format

A data frame with 36 observations on the following 3 variables.

- female  female parent
- male    male parent
- yield   yield, grams/plot
- stems   stems per plot
- nodes   podded nodes per stem
- pods    pods per podded node
- seeds   seeds per pod
- weight  weight (g) per 100 seeds
- height  height (cm) in April
- width   width (cm) in April
- flower  mean flowering date in May
Details

Yield in grams/plot for full diallel cross between 6 inbred lines of winter beans. Values are means over two years.

Source


References


Examples

```r
## Not run:
library(agridat)
data(bond.diallel)
dat <- bond.diallel
# Because these data are means, we will not be able to reproduce
# the anova table in Bond. More useful as a multivariate example.

libs(corrgram)
corrgram(dat[, 3:11], main="bond.diallel",
         lower=panel.pts)
# Multivariate example from sommer package
corrgram(dat[,c("stems","pods","seeds")],
          lower=panel.pts, upper=panel.conf, main="bond.diallel")

libs(sommer)
m1 <- mmer(cbind(stems,pods,seeds) ~ 1,
            random= ~ vs(female)+vs(male),
            rcov= ~ vs(units),
            data=dat)

#### genetic variance covariance
cov2cor(m1$sigma$`u:female`)  
cov2cor(m1$sigma$`u:male`)  
cov2cor(m1$sigma$`u:units`)  

## End(Not run)
```
**bose.multi.uniformity**  
*Uniformity trials of barley, wheat, lentils*

**Description**
Uniformity trials of barley, wheat, lentils in India 1930-1932.

**Usage**
```r
data("bose.multi.uniformity")
```

**Format**
A data frame with 1170 observations on the following 5 variables.

- `year` year
- `crop` crop
- `row` row ordinate
- `col` column ordinate
- `yield` plot yield, grams

**Details**
A field about 1/4 acre was sown in three consecutive years (beginning in 1929-1930) with barley, wheat, and lentil.

At harvest, borders 3 feet on east and west and 6 feet on north and south were removed. The field was divided into plots four feet square, which were harvested separately, measured in grams.

Fertility contours of the field were somewhat similar across years, with correlation values across years 0.45, 0.48, 0.21.

Field width: 15 plots * 4 feet = 60 feet.
Field length: 26 plots * 4 feet = 104 feet.

**Conclusions:**
"An experimental field which may be sensibly uniform for one crop or for one season may not be so for another crop or in a different season" p. 592.

**Source**

**References**
Examples

```r
## Not run:
library(agridat)
data(bose.multi.uniformity)
dat <- bose.multi.uniformity

# match sum at bottom of Bose tables 1, 4, 5
# library(dplyr)
# dat
libs(desplot, dplyr)
# Calculate percent of mean yield for each year
dat <- group_by(dat, year)
dat <- mutate(dat, pctyld = (yield-mean(yield))/mean(yield))

dat <- ungroup(dat)
dat <- mutate(dat, year=as.character(year))
# Bose smoothed the data by averaging 2×3 plots together before drawing
# contour maps. Heatmaps of raw data have similar structure to Bose Fig 1.
desplot(dat, pctyld ~ col*row|year,
     tick=TRUE, flip=TRUE, aspect=(26)/(15),
     main="bose.multi.* - Percent of mean yield")

# contourplot() results need to be mentally flipped upside down
# contourplot(pctyld ~ col*row|year, dat,
#     region=TRUE, as.table=TRUE, aspect=26/15)

## End(Not run)
```

box.cork

Weight of cork samples on four sides of trees

Description

The cork data gives the weights of cork borings of the trunk for 28 trees on the north (N), east (E), south (S) and west (W) directions.

Format

Data frame with 28 observations on the following 5 variables.

- **tree**  tree number
- **dir**  direction N,E,S,W
- **y**  weight of cork deposit (centigrams), north direction

Source

box.cork

References


Examples

```r
## Not run:
library(agridat)
data(box.cork)
dat <- box.cork

libs(reshape2, lattice)
dat2 <- acast(dat, tree ~ dir, value.var="y")
splom(dat2, pscales=3,
    prepanel.limits = function(x) c(25,100),
    main="box.cork", xlab="Cork yield on side of tree",
    panel=function(x,y,...){
        panel.splom(x,y,...)
        panel.abline(0,1,col="gray80")
    })

## Radial star plot, each tree is one line
libs(plotrix)
libs(reshape2)
dat2 <- acast(dat, tree ~ dir, value.var="y")
radial.plot(dat2, start=pi/2, rp.type="p", clockwise=TRUE,
    radial.lim=c(0,100), main="box.cork",
    lw=2, labels=c("North","East","South","West"),
    line.col=rep(c("royalblue","red","#009900","dark orange",
        "#999999","#a6761d","deep pink"),
        length=nrow(dat2)))

# asreml 4
libs(asreml)

# Unstructured covariance
dat$dir <- factor(dat$dir)
dat$tree <- factor(dat$tree)
dat <- dat[order(dat$tree, dat$dir),]

# Unstructured covariance matrix
m1 <- asreml(y~dir, data=dat, residual = ~ tree:us(dir))

libs(lucid)
vct(m1)

# Note: 'rcor' is a personal function to extract the correlations
# into a matrix format
# round(kw::rcor(m1)$dir, 2)
```
bradley.multi.uniformity

Uniformity trial of 4 crops on the same land

Description

Uniformity trial of 4 crops on the same land in Trinidad.

Usage

data("bradley.multi.uniformity")

Format

A data frame with 440 observations on the following 5 variables.

row row
col column
yield yield, pounds per plot
season season
crop crop
Details

Experiments conducted in Trinidad.

Plots were marked in May 1939 in Fields 1, 2, and 3. Prior to 1939 it was difficult to obtain significant results on this land.

Plots were 1/40 acre each, 33 feet square. Discard between blocks (the rows) was 7 feet and between plots (the columns) was 4 feet. For roadways, a gap of 14 feet is between blocks 10 and 11 and a gap of 10 feet between plots E/F (which we call columns 5/6).

Data was collected for 4 crops. Two other crops had poor germination and were omitted.

Field width: 10 plots * 33 feet + 8 gaps * 4 feet + 1 gap * 10 = 372 feet
Field length: 11 blocks (plots) * 33 feet + 9 gaps * 7 feet + 1 gap * 14 feet = 440 feet
Crop 1. Woolly Pyrol. Crop cut at flowering and weighed in pounds. Note, woolly pyrol appears to be a bean also called black gram, phaseolus mungo.
Crop 2. Woolly Pyrol. Crop cut at flowering and weighed in pounds.
Crop 3. Maize. Net weight of cobs in pounds. Source document also has number of cobs.
Crop 4. Yams. Weights in pounds. Source document has weight to 1/4 pound, which has here been rounded to the nearest pound. (Half pounds were rounded to nearest even pound.) Source document also has number of yams.

Notes by Bradley.

The edges of the field tended to be slightly higher yielding. Thought to be due to the heavier cultivation which the edges recieve (p. 18).

The plot in row 9, col 7 (9G in Bradley) is higher yielding than its neighbors, thought to be the site of a saman tree dug up and burned when the field was plotted. Bits of charcoal were still in the soil.

Bradley also examined soil samples on selected plots and looked at nutrients, moisture, texture, etc. The selected plots were 4 high-yielding plots and 4 low-yielding plots. Little difference was observed. Unexpectedly, yams gave higher yield on plots with more compaction.

Source

P. L. Bradley (1941). A study of the variation in productivity over a number of fixed plots in field 2. Dissertation: The University of the West Indies. Appendix 1a, 1b, 1c, 1d. https://hdl.handle.net/2139/41264

The data are repeated in: C. E. Wilson. Study of the plots laid out on field II with a view to obtaining plot-fertility data for use in future experiments on these plots, season 1940-41. Dissertation: The University of the West Indies. Page 36-39. https://uwispace.sta.uwi.edu/dspace/handle/2139/43658

References

None

Examples

## Not run:

library(agridat)
data(bradley.multi.uniformity)


```r
dat <- bradley.multi.uniformity

# figures similar to Bradley, pages 11-15
libs(desplot)
desplot(dat, yield ~ col*row, subset=season==1,
    flip=TRUE, aspect=433/366, # true aspect (omits roadways)
    main="bradley.multi.uniformity - season 1, woolly pyrol")

desplot(dat, yield ~ col*row, subset=season==2,
    flip=TRUE, aspect=433/366, # true aspect (omits roadways)
    main="bradley.multi.uniformity - season 2, woolly pyrol")

desplot(dat, yield ~ col*row, subset=season==3,
    flip=TRUE, aspect=433/366, # true aspect (omits roadways)
    main="bradley.multi.uniformity - season 3, maize")

desplot(dat, yield ~ col*row, subset=season==4,
    flip=TRUE, aspect=433/366, # true aspect (omits roadways)
    main="bradley.multi.uniformity - season 4, yams")

dat1 <- subset(bradley.multi.uniformity, season==1)
dat2 <- subset(bradley.multi.uniformity, season==2)
dat3 <- subset(bradley.multi.uniformity, season==3)
dat4 <- subset(bradley.multi.uniformity, season==4)

# to combine plots across seasons, each yield value was converted to percent
# of maximum yield in that season. Same as Bradley, page 17.
dat1$percent <- dat1$yield / max(dat1$yield) * 100
dat2$percent <- dat2$yield / max(dat2$yield) * 100
dat3$percent <- dat3$yield / max(dat3$yield) * 100
dat4$percent <- dat4$yield / max(dat4$yield) * 100

# make sure data is in same order, then combine
dat1 <- dat1[order(dat1$col, dat1$row),]
dat2 <- dat2[order(dat2$col, dat2$row),]
dat3 <- dat3[order(dat3$col, dat3$row),]
dat4 <- dat4[order(dat4$col, dat4$row),]
dat14 <- dat1[,c('Var.row','Var.col')]
dat14$fertility <- dat1$percent + dat2$percent + dat3$percent + dat4$percent

libs(desplot)
desplot(dat14, fertility ~ col*row,
    tick=TRUE, flip=TRUE, aspect=433/366, # true aspect (omits roadways)
    main="bradley.multi.uniformity - fertility")
```

## End(Not run)
Description

Rape seed yields for 5 genotypes, 3 years, 9 locations.

Format

A data frame with 135 observations on the following 4 variables.

- gen: genotype
- year: year, numeric
- loc: location, 9 levels
- yield: yield, kg/ha

Details

The yields are the mean of 4 reps.

Note, in table 2 of Brandle, the value of Triton in 1985 at Bagot is shown as 2355, but should be 2555 to match the means reported in the paper.

Used with permission of P. McVetty.

Source


Examples

## Not run:

```r
library(aagridat)
data(brandle.rape)
dat <- brandle.rape

libs(lattice)
dotplot(gen~yield|loc, dat, group=year, auto.key=list(columns=3),
main="brandle.rape, yields per location", ylab="Genotype")
```

# Matches table 4 of Brandle
# round(tapply(dat$yield, dat$gen, mean),0)

# Brandle reports variance components:
# sigma^2 gl: 9369    gy: 14027     g: 72632     resid: 150000
# Brandle analyzed rep-level data, so the residual variance is different.
# The other components are matched by the following analysis.

```r
libs(lme4)
libs(lucid)
dat$year <- factor(dat$year)
m1 <- lmer(yield ~ year + loc + year:loc + (1|gen) +
(1|gen:loc) + (1|gen:year), data=dat)
vc(m1)
```
## Description
Switchback experiment on dairy cattle, milk yield for two treatments

## Usage
data("brandt.switchback")

## Format
A data frame with 30 observations on the following 5 variables.

- **group**: group: A, B
- **cow**: cow, 10 levels
- **trt**: treatment, 2 levels
- **period**: period, 3 levels
- **yield**: milk yield, pounds

## Details
In this experiment, 10 cows were selected from the Iowa State College Holstein-Friesian herd and divided into two equal groups. Care was taken to have the groups as nearly equal as possible with regard to milk production, stage of gestation, body weight, condition and age. These cows were each given 10 pounds of timothy hay and 30 pounds of corn silage daily but were fed different grain mixtures. Treatment T1, then, consisted of feeding a grain mixture of 1 part of corn and cob meal to 1 part of ground oats, while treatment T2 consisted of feeding a grain mixture of 4 parts corn and cob meal, 4 parts of ground oats and 3 parts of gluten feed. The three treatment periods covered 105 days – three periods of 35 days each. The yields for the first 7 days of each period were not considered because of the possible effect of the transition from one treatment to the other. The data, together with sums and differences which aid in the calculations incidental to testing, are given in table 2.

It seems safe to conclude that the inclusion of gluten feed in the grain mixture fed in a timothy hay ration to Holstein-Friesian cows increased the production of milk. The average increase was 21.7 pounds per cow for a 28-day period.
**Source**


**Examples**

```r
## Not run:

library(agridat)

data(brandt.switchback)
dat <- brandt.switchback

# In each period, treatment 2 is slightly higher
# bwplot(yield~trt|period,dat, layout=c(3,1), main="brandt.switchback",
# xlab="Treatment", ylab="Milk yield")

# Yield at period 2 (trt T2) is above the trend in group A,
# below the trend (trt T1) in group B.
# Equivalently, treatment T2 is above the trend line

# Similar to Brandt Table 10
m1 <- aov(yield~period+group+cow:group+period:group, data=dat)
anova(m1)

## End(Not run)
```

**bridges.cucumber**

Multi-environment trial of cucumbers in a latin square design

**Description**

Cucumber yields in latin square design at two locs.

**Format**

A data frame with 32 observations on the following 5 variables.

- `loc` location
- `gen` genotype/cultivar
- `row` row
- `col` column
- `yield` weight of marketable fruit per plot
Details

Conducted at Clemson University in 1985. Four cucumber cultivars were grown in a Latin square design at Clemson, SC, and Tifton, GA.

Separate variances are modeled each location.

Plot dimensions are not given.

Bridges (1989) used this data to illustrate fitting a heterogeneous mixed model.

Used with permission of William Bridges.

Source


Examples

```r
## Not run:

library(agridat)
data(bridges.cucumber)
dat <- bridges.cucumber
dat <- transform(dat, rowf=factor(row), colf=factor(col))

libs(desplot)
desplot(dat, yield~col*row|loc,
     # aspect unknown
     text=gen, cex=1,
     main="bridges.cucumber")

# Graphical inference test for heterogenous variances
libs(nullabor)
# Create a lineup of datasets
fun <- null_permute("loc")
dat20 <- lineup(fun, dat, n=20, pos=9)

# Now plot
libs(lattice)
bwpplot(yield ~ loc|factor(.sample), dat20,
     main="bridges.cucumber - graphical inference")

libs(asreml) # asreml4

## Random row/col/resid. Same as Bridges 1989, p. 147
m1 <- asreml(yield ~ 1 + gen + loc + loc:gen,
             random = ~ rowf:loc + colf:loc, data=dat)

libs(lucid)
lucid::vc(m1)
## effect component std.error z.ratio bound
## rowf:loc 31.62 23.02 1.4 P 0
broadbalk.wheat

---

**Description**

Long term wheat yields on Broadbalk fields at Rothamsted.

**Format**

A data frame with 1258 observations on the following 4 variables.

- **year**: year
- **plot**: plot
- **grain**: grain yield, tonnes
- **straw**: straw yield, tonnes

---

## colf:loc
18.08 15.32 1.2 P 0
## units(R)
31.48 12.85 2.4 P 0

Random row/col/resid at each loc. Matches p. 147

```r
m2 <- asreml(yield ~ 1 + gen + loc + loc:gen,
             random = ~ at(loc):rowf + at(loc):colf, data=dat,
             resid = ~ dsum(~ units|loc))
lucid::vc(m2)
```

<table>
<thead>
<tr>
<th>effect</th>
<th>component</th>
<th>std.error</th>
<th>z.ratio</th>
<th>bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>at(loc, Clemson):rowf</td>
<td>32.32</td>
<td>36.58</td>
<td>0.88</td>
<td>P 0</td>
</tr>
<tr>
<td>at(loc, Tifton):rowf</td>
<td>30.92</td>
<td>28.63</td>
<td>1.1</td>
<td>P 0</td>
</tr>
<tr>
<td>at(loc, Clemson):colf</td>
<td>22.55</td>
<td>28.78</td>
<td>0.78</td>
<td>P 0</td>
</tr>
<tr>
<td>at(loc, Tifton):colf</td>
<td>13.62</td>
<td>14.59</td>
<td>0.93</td>
<td>P 0</td>
</tr>
<tr>
<td>loc_Clemson(R)</td>
<td>46.85</td>
<td>27.05</td>
<td>1.7</td>
<td>P 0</td>
</tr>
<tr>
<td>loc_Tifton(R)</td>
<td>16.11</td>
<td>9.299</td>
<td>1.7</td>
<td>P 0</td>
</tr>
</tbody>
</table>

```r
predict(m2, data=dat, classify='loc:gen')$pvals
```

<table>
<thead>
<tr>
<th>loc</th>
<th>gen</th>
<th>predicted.value</th>
<th>std.error</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clemson Dasher</td>
<td>45.6</td>
<td>5.04</td>
<td>Estimable</td>
</tr>
<tr>
<td>2</td>
<td>Clemson Guardian</td>
<td>31.6</td>
<td>5.04</td>
<td>Estimable</td>
</tr>
<tr>
<td>3</td>
<td>Clemson Poinsett</td>
<td>21.4</td>
<td>5.04</td>
<td>Estimable</td>
</tr>
<tr>
<td>4</td>
<td>Clemson Sprint</td>
<td>26</td>
<td>5.04</td>
<td>Estimable</td>
</tr>
<tr>
<td>5</td>
<td>Tifton Dasher</td>
<td>50.5</td>
<td>3.89</td>
<td>Estimable</td>
</tr>
<tr>
<td>6</td>
<td>Tifton Guardian</td>
<td>38.7</td>
<td>3.89</td>
<td>Estimable</td>
</tr>
<tr>
<td>7</td>
<td>Tifton Poinsett</td>
<td>33</td>
<td>3.89</td>
<td>Estimable</td>
</tr>
<tr>
<td>8</td>
<td>Tifton Sprint</td>
<td>39.2</td>
<td>3.89</td>
<td>Estimable</td>
</tr>
</tbody>
</table>

# Is a heterogeneous model justified? Maybe not.

```r
m1$loglik
## -67.35585
m2$loglik
## -66.35621
```

## End(Not run)
Details

Note: This data is only 1852-1925. You can find recent data for these experiments at the Electronic Rothamsted Archive: https://www.era.rothamsted.ac.uk/

Rothamsted Experiment station conducted wheat experiments on the Broadbalk Fields beginning in 1844 with data for yields of grain and straw collected from 1852 to 1925. Ronald Fisher was hired to analyze data from the agricultural trials. Organic manures and inorganic fertilizer treatments were applied in various combinations to the plots.

N1 is 48kg, N1.5 is 72kg, N2 is 96kg, N4 is 192kg nitrogen.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2b</td>
<td>manure</td>
</tr>
<tr>
<td>3</td>
<td>No fertilizer or manure</td>
</tr>
<tr>
<td>5</td>
<td>P K Na Mg (No N)</td>
</tr>
<tr>
<td>6</td>
<td>N1 P K Na Mg</td>
</tr>
<tr>
<td>7</td>
<td>N2 P K Na Mg</td>
</tr>
<tr>
<td>8</td>
<td>N3 P K Na Mg</td>
</tr>
<tr>
<td>9</td>
<td>N1* P K Na Mg since 1894; 9A and 9B received different treatments 1852-93</td>
</tr>
<tr>
<td>10</td>
<td>N2</td>
</tr>
<tr>
<td>11</td>
<td>N2 P</td>
</tr>
<tr>
<td>12</td>
<td>N2 P Na*</td>
</tr>
<tr>
<td>13</td>
<td>N2 P K</td>
</tr>
<tr>
<td>14</td>
<td>N2 P Mg*</td>
</tr>
<tr>
<td>15</td>
<td>N2 P K Na Mg (timing of N application different to other plots, see below)</td>
</tr>
<tr>
<td>16</td>
<td>N4 P K Na Mg 1852-64; unmanured 1865-83; N2*P K Na Mg since 1884</td>
</tr>
<tr>
<td>17</td>
<td>N2 applied in even years; P K Na Mg applied in odd years</td>
</tr>
<tr>
<td>18</td>
<td>N2 applied in odd years; P K Na Mg applied in even years</td>
</tr>
<tr>
<td>19</td>
<td>N1.5 P and rape cake 1852-78, 1879-1925 rape cake only</td>
</tr>
</tbody>
</table>

Source


Retrieved from https://lib.stat.cmu.edu/datasets/Andrews/

References


Examples

```r
## Not run:

library(agridat)
data(broadbalk.wheat)
dat <- broadbalk.wheat
libs(lattice)
```
bryan.corn.uniformity

# xyplot(grain~straw|plot, dat, type=c('p','smooth'), as.table=TRUE,
#       main="broadbalk.wheat")
xyplot(grain~year|plot, dat, type=c('p','smooth'), as.table=TRUE,
       main="broadbalk.wheat") # yields are decreasing

# See the treatment descriptions to understand the patterns
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(grain~year*plot, dat, main="broadbalk.wheat: Grain", col.regions=redblue)

## End(Not run)

bryan.corn.uniformity  Uniformity trial of corn at 3 locations in Iowa.

Description
Uniformity trial of corn at 3 locations in Iowa.

Usage
data("bryan.corn.uniformity")

Format
A data frame with 1728 observations on the following 4 variables.

expt  experiment (variety/orientation)
row   row
col   column
yield yield, pounds per plot

Details
Three varieties of corn were planted. Each experiment was 48 rows, each row 48 hills long, .65 acres. (A "hill" is a single hole with possibly multiple seeds). Spacing between the hills would be sqrt(43560 sq ft * .64) / 48 = 3.5 feet.

In the experiment code, K=Krug, I=Iodent, M=McCulloch (varieties of corn), 23=1923, 25=1925, E=East/West, N=North/South.

Each experiment was aggregated into experimental units by combining 8 hills, both in East/West direction and also in North/South direction. Thus, each field is represented twice in the data.

Source
Arthur Bryan (1933). Factors Affecting Experimental Error in Field Plot Tests With Corn. Agricultural Experiment Station, Iowa State College. Tables 22-27. https://hdl.handle.net/2027/uiug.30112019568168
buntaran.wheat


Description


Usage

data("buntaran.wheat")

Format

A data frame with 1069 observations on the following 7 variables.

zone  Geographic zone: south, middle, north
loc  Location
rep  Block replicate (up to 4)
Data from wheat trials in Sweden in 2016. The experiments in each location were multi-rep with incomplete blocks in an alpha design.

Electronic data are from the online supplement of Buntaran (2020) and also from the "init" package at https://github.com/Flavjack/inti.

### Source


### Examples

```r
data(buntaran.wheat)
library(agridat)
dat <- buntaran.wheat
library(lattice)
bwplot(yield~loc|zone, dat, layout=c(1,3),
      scales=list(x=list(rot=90)),
      main="buntaran.wheat")
```

### Description

Incomplete block alpha design

### Usage

```r
data("burgueno.alpha")
```
Format

A data frame with 48 observations on the following 6 variables.

- `rep`: rep, 3 levels
- `block`: block, 12 levels
- `row`: row
- `col`: column
- `gen`: genotype, 16 levels
- `yield`: yield

Details

A field experiment with 3 reps, 4 blocks per rep, laid out as an alpha design.
The plot size is not given.
Electronic version of the data obtained from CropStat software.
Used with permission of Juan Burgueno.

Source


Examples

```r
## Not run:

library(agridat)
data(burgueno.alpha)
dat <- burgueno.alpha

library(desplot)
desplot(dat, yield~col*row, 
       out1=rep, out2=block, # aspect unknown
       text=gen, cex=1, shorten="none",
       main='burgueno.alpha')

library(lme4,lucid)
# Inc block model
m0 <- lmer(yield ~ gen + (1|rep/block), data=dat)
vce(m0) # Matches Burgueno p. 26
## grp var1 var2 vcov sdcor
## block:rep (Intercept) <NA> 86900 294.8
## rep (Intercept) <NA> 200900 448.2
## Residual <NA> <NA> 133200 365

library(asreml) # asreml4
```
dat <- transform(dat, xf=factor(col), yf=factor(row))
dat <- dat[order(dat$xf, dat$yf),]

# Sequence of models on page 36
m1 <- asreml(yield ~ gen, data=dat)
m1$loglik # -232.13
m2 <- asreml(yield ~ gen, data=dat, random = ~ rep)
m2$loglik # -223.48

# Inc Block model
m3 <- asreml(yield ~ gen, data=dat, random = ~ rep/block)
m3$loglik # -221.42
m3$coef$fixed # Matches solution on p. 27

# AR1xAR1 model
m4 <- asreml(yield ~ 1 + gen, data=dat, resid = ~ar1(xf):ar1(yf))
m4$loglik # -221.47
plot(varioGram(m4), main="burgueno.alpha") # Figure 1

m5 <- asreml(yield ~ 1 + gen, data=dat, random= ~ yf, resid = ~ar1(xf):ar1(yf))
m5$loglik # -220.07

m6 <- asreml(yield ~ 1 + gen + pol(yf,-2), data=dat, resid = ~ar1(xf):ar1(yf))
m6$loglik # -204.64

m7 <- asreml(yield ~ 1 + gen + lin(yf), data=dat, random= ~ spl(yf), resid = ~ar1(xf):ar1(yf))
m7$loglik # -212.51

m8 <- asreml(yield ~ 1 + gen + lin(yf), data=dat, random= ~ spl(yf))
m8$loglik # -213.91

# Polynomial model with predictions
m9 <- asreml(yield ~ 1 + gen + pol(yf,-2) + pol(xf,-2), data=dat, random= ~ spl(yf), resid = ~ar1(xf):ar1(yf))
m9 <- update(m9)
m9$loglik # -191.44 vs -189.61
p9 <- predict(m9, classify="gen:xf:yf", levels=list(xf=1,yf=1))
p9

m10 <- asreml(yield ~ 1 + gen + lin(yf)+lin(xf), data=dat, resid = ~ar1(xf):ar1(yf))
m10$loglik # -211.56
burgueno.rowcol

### Description

Row-column design

### Usage

```r
data("burgueno.rowcol")
```

### Format

A data frame with 128 observations on the following 5 variables.

- **rep** rep, 2 levels
- **row** row
- **col** column
- **gen** genotype, 64 levels
- **yield** yield, tons/ha

### Details

A field experiment with two contiguous replicates in 8 rows, 16 columns.
The plot size is not given.
Electronic version of the data obtained from CropStat software.
Used with permission of Juan Burgueno.
Source


Examples

```r
## Not run:
library(agridat)
data(burgueno.rowcol)
dat <- burgueno.rowcol

# Two contiguous reps in 8 rows, 16 columns
libs(desplot)
desplot(dat, yield ~ col*row,
        out1=rep, # aspect unknown
text=gen, shorten="none", cex=.75,
        main="burgueno.rowcol")

libs(lme4,lucid)

# Random rep, row and col within rep
# m1 <- lmer(yield ~ gen + (1|rep) + (1|rep:row) + (1|rep:col), data=dat)
# vc(m1) # Match components of Burgueno p. 40
## grp var1 var2 vcov sdcor
## rep:col (Intercept) <NA> 0.2189 0.4679
## rep:row (Intercept) <NA> 0.1646 0.4057
## rep (Intercept) <NA> 0.1916 0.4378
## Residual <NA> <NA> 0.1796 0.4238

libs(asreml) ## asreml4

# AR1 x AR1 with linear row/col effects, random spline row/col
dat <- transform(dat, xf=factor(col), yf=factor(row))
dat <- dat[order(dat$xf,dat$yf),]
m2 <- asreml(yield ~ gen + lin(yf) + lin(xf), data=dat,
            random = ~ spl(yf) + spl(xf),
            resid = ~ ar1(xf):ar1(yf))
m2 <- update(m2) # More iterations

## Scaling of spl components has changed in asreml from old versions
libs(lucid)

vc(m2) # Match Burgueno p. 42
## effect component std.error z.ratio bound
## spl(yf) 0.09077 0.08252 1.1 P 0
## spl(xf) 0.08107 0.08209 0.99 P 0
## xf:yf(R) 0.1482 0.03119 4.8 P 0
## xf:yf!xf!cor 0.1152 0.2269 0.51 U 0.1
## xf:yf!yf!cor 0.009467 0.2414 0.039 U 0.9

plot(varioGram(m2), main="burgueno.rowcol")
```
burgueno.unreplicated  

Field experiment with unreplicated genotypes plus one repeated check.

Description
Field experiment with unreplicated genotypes plus one repeated check.

Usage
data("burgueno.unreplicated")

Format
A data frame with 434 observations on the following 4 variables.

gen  genotype, 281 levels
col  column
row  row
yield  yield, tons/ha

Details
A field experiment with 280 new genotypes. A check genotype is planted in every 4th column.
The plot size is not given.
Electronic version of the data obtained from CropStat software.
Used with permission of Juan Burgueno.

Source

Examples
## Not run:
library(agridat)
data(burgueno.unreplicated)
dat <- burgueno.unreplicated

# Define a 'check' variable for colors
dat$check <- ifelse(dat$gen=="G000", 2, 1)
# Every fourth column is the 'check' genotype
libs(desplot)
Multi-environment trial of maize with pedigrees

**Description**


**Format**

A data frame with 245 observations on the following 5 variables.
Details

Ten inbreds were crossed to produce a diallel without reciprocals. The 45 F1 crosses were evaluated along with 4 checks in a triple-lattice 7x7 design. Pink stem borer infestation was natural.

Experiments were performed in 1995 and 1996 at three sites in northwestern Spain: Pontevedra (42 deg 24 min N, 8 deg 38 min W, 20 m over sea), Pontecaldelas (42 deg 23 N, 8 min 32 W, 300 m above sea), Ribadumia (42 deg 30 N, 8 min 46 W, 50 m above sea).

A two-letter location code and the year are concatenated to define the environment.

The average number of larvae per plant in each environment:

<table>
<thead>
<tr>
<th>Env</th>
<th>Larvae</th>
</tr>
</thead>
<tbody>
<tr>
<td>pc95</td>
<td>0.54</td>
</tr>
<tr>
<td>pc96</td>
<td>0.91</td>
</tr>
<tr>
<td>ri96</td>
<td>1.78</td>
</tr>
<tr>
<td>pv95</td>
<td>2.62</td>
</tr>
<tr>
<td>pv96</td>
<td>3.35</td>
</tr>
</tbody>
</table>

Used with permission of Ana Butron.

Source


Examples

```r
## Not run:
library(agridat)
data(butron.maize)
dat <- butron.maize

libs(reshape2)
mat <- acast(dat, gen~env, value.var='yield')
mat <- sweep(mat, 2, colMeans(mat))
mat.svd <- svd(mat)
# Calculate PC1 and PC2 scores as in Table 4 of Butron
# Comment out to keep Rcmd check from choking on '
# round(mat.svd$u[,1:2]
```
biplot(princomp(mat), main="butron.maize", cex=.7) # Figure 1 of Butron

if(0){
  # Fixme: This section is broken because synbreed has been removed from CRAN
  # Here we see if including pedigree information is helpful for a
  # multi-environment model
  # Including the pedigree provided little benefit

  # Create the pedigree
  ped <- dat[, c('gen','male','female')]
  ped <- ped[!duplicated(ped),] # remove duplicates
  unip <- unique(c(ped$male, ped$female)) # Unique parents
  unip <- unip[!is.na(unip)]
  # We have to define parents at the TOP of the pedigree
  ped <- rbind(data.frame(gen=c("Dent","Flint"), # genetic groups
    male=c(0,0),
    female=c(0,0)),
    data.frame(gen=c("A589","A637","A661","CM105","EP28",
     "EP31","EP42","F7","PB60","Z77016"),
    male=rep(c("Dent","Flint"),each=5),
    female=rep(c("Dent","Flint"),each=5)),
    ped)
  ped[is.na(ped$male),'male'] <- 0
  ped[is.na(ped$female),'female'] <- 0
}

# asreml 4
if(0){
  libs(asreml)
  ped.ainv <- ainverse(ped)
  m0 <- asreml(yield ~ 1+env, data=dat, random = ~ gen)
  m1 <- asreml(yield ~ 1+env, random = ~ vm(gen, ped.ainv), data=dat)
  m2 <- update(m1, random = ~ idv(env):vm(gen, ped.ainv))
  m3 <- update(m2, random = ~ diag(env):vm(gen, ped.ainv))
  m4 <- update(m3, random = ~ fa(env,1):vm(gen, ped.ainv))
  #summary(m0)$aic
  #summary(m4)$aic
  ## df  AIC
  ## m0  2 229.4037
  ## m1  2 213.2487
  ## m2  2 260.6156
  ## m3  6 296.8061
  ## m4 11 218.1568

  p0 <- predict(m0, data=dat, classify="gen")$pvals
  p1 <- predict(m1, data=dat, classify="gen")$pvals
  ppar <- p1[1:12,] # parents
  p1 <- p1[-c(1:12),] # remove parents
  # Careful! Need to manually sort the predictions
```r
p0 <- p0[order(as.character(p0$gen)),]
p1 <- p1[order(as.character(p1$gen)),]

# lims <- range(c(p0$pred, p1$pred)) * c(.95,1.05)
lims <- c(6,8.25)  # zoom in on the higher-yielding hybrids
plot(p0$predicted.value, p1$predicted.value,
pch="", xlim=lims, ylim=lims, main="butron.maize",
xlab="BLUP w/o pedigree", ylab="BLUP with pedigree")
abline(0,1,col="lightgray")
text(x=p0$predicted.value, y=p1$predicted.value,
p0$gen, cex=.5, srt=-45)
text(x=min(lims), y=p1par$predicted.value, p1par$gen, cex=.5, col="red")
round( cor(p0$predicted.value, p1$predicted.value), 3)
# Including the pedigree provided very little change
```

```r
## End(Not run)
```

---

### byers.apple

**Diameters of apples**

**Description**

Measurements of the diameters of apples

**Format**

A data frame with 480 observations on the following 6 variables.

- `tree` tree, 10 levels
- `apple` apple, 24 levels
- `size` size of apple
- `appleid` unique id number for each apple
- `time` time period, 1-6 = (week/2)
- `diameter` diameter, inches

**Details**

Experiment conducted at the Winchester Agricultural Experiment Station of Virginia Polytechnic Institute and State University. Twentyfive apples were chosen from each of ten apple trees. Of these, there were 80 apples in the largest size class, 2.75 inches in diameter or greater. The diameters of the apples were recorded every two weeks over a 12-week period.

**Source**

Examples

library(agridat)
data(byers.apple)
dat <- byers.apple

libs(lattice)
xyplot(diameter ~ time | factor(appleid), data=dat, type=c('p','l'),
       strip=strip.custom(par.strip.text=list(cex=.7)),
       main="byers.apple")

## Not run:
# Overall fixed linear trend, plus random intercept/slope deviations
# for each apple. Observations within each apple are correlated.
libs(nlme)
libs(lucid)
m1 <- lme(diameter ~ 1 + time, data=dat,
          random = ~ time|appleid, method='ML',
          cor = corAR1(0, form=- time|appleid),
          na.action=na.omit)
vc(m1)
## effect      variance     stddev     corr
## (Intercept) 0.007354 0.08575   NA
## time        0.00003632 0.006027 0.83
## Residual    0.0004555 0.02134   NA

## End(Not run)

caribbean.maize  Multi-environment trial of maize with fertilization

Description

Maize fertilization trial on Antigua and St. Vincent.

Format

A data frame with 612 observations on the following 7 variables.

isle  island, 2 levels
site  site
block block
plot  plot, numeric
trt   treatment factor
ears  number of ears harvested
yield yield in kilograms
Details

Antigua is a coral island in the Caribbean with sufficient level land for experiments and a semi-arid climate, while St. Vincent is volcanic and level areas are uncommon, but the rainfall can be seasonally heavy.

Plots were 16 feet by 18 feet. A central area 12 feet by 12 feet was harvested and recorded.

The number of ears harvested was only recorded on the isle of Antigua.

The digits of the treatment represent the levels of nitrogen, phosphorus, and potassium fertilizer, respectively.

The TEAN site suffered damage from goats on plot 27, 35 and 36. The LFAN site suffered damage from cattle on one boundary–plots 9, 18, 27, 36.

Source


Retrieved from https://lib.stat.cmu.edu/datasets/Andrews/

References

Also in the DAAG package as data sets antigua, stVincent.

Examples

library(agridat)
data(caribbean.maize)
dat <- caribbean.maize

# Yield and ears are correlated
libs(lattice)
xyplot(yield~ears|site, dat, ylim=c(0,10), subset=isle=="Antigua",
       main="caribbean.maize - Antiqua")

# Some locs show large response to nitrogen (as expected), e.g. UISV, OOSV
dotplot(trt~yield|site, data=dat, main="caribbean.maize treatment response")

# The pattern is a bit hard to see, so we split the treatment factor
# into separate factors, and group sites by island
dat <- transform(dat, N=factor(substring(trt,2,2)),
                 P=factor(substring(trt,3,3)),
                 K=factor(substring(trt,4,4)))
dat <- transform(dat, env=paste(substring(isle,1,1),site,sep="-"))
# Now we can see the strong N*site interaction
bwplot(yield~N|env, dat,
       main="caribbean.maize", xlab="nitrogen")
Germination of alfalfa seeds at various salt concentrations

Description
Germination of alfalfa seeds at various salt concentrations

Usage
data("carlson.germination")

Format
A data frame with 120 observations on the following 3 variables.

- gen: genotype factor, 15 levels
- germ: germination percent, 0-100
- nacl: salt concentration percent, 0-2

Details
Data are means averaged over 5, 10, 15, and 20 day counts. Germination is expressed as a percent of the no-salt control to account for differences in germination among the cultivars.

Source

Examples
```r
## Not run:
library(agridat)
data(carlson.germination)
dat <- carlson.germination
dat$germ <- dat$germ/100 # Convert to percent

# Separate response curve for each genotype.
# Really, we should use a glmm with random int/slope for each genotype
m1 <- glm(germ~ 0 + gen*nacl, data=dat, family=quasibinomial)

# Plot data and fitted model
library(latticeExtra)
newd <- data.frame(expand.grid(gen=levels(dat$gen), nacl=seq(0,2,length=100)))
newd$pred <- predict(m1, newd, type="response")

xyplot(germ~nacl|gen, dat, as.table=TRUE, main="carlson.germination",
      xlab="Percent NaCl", ylab="Fraction germinated") +
xyplot(pred~nacl|gen, newd, type='l', grid=list(h=1,v=0))
```
# Calculate LD50 values. Note, Carlson et al used quadratics, not glm.
# MASS::dose.p cannot handle multiple slopes, so do a separate fit for
# each genotype. Results are vaguely similar to Carlson table 5.
## libs(MASS)
## for(ii in unique(dat$gen)){
## cat("n", ii, "n")
## mm <- glm(germ ~ 1 + nacl, data=dat, subset=gen==ii, family=quasibinomial(link="probit"))
## print(dose.p(mm))
## }
## Dose  SE
## Anchor 1.445728 0.05750418
## Apollo 1.305804 0.04951644
## Baker  1.444153 0.07653989
## Drylander 1.351201 0.03111795
## Grimm  1.395735 0.04206377
##
## End(Not run)

carmer.density  Nonlinear maize yield-density model

Description
Nonlinear maize yield-density model.

Format
A data frame with 32 observations on the following 3 variables.

<table>
<thead>
<tr>
<th>gen</th>
<th>genotype/hybrid, 8 levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>pop</td>
<td>population (plants)</td>
</tr>
<tr>
<td>yield</td>
<td>yield, pounds per hill</td>
</tr>
</tbody>
</table>

Details
Eight single-cross hybrids were in the experiment–Hy2xOh7 and WF9xC103 were included because it was believed they had optimum yields at relatively high and low populations. Planted in 1963. Plots were thinned to 2, 4, 6, 8 plants per hill, giving densities 8, 16, 24, 32 thousand plants per acre. Hills were in rows 40 inches apart. One hill = 1/4000 acre. Split-plot design with 5 reps, density is main plot and subplot was hybrid.

Source
Examples

```r
## Not run:
library(agridat)
data(carmer.density)
dat <- carmer.density
dat$gen <- factor(dat$gen, levels=c('Hy2x0h7', 'WF9xC103', 'R61x187-2',
                                          'WF9x38-11', 'WF9xB14', 'C103xB14',
                                          '0h43xB37', 'WF9xH60'))

# Separate analysis for each hybrid
# Model: y = x * a * k^x. Table 1 of Carmer and Jackobs.
out <- data.frame(a=rep(NA,8), k=NA)
preds <- NULL
rownames(out) <- levels(dat$gen)
newdat <- data.frame(pop=seq(2,8,by=.1))
for(i in levels(dat$gen)){
  print(i)
dati <- subset(dat, gen==i)
  mi <- nls(yield ~ pop * a * k^pop, data=dati, start=list(a=10,k=1))
  out[i, ] <- mi$m$getPars()
  # Predicted values
  pi <- cbind(gen=i, newdat, pred= predict(mi, newdat=newdat))
preds <- rbind(preds, pi)
}

# Optimum plant density is -1/log(k)
out$pop.opt <- -1/log(out$k)
round(out, 3)
## a  k  pop.opt
## Hy2x0h7 0.782 0.865 6.875
## WF9xC103 1.039 0.825 5.192
## R61x187-2 0.998 0.798 4.441
## WF9x38-11 1.042 0.825 5.203
## WF9xB14 1.067 0.806 4.647
## C103xB14 0.813 0.860 6.653
## 0h43xB37 0.673 0.862 6.740
## WF9xH60 0.858 0.854 6.358

# Fit an overall fixed-effect with random deviations for each hybrid.
libsnlme
m1 <- nlme(yield ~ pop * a * k^pop,
           fixed = a + k ~ 1,
           random = a + k ~ 1|gen,
           data=dat, start=c(a=10,k=1))
# summary(m1) # Random effect for 'a' probably not needed

libslatticeExtra
# Plot Data, fixed-effect prediction, random-effect prediction.
pdat <- expand.grid(gen=levels(dat$gen), pop=seq(2,8,length=50))
pdat$pred <- predict(m1, pdat)
```
cate.potassium

```
pdat$predf <- predict(m1, pdat, level=0)

xyplot(yield~pop|gen, dat, pch=16, as.table=TRUE,
main="carmer.density models",
key=simpleKey(text=c("Data", "Fixed effect","Random effect"),
     col=c("blue", "red","darkgreen"), columns=3, points=FALSE)) +
xyplot(predf~pop|gen, pdat, type=
     "Var","Var","Var", as.table=TRUE, col="red") +
xyplot(pred~pop|gen, pdat, type='l', col="darkgreen", lwd=2)
```

## End(Not run)

cate.potassium Relative cotton yield for different soil potassium concentrations

**Description**

Relative cotton yield for different soil potassium concentrations

**Format**

A data frame with 24 observations on the following 2 variables.

- **yield**: Relative yield
- **potassium**: Soil potassium, ppm

**Details**

Cate & Nelson used this data to determine the minimum optimal amount of soil potassium to achieve maximum yield.

Note, Fig 1 of Cate & Nelson does not match the data from Table 2. It sort of appears that points with high-concentrations of potassium were shifted left to a truncation point. Also, the calculations below do not quite match the results in Table 1. Perhaps the published data were rounded?

**Source**


**Examples**

```
library(agridat)
data(cate.potassium)
dat <- cate.potassium
names(dat) <- c('y','x')

CateNelson <- function(dat){
```
dat <- dat[order(dat$x),] # Sort the data by x
x <- dat$x
y <- dat$y

# Create a data.frame to store the results
out <- data.frame(x=NA, mean1=NA, css1=NA, mean2=NA, css2=NA, r2=NA)

css <- function(x) { var(x) * (length(x)-1) }
tcss <- css(y) # Total corrected sum of squares

for(i in 2:(length(y)-2)){
y1 <- y[1:i]
y2 <- y[-(1:i)]
out[i, 'x'] <- x[i]
out[i, 'mean1'] <- mean(y1)
out[i, 'mean2'] <- mean(y2)
out[i, 'css1'] <- css1 <- css(y1)
out[i, 'css2'] <- css2 <- css(y2)
out[i, 'r2'] <- ( tcss - (css1+css2)) / tcss
}
return(out)
}
Format

A data frame with 81 observations on the following 10 variables.

yield  yield
block  block
row  row position
col  column position
trt  treatment code
N  nitrogen treatment: 0, 1, 2
P  phosphorous treatment: 0, 1, 2
K  potassium treatment: 0, 1, 2
B  bagasse treatment: 0, 1, 2
F  filter press mud treatment: 0, 1, 2

Details

An experiment grown in 1949 at the Worthy Park Estate in Jamaica.
Nitrogen was applied as sulphate of ammonia at 0, 3, 6 hundred-weight per acre.
Phosphorous was applied as superphosphate at 0, 4, 8 hundred-weight per acre.
Potassium was applied as muriate of potash at 0, 1, 2 hundred-weight per acre.
Bagasse applied pre-plant at 0, 20, 40 tons per acre.
Filter press mud applied pre-plant at 0, 10, 20 tons per acre.
Each plot was 18 yards long by 6 yards (3 rows) wide. Plots were arranged in nine columns of nine,
a 2-yard space separating plots along the rows and two guard rows separating plots across the rows.
Field width: 6 yards * 9 plots + 4 yards * 8 gaps = 86 yards
Field length: 18 yards * 9 plots + 2 yards * 8 gaps = 178 yards

Source


References

None

Examples

```
## Not run:

library(agridat)
data(chinloy.fractionalfactorial)
dat <- chinloy.fractionalfactorial
```
# Treatments are coded with levels 0,1,2. Make sure they are factors
dat <- transform(dat,
        N=factor(N), P=factor(P), K=factor(K), B=factor(B), F=factor(F))

# Experiment layout
libs(desplot)
desplot(dat, yield ~ col*row,
       out1=block, text=trt, shorten="no", cex=0.6,
       aspect=178/86,
       main="chinloy.fractionalfactorial")

# Main effect and some two-way interactions. These match Chinloy table 6.
# Not sure how to code terms like P^2K=B^2F
m1 <- aov(yield ~ block + N + P + K + B + F + N:P + N:K + N:B + N:F, dat)
anova(m1)

## End(Not run)

christidis.competition

*Competition between varieties in cotton*

**Description**

Competition between varieties in cotton, measurements taken for each row.

**Usage**

`data("christidis.competition")`

**Format**

A data frame with 270 observations on the following 8 variables.

- `plot` plot
- `plotrow` row within plot
- `block` block
- `row` row, only 1 row
- `col` column
- `gen` genotype
- `yield` yield, kg
- `height` height, cm
Details

Nine genotypes/varieties of cotton were used in a variety test. The plots were 100 meters long and 2.40 meters wide, each plot having 3 rows 0.80 meters apart.

The layout was an RCB of 5 blocks, each block having 2 replicates of every variety (with the original intention of trying 2 seed treatments). Each row was harvested/weighed separately. After the leaves of the plants had dried up and fallen, the mean height of each row was measured.

Christidis found significant competition between varieties, but not due to height differences. Crude analysis.

TODO: Find a better analysis of this data which incorporates field trends AND competition effects, maybe including a random effect for border rows of all genotype pairs (as neighbors)?

Source


References

None

Examples

```r
## Not run:

library(agridat)
data(christidis.competition)
dat <- christidis.competition

# Match Christidis Table 2 means
# aggregate(yield ~ gen, aggregate(yield ~ gen+plot, dat, sum), mean)

# Each RCB block has 2 replicates of each genotype
# with(dat, table(block,gen))

libs(lattice)

# Tall plants yield more
# xyplot(yield ~ height|gen, data=dat)

# Huge yield variation across field. Also heterogeneous variance.
xyplot(yield ~ col, dat, group=gen, auto.key=list(columns=5),
        main="christidis.competition")

libs(mgcv)
if(is.element("package:gam", search())) detach("package:gam")
# Simple non-competition model to remove main effects
m1 <- gam(yield ~ gen + s(col), data=dat)
p1 <- as.data.frame(predict(m1, type="terms"))
names(p1) <- c('geneff','coleff')
```
dat2 <- cbind(dat, p1)
dat2 <- transform(dat2, res=yield-geneff-coleff)
libs(lattice)
xyplot(res ~ col, data=dat2, group=gen,
       main="christidis.competition - residuals")

## End(Not run)
## christidis.wheat.uniformity

### Uniformity trial of wheat

Uniformity trial of wheat at Cambridge, UK in 1931.

### Usage

```r
data("christidis.wheat.uniformity")
```

### Format

A data frame with 288 observations on the following 3 variables.

- `row` row
- `col` column
- `yield` yield

### Details

Two blocks, 24 rows each. In block A, each 90-foot row was divided into 12 units, each unit 7.5 feet long. Rows were 8 inches wide.
- Field width: 12 units * 7.5 feet = 90 feet
- Field length: 24 rows * 8 inches = 16 feet
Source

References
None

Examples

```r
## Not run:
library(agridat)
data(christidis.wheat.uniformity)
dat <- christidis.wheat.uniformity

# sum(dat$yield) # Matches Christidis
libs(desplot)
desplot(dat, yield ~ col*row,
     flip=TRUE, aspect=16/90, # true aspect
     main="christidis.wheat.uniformity")

## End(Not run)
```

cleveland.soil

*Soil resistivity in a field*

**Description**
Soil resistivity in a field

**Format**
A data frame with 8641 observations on the following 5 variables.

- northing: y ordinate
- easting: x ordinate
- resistivity: Soil resistivity, ohms
- is.ns: Indicator of north/south track
- track: Track number

**Details**
Resistivity is related to soil salinity.
library(agridat)

data(cleveland.soil)
dat <- cleveland.soil

# Similar to Cleveland fig 4.64
## libs(latticeExtra)
## redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
## levelplot(resistivity ~ easting + northing, data = dat,
##       col.regions=redblue,
##       panel=panel.levelplot.points,
##       aspect=2.4, xlab= "Easting (km)", ylab= "Northing (km)",
##       main="cleveland")

# 2D loess plot. Cleveland fig 4.68
sg1 <- expand.grid(easting = seq(.15, 1.410, by = .02),
                  northing = seq(.150, 3.645, by = .02))
lo1 <- loess(resistivity~easting*northing, data=dat, span = 0.1, degree = 2)
fit1 <- predict(lo1, sg1)
libs(lattice)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(fit1 ~ sg1$easting * sg1$northing,
       col.regions=redblue,
       cuts = 9,
       aspect=2.4, xlab = "Easting (km)", ylab = "Northing (km)",
       main="cleveland.soil - 2D smooth of Resistivity")

# Not run:
## 3D loess plot with data overlaid
libs(rgl)
bg3d(color = "white")
clear3d()
points3d(dat$easting, dat$northing, dat$resistivity / 100,
         col = rep("gray50", nrow(dat)))
rgl::surface3d(seq(.15, 1.410, by = .02),
             seq(.150, 3.645, by = .02),
             fit1/100, alpha=0.9, col=rep("wheat", length(fit1)),
             front="fill", back="fill")

rgl.close()

## End(Not run)
Yield and number of plants in a sugarbeet fertilizer experiment

Description

Yield and number of plants in a sugarbeet fertilizer experiment.

Usage

data("cochran.beets")

Format

A data frame with 42 observations on the following 4 variables.

fert fertilizer treatment
block block
yield yield, tons/acres
plants number of plants per plot

Details

Yield (tons/acre) and number of beets per plot. Fertilizer treatments combine superphosphate (P), muriate of potash (K), and sodium nitrate (N).

Source


References


Examples

library(agridat)
data(cochran.beets)
dat = cochran.beets

# P has strong effect
libs(lattice)
xyplot(yield ~ plants|fert, dat, main="cochran.beets")
cochran.bib  Multi-environment trial of corn, balanced incomplete block design

Description

Balanced incomplete block design in corn

Format

A data frame with 52 observations on the following 3 variables.

- loc  location/block, 13 levels
- gen  genotype/line, 13 levels
- yield  yield, pounds/plot

Details

Incomplete block design. Each loc/block has 4 genotypes/lines. The blocks are planted at different locations.

Conducted in 1943 in North Carolina.

Source

North Carolina Agricultural Experiment Station, United States Department of Agriculture.

References


Examples

```r
library(agridat)

data(cochran.bib)
dat <- cochran.bib

# Show the incomplete-block structure
libs(lattice)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(yield~loc*gen, dat,
  col.regions=redblue,
  xlab="loc (block)", main="cochran.bib - incomplete blocks")

with(dat, table(gen,loc))
rowSums(as.matrix(with(dat, table(gen,loc))))
colSums(as.matrix(with(dat, table(gen,loc))))
```
m1 = aov(yield ~ gen + Error(loc), data=dat)
summary(m1)

libs(nlme)
m2 = lme(yield ~ -1 + gen, data=dat, random=-1|loc)

---

cochran.crd  Potatoes scab infection with sulfur treatments

Description

Potato scab infection with sulfur treatments

Format

A data frame with 32 observations on the following 5 variables.

inf  infection percent
trt  treatment factor
row  row
col  column

Details

The experiment was conducted to investigate the effect of sulfur on controlling scab disease in potatoes. There were seven treatments. Control, plus spring and fall application of 300, 600, 1200 pounds/acre of sulfur. The response variable was infection as a percent of the surface area covered with scab. A completely randomized design was used with 8 replications of the control and 4 replications of the other treatments.

Although the original analysis did not show significant differences in the sulfur treatments, including a polynomial trend in the model uncovered significant differences (Tamura, 1988).

Source


References

https://doi.org/10.2134/agronj1988.00021962008000050003x
Examples

```r
## Not run:

library(agridat)
data(cochran.crd)
dat <- cochran.crd

# Field plan
lib(desplot)
desplot(dat, inf~col*row,
       text=trt, cex=1, # aspect unknown
       main="cochran.crd")

# CRD anova. Table 6 of Tamura 1988
contrasts(dat$trt) <- cbind(c1=c(1,1,1,-6,1,1,1), # Control vs Sulf
c2=c(-1,-1,-1,0,1,1,1)) # Fall vs Sp
m1 <- aov(inf ~ trt, data=dat)
anova(m1)
summary(m1, split=list(trt=list("Control vs Sulf"=1, "Fall vs Spring"=2)))

# Quadratic polynomial for columns...slightly different than Tamura 1988
m2 <- aov(inf ~ trt + poly(col,2), data=dat)
anova(m2)
summary(m2, split=list(trt=list("Control vs Sulf"=1, "Fall vs Spring"=2)))

## End(Not run)
```

cochran.eelworms

Counts of eelworms before and after fumigant treatments

Description

Counts of eelworms before and after fumigant treatments

Format

A data frame with 48 observations on the following 7 variables.

- **block**: block factor, 4 levels
- **row**: row
- **col**: column
- **fumigant**: fumigant factor
- **dose**: dose, Numeric 0,1,2. Maybe should be a factor?
- **initial**: count of eelworms pre-treatment
- **final**: count of eelworms post-treatment
Details

In the original experiment plan (as shown in Bailey 2008), columns 9, 10, 11 are shifted up slightly. Treatment codes: Con = Control, Chl = Chlorodinitrobenzen, Cym = Cymag, Car = Carbon Disulphide jelly, See = Seekay.

Experiment was conducted in 1935 at Rothamsted Experiment Station. In early March 400 grams of soil were sampled and the number of eelworm cysts were counted. Fumigants were added to the soil, oats were sown and later harvested. In October, the plots were again sampled and the final count of cysts recorded.

Source

Cochran and Cox, 1950. Experimental Designs. Table 3.1.

References


Examples

```r
## Not run:

library(agridat)
data(cochran.eelworms)
dat <- cochran.eelworms

# Very strong spatial trends
libs(desplot)
desplot(dat, initial ~ col*row,
       flip=TRUE, # aspect unknown
       main="cochran.eelworms")

# final counts are strongly related to initial counts
libs(lattice)
xyplot(final~initial|factor(dose), data=dat, group=fumigant,
       main="cochran.eelworms - by dose (panel) & fumigant",
       xlab="Initial worm count",
       ylab="Final worm count", auto.key=list(columns=5))

# One approach...log transform, use 'initial' as covariate, create 9 treatments
dat <- transform(dat, trt=factor(paste0(fumigant, dose)))
m1 <- aov(log(final) ~ block + trt + log(initial), data=dat)
anova(m1)

## End(Not run)
```
**coughran.factorial**

**Factorial experiment in beans**

**Description**

Factorial experiment in beans

**Usage**

```r
data("coughran.factorial")
```

**Format**

A data frame with 32 observations on the following 4 variables.

- **rep**: rep factor
- **block**: block factor
- **trt**: treatment factor, 16 levels
- **yield**: yield (pounds)

**Details**

Conducted by Rothamsted Experiment Station in 1936. The treatments are listed below.

- **d**: dung: None, 10 tons/acre.
- **n**: nitrochalk: None, 0.4 hundredweight nitrogen per acre.
- **p**: Superphosphate: None, 0.6 hundredweight per acre
- **k**: muriate of potash: None, 1 hundredweight K20 per acres.

The response variable is the yield of beans.

**Source**


**Examples**

```r
## Not run:
library(agridat)
data(coughran.factorial)
dat <- coughran.factorial

# Split treatment into individual factors
dat <- transform(dat,
                d = -1 + 2 * grepl("d", trt),
                n = -1 + 2 * grepl("n", trt),
```
p = -1 + 2 * grepl('p',trt),
k = -1 + 2 * grepl('k',trt))
dat <- transform(dat, d=factor(d), n=factor(n), p=factor(p), k=factor(k))

# Cochran table 6.5.
m1 <- lm(yield ~ rep * block + (d+n+p+k)^3, data=dat)
anova(m1)

libs(FrF2)
aliases(m1)
MEPlot(m1, select=3:6,
  main="cochran.factorial - main effects plot")

## End(Not run)

---

**cochran.latin**

*Latin square design in wheat*

**Description**

Six wheat plots were sampled by six operators and shoot heights measured. The operators sampled plots in six ordered sequences. The dependent variate was the difference between measured height and true height of the plot.

**Format**

A data frame with 36 observations on the following 4 variables.

- **row** row
- **col** column
- **operator** operator factor
- **diff** difference between measured height and true height

**Source**


**Examples**

```r
library(agridat)
data(cochran.latin)
dat <- cochran.latin
libs(desplot)
desplot(dat, diff~col*row,
```
cochran.lattice

Balanced lattice experiment in cotton

Description

Balanced lattice experiment in cotton

Usage

data("cochran.lattice")

Format

A data frame with 80 observations on the following 5 variables.

y  percent of affected flower buds
rep replicate
row row
col column
trt treatment factor

Details

The experiment is a balanced lattice square with 16 treatments in a 4x4 layout in each of 5 replicates. The treatments were applied to cotton plants. Each plot was ten rows wide by 70 feet long (about 1/18 of an acre). (Estimated plot width is 34.5 feet.) Data were collected from the middle 4 rows. The data are the percentages of squares showing attack by boll weevils. A 'square' is the name given to a young flower bud.

The plot orientation is not clear.

Source


References

Examples

```r
## Not run:

library(agridat)
data(cochran.lattice)
dat <- cochran.lattice

libs(desplot)
desplot(dat, y~row*col|rep,
    text=trt, # aspect unknown, should be 2 or .5
    main="cochran.lattice")

# Random rep,row,column model often used by Federer
libs(lme4)
dat <- transform(dat, rowf=factor(row), colf=factor(col))
m1 <- lmer(y ~ trt + (1|rep) + (1|rep:row) + (1|rep:col), data=dat)
summary(m1)

## End(Not run)
```

cochran.wireworms

Wireworms controlled by fumigants in a latin square

Description
Wireworms controlled by fumigants in a latin square

Format
A data frame with 25 observations on the following 4 variables.

- **row**: row
- **col**: column
- **trt**: fumigant treatment, 5 levels
- **worms**: count of wireworms per plot

Details
Plots were approximately 22 cm by 13 cm. Layout of the experiment was a latin square. The number of wireworms in each plot was counted, following soil fumigation the previous year.
Source


References


Examples

```r
## Not run:

library(agridat)
data(cochran.wireworms)
dat <- cochran.wireworms

libs(desplot)
desplot(dat, worms ~ col*row,
  text=trt, cex=1, # aspect unknown
  main="cochran.wireworms")

# Trt K is effective, but not the others. Really, this says it all.
libs(lattice)
bwplot(worms ~ trt, dat, main="cochran.wireworms", xlab="Treatment")

# Snedecor and Cochran do ANOVA on sqrt(x+1).
dat <- transform(dat, rowf=factor(row), colf=factor(col))
m1 <- aov(sqrt(worms+1) ~ rowf + colf + trt, data=dat)
anova(m1)

# Instead of transforming, use glm
m2 <- glm(worms ~ trt + rowf + colf, data=dat, family="poisson")
anova(m2)

# GLM with random blocking.
libs(lme4)
m3 <- glmer(worms ~ -1 +trt +(1|rowf) +(1|colf), data=dat, family="poisson")
summary(m3)
```

## Fixed effects:

| Estimate | Std. Error | z value | Pr(>|z|) |
|----------|------------|---------|----------|
| trtK     | 0.1393     | 0.4275  | 0.326    | 0.745    |
| trtM     | 1.7814     | 0.2226  | 8.002    | 1.22e-15***|
| trtN     | 1.9028     | 0.2142  | 8.881    | < 2e-16 ***|
| trtO     | 1.7147     | 0.2275  | 7.537    | 4.80e-14 ***|
connolly.potato

## End(Not run)

---

connolly.potato  

Potato yields in single-drill plots

### Description

Potato yields in single-drill plots

### Usage

```r
data("connolly.potato")
```

### Format

A data frame with 80 observations on the following 6 variables.

- **rep**  
- **block**  
- **gen**  
- **variety**  
- **row**  
- **col**  
- **yield** yield, kg/ha  
- **matur**  maturity group

### Details

Connolly et el use this data to illustrate how yield can be affected by competition from neighboring plots.

This data uses M1, M2, M3 for maturity, while Connolly et al use FE (first early), SE (second early) and M (maincrop).

The trial was 20 sections, each of which was an independent row of 20 drills. The data here are four reps of single-drill plots from sections 1, 6, 11, and 16.

The neighbor covariate for a plot is defined as the average of the plots to the left and right. For drills at the edge of the trial, the covariate was the average of the one neighboring plot yield and the section (i.e. rep) mean.

It would be interesting to fit a model that uses differences in maturity between a plot and its neighbor as the actual covariate.


Used with permission of Iain Currie.

### Source

Examples

## Not run:

library(agridat)

data(connolly.potato)
dat <- connolly.potato

# Field plan
libs(desplot)
desplot(dat, yield~col*row,
outl=rep, # aspect unknown
main="connolly.potato yields (reps not contiguous)"

# Later maturities are higher yielding
libs(lattice)
bwplot(yield~matur, dat, main="connolly.potato yield by maturity")

# Observed raw means. Matches Connolly table 2.
mm <- aggregate(yield~gen, data=dat, FUN=mean)
mn[rev(order(mn$yield)),]

# Create a covariate which is the average of neighboring plot yields
libs(reshape2)
mat <- acast(dat, row~col, value.var='yield')
mat2 <- matrix(NA, nrow=4, ncol=20)
mat2[, 1] <- (mat[, 1] + apply(mat, 1, mean))/2
mat2[, 20] <- (mat[, 20] + apply(mat, 1, mean))/2
dat2 <- melt(mat2)
colnames(dat2) <- c('row','col','cov')
dat <- merge(dat, dat2)
# xyplot(yield ~ cov, data=dat, type=c('p','r'))

# Connolly et al fit a model with avg neighbor yield as a covariate
m1 <- lm(yield ~ 0 + gen + rep + cov, data=dat)
coef(m1)['cov'] # = -0.303 (Connolly obtained -0.31)

# Block names and effects
bmn <- c("R1","R2","R3","R4")
beff <- c(0, coef(m1)[c('repR2','repR3','repR4')])

# Variety names and effects
vnm <- paste0("V", formatC(1:20, width=2, flag='0'))
veff <- coef(m1)[1:20]

# Adjust yield for variety and block effects
dat <- transform(dat, yadj = yield - beff[match(rep,bmn)]
- veff[match(gen,vnm)])

# Similar to Connolly Fig 1. Point pattern doesn’t quite match
xyplot(yadj~cov, data=dat, type=c('p','r'),
cornelius.maize

```r
main="connolly.potato",
   xlab="Avg yield of nearest neighbors",
   ylab="Yield, adjusted for variety and block effects"
)

## End(Not run)
```

cornelius.maize  
Multi-environment trial of maize for 9 cultivars at 20 locations.

Description

Maize yields for 9 cultivars at 20 locations.

Usage

data("cornelius.maize")

Format

A data frame with 180 observations on the following 3 variables.

- env  environment factor, 20 levels
- gen  genotype/cultivar, 9 levels
- yield  yield, kg/ha

Details

Cell means (kg/hectare) for the CIMMYT EVT16B maize yield trial.

Source


References

Examples

```r
## Not run:
library(agridat)
data(cornelius.maize)
dat <- cornelius.maize

dotplot(gen~yield|env,dat) # We cannot compare genotype yields easily
# Subtract environment mean from each observation
libs(reshape2)
mat <- acast(dat, gen~env)
mat <- scale(mat, scale=FALSE)
dat2 <- melt(mat)
names(dat2) <- c('gen', 'env', 'yield')
libs(lattice)
bwplot(yield ~ gen, dat2,
       main="cornelius.maize - environment centered yields")

if(0){
  # This reproduces the analysis of Forkman and Piepho.

test.pc <- function(Y0, type="AMMI", n.boot=10000, maxpc=6) {
  # Test the significance of Principal Components in GGE/AMMI

  # Singular value decomposition of centered/double-centered Y
  Y <- sweep(Y0, 1, rowMeans(Y0)) # subtract environment means
  if(type="AMMI") {
    Y <- sweep(Y, 2, colMeans(Y0)) # subtract genotype means
    Y <- Y + mean(Y0)
  }
  lam <- svd(Y)$d

  # Observed value of test statistic.
  # t.obs[k] is the proportion of variance explained by the kth term out of
  # the k...M terms, e.g. t.obs[2] is lam[2]^2 / sum(lam[2:M]^2)
  t.obs <- { lam^2/rev(cumsum(rev(lam^2))) }[1:(M-1)]
  t.boot <- matrix(NA, nrow=n.boot, ncol=M-1)
  I <- if(type="AMMI") nrow(Y0)-1 else nrow(Y0)
  J <- ncol(Y0)-1
  M <- min(I, J) # rank of Y, maximum number of components
  M <- min(M, maxpc) # Optional step: No more than 5 components
  for(K in 0:(M-2)){ # 'K' multiplicative components in the svd
    for(bb in 1:n.boot){
      E.b <- matrix(rnorm((I-K) * (J-K)), nrow = I-K, ncol = J-K)
      lam.b <- svd(E.b)$d
      t.boot[bb, K+1] <- lam.b[1]^2 / sum(lam.b^2)
    }
  }
}
```
corsten.interaction

Multi-environment trial of corn

Description

The data is the yield (kg/acre) of 20 genotypes of corn at 7 locations.

Format

A data frame with 140 observations on the following 3 variables.

gen  genotype, 20 levels
loc  location, 7 levels
yield  yield, kg/acre

Details

The data is used by Corsten & Denis (1990) to illustrate two-way clustering by minimizing the interaction sum of squares.

In their paper, the labels on the location dendrogram have a slight typo. The order of the loc labels shown is 1 2 3 4 5 6 7. The correct order of the loc labels is 1 2 4 5 6 7 3.

Used with permission of Jean-Baptiste Denis.
Source


Examples

```r
## Not run:
library(agridat)
data(corsten.interaction)
dat <- corsten.interaction

libs(reshape2)
m1 <- melt(dat, measure.var='yield')
dmat <- acast(m1, loc~gen)

# Corsten (1990) uses this data to illustrate simultaneous row and column clustering based on interaction sums-of-squares.
# There is no (known) function in R to reproduce this analysis (please contact the package maintainer if this is not true).
# For comparison, the 'heatmap' function clusters the rows and columns _independently_ of each other.
heatmap(dmat, main="corsten.interaction")

## End(Not run)
```

---

cox.stripsplit

`cox.stripsplit`  
*Strip-split-plot of barley with fertilizer, calcium, and soil factors.*

Description

Strip-split-plot of barley with fertilizer, calcium, and soil factors.

Format

A data frame with 96 observations on the following 5 variables.

- `rep` replicate, 4 levels
- `soil` soil, 3 levels
- `fert` fertilizer, 4 levels
- `calcium` calcium, 2 levels
- `yield` yield of winter barley
Details

Four different fertilizer treatments are laid out in vertical strips, which are then split into subplots with different levels of calcium. Soil type is stripped across the split-plot experiment, and the entire experiment is then replicated three times.

Sometimes called a split-block design.

Source

Comes from the notes of Gertrude Cox and A. Rotti.

References


Examples

```r
## Not run:
library(agridat)
data(cox.stripsplit)
dat <- cox.stripsplit

# Raw means
# aggregate(yield ~ calcium, data=dat, mean)
# aggregate(yield ~ soil, data=dat, mean)
# aggregate(yield ~ calcium, data=dat, mean)

libs(HH)
interaction2wt(yield ~ rep + soil + fert + calcium, dat,
x.between=0, y.between=0,
main="cox.stripsplit")

# Traditional AOV
m1 <- aov(yield~ fert*calcium*soil +
 Error(rep/(fert+soil+calcium:fert+soil:fert)),
data=dat)
summary(m1)

with balanced data, the following are all basically identical

libs(lme4)
# The 'rep:soil:fert' term causes problems...so we drop it.
m2 <- lmer(yield ~ fert*soil*calcium + (1|rep) + (1|rep:fert) +
(1|rep:soil) + (1|rep:fert:calcium), data=dat)

if(0){
# afex uses Kenward-Rogers approach for denominator d.f.
libs(afex)
mixed(yield ~ fert*soil*calcium + (1|rep) + (1|rep:fert) +
(1|rep:soil) + (1|rep:fert:calcium) + (1|rep:soil:fert), data=dat,
...}
```
control=lmerControl(check.nobs.vs.rankZ="ignore"))

## Effect     stat  ndf  ddf F.scaling  p.value
## 1 (Intercept) 1350.8113 1 3.0009 1 0.0000
## 2 fert        3.5619  3 9.0000 1 0.0604
## 3 soil        3.4659  2 6.0000 1 0.0999
## 4 calcium     1.8835  1 12.0000 1 0.1950
## 5 fert:soil   1.2735  6 18.0000 1 0.3179
## 6 fert:calcium 4.4457  3 12.0000 1 0.0255
## 7 soil:calcium 0.2494  2 24.0000 1 0.7813
## 8 fert:soil:calcium 0.3504  6 24.0000 1 0.9027

} # End(Not run)

cramer.cucumber

Cucumber yields and quantitative traits

Description

Cucumber yields and quantitative traits

Usage

data("cramer.cucumber")

Format

A data frame with 24 observations on the following 9 variables.

- cycle
- rep
- plants
- flowers
- branches
- leaves
- totalfruit
- culledfruit
- earlyfruit

Details

The data are used to illustrate path analysis of the correlations between phenotypic traits.

Used with permission of Christopher Cramer.
Source


References


Examples

```r
## Not run:
library(agridat)
data(cramer.cucumber)
dat <- cramer.cucumber

libs(lattice)
splom(dat[3:9], group=dat$cycle,
    main="cramer.cucumber - traits by cycle",
    auto.key=list(columns=3))

# derived traits
dat <- transform(dat,
    marketable = totalfruit-culledfruit,
    branchesperplant = branches/plants,
    nodesperbranch = leaves/(branches+plants),
    femalenodes = flowers+totalfruit)
dat <- transform(dat,
    perfenod = (femalenodes/leaves),
    fruitset = totalfruit/flowers,
    fruitperplant = totalfruit / plants,
    marketableperplant = marketable/plants,
    earlyperplant=earlyfruit/plants)

# just use cycle 1
dat1 <- subset(dat, cycle==1)

# define independent and dependent variables
indep <- c("branchesperplant", "nodesperbranch", "perfenod", "fruitset")
dep0 <- "fruitperplant"
dep <- c("marketable","earlyperplant")

# standardize trait data for cycle 1
sdat <- data.frame(scale(dat1[1:8, c(indep,dep0,dep)]))

# slopes for dep0 ~ indep
X <- as.matrix(sdat[,indep])
Y <- as.matrix(sdat[,c(dep0)])
# estdep <- solve(t(X))
estdep <- solve(crossprod(X), crossprod(X,Y))
```
estdep
## branchesperplant 0.7160269
## nodesperbranch 0.3415537
## perfenod 0.2316693
## fruitset 0.2985557

# slopes for dep ~ dep0
X <- as.matrix(sdat[,dep0])
Y <- as.matrix(sdat[,c(dep)])
estind2 <- solve(t(X), crossprod(X,Y))
estind2

## marketable earlyperplant
## 0.97196 0.8828393

# correlation coefficients for indep variables
corrind=cor(sdat[,indep])
round(corrind,2)

## branchesperplant nodesperbranch perfenod fruitset
## branchesperplant 1.00 0.52 -0.24 0.09
## nodesperbranch 0.52 1.00 -0.44 0.14
## perfenod -0.24 -0.44 1.00 0.04
## fruitset 0.09 0.14 0.04 1.00

# Correlation coefficients for dependent variables
corrdep=cor(sdat[,c(dep0, dep)])
round(corrdep,2)

## fruitperplant marketable earlyperplant
## fruitperplant 1.00 0.97 0.88
## marketable 0.97 1.00 0.96
## earlyperplant 0.88 0.96 1.00

result = corrind
result = result*matrix(estdep,ncol=4,nrow=4,byrow=TRUE)
round(result,2) # match SAS output columns 1-4

## branchesperplant nodesperbranch perfenod fruitset
## branchesperplant 0.72 0.18 -0.06 0.03
## nodesperbranch 0.37 0.34 -0.10 0.04
## perfenod -0.17 -0.15 0.23 0.01
## fruitset 0.07 0.05 0.01 0.30

resdep0 = rowSums(result)
resdep <- cbind(resdep0,resdep0)*matrix(estind2, nrow=4,ncol=2,byrow=TRUE)
colnames(resdep) <- dep

# slightly different from SAS output last 2 columns
round(cbind(fruitperplant=resdep0, round(resdep,2)),2)

## fruitperplant marketable earlyperplant
## branchesperplant 0.87 0.84 0.76
## nodesperbranch 0.65 0.63 0.58
## perfenod -0.08 -0.08 -0.07
## fruitset 0.42 0.41 0.37

## End(Not run)
Weight gain in pigs for different treatments, with initial weight and feed eaten as covariates.

```r
data("crampton.pig")
```

A data frame with 50 observations on the following 5 variables.

- `treatment`: feed treatment
- `rep`: replicate
- `weight1`: initial weight
- `feed`: feed eaten
- `weight2`: final weight

A study of the effect of initial weight and feed eaten on the weight gaining ability of pigs with different feed treatments.

The data are extracted from Ostle. It is not clear that 'replicate' is actually a blocking replicate as opposed to a repeated measurement. The original source document needs to be consulted.

Crampton, EW and Hopkins, JW. (1934). The Use of the Method of Partial Regression in the Analysis of Comparative Feeding Trial Data, Part II. The Journal of Nutrition, 8, 113-123. https://doi.org/10.1093/jn/8.3.329


```r
## Not run:
library(agridat)
data(crampton.pig)
dat <- crampton.pig
```
dat <- transform(dat, gain=weight2-weight1)
libs(lattice)
# Trt 4 looks best
xyplot(gain ~ feed, dat, group=treatment, type=c('p','r'),
       auto.key=list(columns=5),
       xlab="Feed eaten", ylab="Weight gain", main="crampton.pig")

# Basic Anova without covariates
m1 <- lm(weight2 ~ treatment + rep, data=dat)
anova(m1)
# Add covariates
m2 <- lm(weight2 ~ treatment + rep + weight1 + feed, data=dat)
anova(m2)
# Remove treatment, test this nested model for significant treatments
m3 <- lm(weight2 ~ rep + weight1 + feed, data=dat)
anova(m2,m3) # p-value .07. F=2.34 matches Ostle

## End(Not run)

crossa.wheat  Multi-environment trial of wheat for 18 genotypes at 25 locations

Description
Wheat yields for 18 genotypes at 25 locations

Format
A data frame with 450 observations on the following 3 variables.

  loc  location
  locgroup  location group: Grp1-Grp2
  gen  genotype
  gengroup  genotype group: W1, W2, W3
  yield  grain yield, tons/ha

Details
Grain yield from the 8th Elite Selection Wheat Yield Trial to evaluate 18 bread wheat genotypes at 25 locations in 15 countries.
Cross et al. used this data to cluster locations into 2 mega-environments and clustered genotypes into 3 wheat clusters.

Locations

<table>
<thead>
<tr>
<th>Code</th>
<th>Country</th>
<th>Location</th>
<th>Latitude (N)</th>
<th>Elevation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK</td>
<td>Algeria</td>
<td>El Khroub</td>
<td>36</td>
<td>640</td>
</tr>
<tr>
<td>AL</td>
<td>Algeria</td>
<td>Setif</td>
<td>36</td>
<td>1,023</td>
</tr>
</tbody>
</table>
crossa.wheat

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ</td>
<td>Bangladesh</td>
<td>Joydebpur</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>CA</td>
<td>Cyprus</td>
<td>Athalassa</td>
<td>35</td>
<td>142</td>
</tr>
<tr>
<td>EG</td>
<td>Egypt</td>
<td>El Gemmeiza</td>
<td>31</td>
<td>8</td>
</tr>
<tr>
<td>ES</td>
<td>Egypt</td>
<td>Sakha</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>EB</td>
<td>Egypt</td>
<td>Beni-Suef</td>
<td>29</td>
<td>28</td>
</tr>
<tr>
<td>IL</td>
<td>India</td>
<td>Ludhiana</td>
<td>31</td>
<td>247</td>
</tr>
<tr>
<td>ID</td>
<td>India</td>
<td>Delhi</td>
<td>29</td>
<td>228</td>
</tr>
<tr>
<td>JM</td>
<td>Jordan</td>
<td>Madaba</td>
<td>36</td>
<td>785</td>
</tr>
<tr>
<td>KN</td>
<td>Kenya</td>
<td>Njoro</td>
<td>0</td>
<td>2,165</td>
</tr>
<tr>
<td>MG</td>
<td>Mexico</td>
<td>Guanajuato</td>
<td>21</td>
<td>1,765</td>
</tr>
<tr>
<td>MS</td>
<td>Mexico</td>
<td>Sonora</td>
<td>27</td>
<td>38</td>
</tr>
<tr>
<td>MM</td>
<td>Mexico</td>
<td>Michoacfin</td>
<td>20</td>
<td>1,517</td>
</tr>
<tr>
<td>NB</td>
<td>Nepal</td>
<td>Bhairahwa</td>
<td>27</td>
<td>105</td>
</tr>
<tr>
<td>PI</td>
<td>Pakistan</td>
<td>Islamabad</td>
<td>34</td>
<td>683</td>
</tr>
<tr>
<td>PA</td>
<td>Pakistan</td>
<td>Ayub</td>
<td>32</td>
<td>213</td>
</tr>
<tr>
<td>SR</td>
<td>Saudi Arabia</td>
<td>Riyadh</td>
<td>24</td>
<td>600</td>
</tr>
<tr>
<td>SG</td>
<td>Sudan</td>
<td>Gezira</td>
<td>14</td>
<td>411</td>
</tr>
<tr>
<td>SE</td>
<td>Spain</td>
<td>Encinar</td>
<td>38</td>
<td>20</td>
</tr>
<tr>
<td>SJ</td>
<td>Spain</td>
<td>Jerez</td>
<td>37</td>
<td>180</td>
</tr>
<tr>
<td>SC</td>
<td>Spain</td>
<td>Cordoba</td>
<td>38</td>
<td>110</td>
</tr>
<tr>
<td>SS</td>
<td>Spain</td>
<td>Sevilla</td>
<td>38</td>
<td>20</td>
</tr>
<tr>
<td>TB</td>
<td>Tunisia</td>
<td>Beja</td>
<td>37</td>
<td>150</td>
</tr>
<tr>
<td>TC</td>
<td>Thailand</td>
<td>Chiang Mai</td>
<td>18</td>
<td>820</td>
</tr>
</tbody>
</table>

Used with permission of Jose’ Crossa.

Source


References


Examples

```r
## Not run:
library(agridat)
data(crossa.wheat)
dat <- crossa.wheat

# AMMI biplot. Fig 3 of Crossa et al.
libs(agricolae)
m1 <- with(dat, AMMI(E=loc, G=gen, R=1, Y=yield))
```
crowder.seeds

Germination of Orobanche seeds for two genotypes and two treatments.

Description

Number of Orobanche seeds tested/germinated for two genotypes and two treatments.

Format

- plate Factor for replication
- gen Factor for genotype with levels 073, 075
- extract Factor for extract from bean, cucumber
- germ Number of seeds that germinated
- n Total number of seeds tested
Details

Egyptian broomrape, * Orobanche aegyptiaca* is a parasitic plant family. The plants have no chlorophyll and grow on the roots of other plants. The seeds remain dormant in soil until certain compounds from living plants stimulate germination.

Two genotypes were studied in the experiment, *O. aegyptiaca 73* and *O. aegyptiaca 75*. The seeds were brushed with one of two extracts prepared from either a bean plant or cucumber plant.

The experimental design was a 2x2 factorial, each with 5 or 6 reps of plates.

Source


References


Examples

```r
## Not run:

library(agridat)
data(crowder.seeds)
dat <- crowder.seeds
m1.glm <- m1.glmm <- m1.bb <- m1.hglm <- NA

# ----- Graphic
libs(lattice)
dotplot(germ/n~gen|extract, dat, main="crowder.seeds")

# ----- GLM.
# family=binomial() fixes dispersion at 1
# family=quasibinomial() estimates dispersion, had larger std errors
m1.glm <- glm(cbind(germ,n-germ) ~ gen*extract,
              data=dat,
              family=binomial(),
              family=quasibinomial()
)
summary(m1.glm)

# --- GLMM. Assumes Gaussian random effects
libs(MASS)
m1.glmm <- glmmPQL(cbind(germ, n-germ) ~ gen*extract, random= ~1|plate,
                   family=binomial(), data=dat)
summary(m1.glmm)
```
libs(brms)
m1.brms <- brm(

  # ----- HGML package. Beta-binomial with beta-distributed random effects
  # libs(hglm)
  # m1.hglm <- hglm(fixed= germ/n - I(gen=="O75")*extract, weights=n, data=dat,
  # random=~1|plate, family=binomial(), rand.family=Beta(),
  # fix.disp=1)

  # ----- INLA package. See: https://haakonbakka.bitbucket.io/btopic102.html
  # libs(INLA)
  # gen,extract are fixed. plate is a random effect
  # Priors for hyper parameters. See: inla.doc("pc.prec")
  # hyper1 = list(theta = list(prior="pc.prec", param=c(1,0.01)))
  # m1.inla = inla(germ ~ gen*extract + f(plate, model="iid", hyper=hyper1),
  #   data=crowder.seeds,
  #   family="binomial", Ntrials=n,
  #   control.family=list(control.link=list(model="logit")))

  # Compare coefficients

  ## round(summary(m1.glm)$coef,2)
  ## Estimate Std. Error t value Pr(>|t|)
  ## (Intercept) -0.41 0.25 -1.64 0.12
  ## genO75 -0.15 0.30 -0.48 0.64
  ## extractcucumber 0.54 0.34 1.58 0.13
  ## genO75:extractcucumber 0.78 0.42 1.86 0.08

  ## round(summary(m1.glmm)$tTable,2)
  ## Value Std.Error DF t-value p-value
  ## (Intercept) -0.44 0.25 17 -1.80 0.09
  ## genO75 -0.10 0.31 17 -0.34 0.74
  ## extractcucumber 0.52 0.34 17 1.56 0.14
  ## genO75:extractcucumber 0.80 0.42 17 1.88 0.08

  ## round(summary(m1.bb)$BCoef,2)
  ## Estimate Std. Error z value Pr(>|z|)
  ## (Intercept) -0.44 0.22 -2.04 0.04
  ## genO75 -0.10 0.27 -0.36 0.72
  ## extractcucumber 0.52 0.30 1.76 0.08
  ## genO75:extractcucumber 0.80 0.38 2.11 0.03

  ## round(summary(m1.hglm)$FixCoefMat,2)
  ## Estimate Std. Error t-value Pr(>|t|)
  ## (Intercept) -0.47 0.24 -1.92 0.08
  ## I(gen == "O75")TRUE -0.08 0.31 -0.25 0.81
  ## extractcucumber 0.51 0.33 1.53 0.16
  ## I(gen == "O75")TRUE:extractcucumber 0.83 0.43 1.92 0.08

  ## round(summary(m1.inla$summary.fixed,2)
  ## Estimate Std. Error t-value Pr(>|t|)
  ## (Intercept) -0.46 0.24 -1.92 0.08
  ## I(gen == "O75")TRUE -0.08 0.31 -0.25 0.81
  ## extractcucumber 0.51 0.33 1.53 0.16
  ## I(gen == "O75")TRUE:extractcucumber 0.83 0.43 1.92 0.08
## genO75 -0.09 0.30 -0.67 -0.10 0.52 -0.11 0
## extractcucumber 0.53 0.32 -0.12 0.53 1.17 0.53 0
## genO75:extractcucumber 0.82 0.42 0.01 0.82 1.66 0.81 0

# ----- Stan using pre-built models from rstanarm
libs(tidyverse, rstan, rstanarm,bayesplot)
m1.stan <- stan_glm( cbind(germ,n-germ) ~ gen*extract, 
data=dat, 
  family = binomial(link="logit") )
# summary(m1.stan)
round(posterior_interval(m1.stan, prob=.90),3)
# 5
# (Intercept) -0.715 -0.111
# genO75 -0.512 0.228
# extractcucumber 0.123 0.977
# genO75:extractcucumber 0.248 1.284

libs(bayesplot)
mcmc_areas(m1.stan, prob = 0.9) +
ggtitle("Posterior distributions", "with medians and 95

if(0) {

  # --- rjags version ---

  # JAGS/BUGS. See https://mathstat.helsinki.fi/openbugs/Examples/Seeds.html
  # Germination rate depends on p, which is a logit of a linear predictor
  # based on genotype and extract, plus random deviation to intercept

  # To match the output on the BUGS web page, use: dat$gen=="O73".
  # We use dat$gen=="O75" to compare with the parameterization above.
  jdat =list(germ = dat$germ, n = dat$n,
    root = as.numeric(dat$extract=="cucumber"),
    gen = as.numeric(dat$gen=="O75"),
    nobs = nrow(dat))

  jinit = list(int = 0, genO75 = 0, extcuke = 0, g75ecuke = 0, tau = 10)

  # Use logical names (unlike BUGS documentation)
  mod.bug =
"model {
  for(i in 1:nobs) {
    germ[i] ~ dbin(p[i], n[i])
    b[i] ~ dnorm(0.0, tau)
    logit(p[i]) <- int + genO75 * gen[i] + extcuke * root[i] +
    g75ecuke * gen[i] * root[i] + b[i]
  }
  int ~ dnorm(0.0, 1.0E-6)
  genO75 ~ dnorm(0.0, 1.0E-6)
  extcuke ~ dnorm(0.0, 1.0E-6)
  g75ecuke ~ dnorm(0.0, 1.0E-6)
  tau ~ dgamma(0.001, 0.001)
  sigma <- 1 / sqrt(tau)"
```r
library(rjags)

oo <- textConnection(mod.bug)
jj <- jags.model(oo, data=jdat, inits=jinit, n.chains=1)
close(oo)

c1 <- coda.samples(jj, c("int","genO75","g75ecuke","extcuke","sigma"),
                  n.iter=20000)

summary(c1) # Medians are very similar to estimates from hglm

# Plot observed data with HPD intervals for germination probability

obs <- dotplot(1:21 ~ germ/n, dat,
                main="crowder.seeds", ylab="plate",
                col=as.numeric(dat$gen), pch=substring(dat$extract,1))

obs + segplot(1:21 ~ lower + upper, data=fit, centers=med)

# --- R2jags version ---

library("agridat")
library("R2jags")

dat <- crowder.seeds

# To match the output on the BUGS web page, use: dat$gen=="O73".
# We use dat$gen=="O75" to compare with the parameterization above.

jdat =list(germ = dat$germ, n = dat$n,
            root = as.numeric(dat$extract=="cucumber"),
            gen = as.numeric(dat$gen=="O75"),
            nobs = nrow(dat))

jinit = list(list(int = 0, genO75 = 0, extcuke = 0, g75ecuke = 0, tau = 10))

mod.bug = function()
{
  for(i in 1:nobs)
  {
    germ[i] ~ dbin(p[i], n[i])
    b[i] ~ dnorm(0.0, tau)
    logit(p[i]) <- int + genO75 * gen[i] + extcuke * root[i] +
                   g75ecuke * gen[i] * root[i] + b[i]
  }

  return()
}
```
cullis.earlygen

```r
int ~ dnorm(0.0, 1.0E-6)
genO75 ~ dnorm(0.0, 1.0E-6)
extcuke ~ dnorm(0.0, 1.0E-6)
g75ecuke ~ dnorm(0.0, 1.0E-6)
tau ~ dgamma(0.001, 0.001)
sigma <- 1 / sqrt(tau)
```

```r
parms <- c("int","genO75","g75ecuke","extcuke","sigma")
j1 <- jags(data=jdat, inits=jinit, parms, model.file=mod.bug,
n.iter=20000, n.chains=1)
print(j1)
## mu.vect sd.vect 2.5
## extcuke  0.519  0.325 -0.140 0.325 0.531 0.728 1.158
## g75ecuke 0.834  0.429 -0.019 0.552 0.821 1.101 1.710
## genO75  -0.096  0.305 -0.670 -0.295 -0.115 0.089 0.552
## int     -0.461  0.236 -0.965 -0.603 -0.455 -0.312 0.016
## sigma   0.255  0.148  0.033 0.140 0.240 0.352 0.572
## deviance 103.319  7.489  90.019 98.010 102.770 108.689 117.288
```

densityplot(as.mcmc(j1))
HPDinterval(as.mcmc(j1))

```r
## End(Not run)
```

---

cullis.earlygen Early generation variety trial in wheat

**Description**

Early generation variety trial in wheat

**Format**

A data frame with 670 observations on the following 5 variables.

- `gen` genotype factor
- `row` row
- `col` column
- `entry` entry (genotype) number
- `yield` yield of each plot, kg/ha
- `weed` weed score
Details

The data are from an unreplicated field experiment conducted at Tullibigeal, New South Wales, Australia in 1987-88. In each row, every 6th plot was the variety 'Kite'. Six other standard varieties were randomly interspersed over the trial. Each plot was 15m x 1.8m, "oriented with the longest side with rows".

The 'weed' variable is a visual score on a 0 to 10 scale, 0 = no weeds, 10 = 100 percent weeds.

The replicated check variety was numbered 526. A further 6 replicated commercially available varieties (numbered 527 to 532) were also randomly assigned to plots with between 3 to 5 plots of each. The aim of these trials is to identify and retain the top, say 20 percent of lines for further testing. Cullis et al. (1989) presented an analysis of early generation variety trials that included a one-dimensional spatial analysis. Below, a two-dimensional spatial analysis is presented.

Note: The 'row' and 'col' variables are as in the VSN link below (switched compared to the paper by Cullis et al.)

Field width: 10 rows * 15 m = 150 m
Field length: 67 plots * 1.8 m = 121 m

The orientation is not certain, but the alternative orientation would have a field roughly 20m x 1000m, which seems unlikely.

Source


References


Examples

```r
## Not run:
library(agridat)
data(cullis.earlygen)
dat <- cullis.earlygen

# Show field layout of checks. Cullis Table 1.
dat$check <- ifelse(dat$entry < 8, dat$entry, NA)
libss(desplot)
desplot(dat, yield ~ col*row,
  col="check", cex=0.5, flip=TRUE, aspect=121/150, # true aspect
  main="cullis.earlygen (yield)"
)

grays <- colorRampPalette(c(white, "#252525"))
desplot(dat, weed ~ col*row,
  at=0:6-0.5, col.regions=grays(7)[-1],
  flip=TRUE, aspect=121/150, # true aspect
  main="cullis.earlygen (weed)"
)`
libs(lattice)
bwplot(yield ~ as.character(weed), dat,  
  horizontal=FALSE,  
  xlab="Weed score", main="cullis.earlygen")

# Moving Grid
libs(mvngGrAd)
shape <- list(c(1),
  c(1),
  c(1:4),
  c(1:4))
# sketchGrid(10,10,20,20,shapeCross=shape, layers=1, excludeCenter=TRUE)
m0 <- movingGrid(rows=dat$row, columns=dat$col, obs=dat$yield,  
  shapeCross=shape, layers=NULL)
data$mov.avg <- fitted(m0)

libs(asreml) # asreml4

# Start with the standard AR1xAR1 analysis
dat <- transform(dat, xf=factor(col), yf=factor(row))
dat <- dat[order(dat$xf, dat$yf),]
m2 <- asreml(yield ~ weed, data=dat, random= ~gen,  
  resid = ~ ar1(xf):ar1(yf))

# Variogram suggests a polynomial trend
m3 <- update(m2, fixed= yield~weed+pol(col,-1))

# Now add a nugget variance
m4 <- update(m3, random= ~ gen + units)

libs(lucid)
vc(m4)
## effect component std.error z.ratio bound
## gen 73780 10420 7.1 P 0
## units 30440 8073 3.8 P 0.1
## xf:yd(R) 54730 10630 5.1 P 0
## xf:yd!xf!cor 0.38 0.115 3.3 U 0
## xf:yd!yf!cor 0.84 0.045 19 U 0

## # Predictions from models m3 and m4 are non-estimable. Why?
## # Use model m2 for predictions
## predict(m2, classify="gen")$pvals
## #
## ## gen predicted.value std.error status
## ## 1 Banks 2723.534 93.14719 Estimable
## ## 2 Eno008 2981.056 162.85241 Estimable
## ## 3 Eno009 2978.088 161.57129 Estimable
## ## 4 Eno010 2821.399 153.96943 Estimable
## ## 5 Eno011 2991.612 161.53507 Estimable

## # Compare AR1 with Moving Grid
## dat$ar1 <- fitted(m2)
damesa.maize

Incomplete-block experiment of maize in Ethiopia.

Description

Incomplete-block experiment of maize in Ethiopia.

Usage

data("damesa.maize")

Format

A data frame with 264 observations on the following 8 variables.

- **site**: site, 4 levels
- **rep**: replicate, 3 levels
- **block**: incomplete block
- **plot**: plot number
- **gen**: genotype, 22 levels
- **row**: row ordinate
- **col**: column ordinate
- **yield**: yield, t/ha

Details

An experiment harvested in 2012, evaluating drought-tolerant maize hybrids at 4 sites in Ethiopia. At each site, an incomplete-block design was used.

Damesa et al use this data to compare single-stage and two-stage analyses.

Source

## Examples

```r
## Not run:
library(agridat)
libs(desplot)
desplot(damesa.maize,
   yield ~ col*row|site,
   main="damesa.maize",
   out1=rep, out2=block, num=gen, cex=1)

if(0){
  # Fit the single-stage model in Damesa
  lib(asreml)
  m0 <- asreml(data=damesa.maize,
               fixed = yield ~ gen,
               random = ~ site + gen:site + at(site):rep/block,
               residual = ~ dsum( ~ units|site) )
lucid::vc(m0) # match Damesa table 1 column 3
  ## effect component std.error z.ratio bound
  ## at(site, S1):rep 0.08819 0.1814 0.49 P 0
  ## at(site, S2):rep 1.383 1.426 0.97 P 0
  ## at(site, S3):rep 0 NA NA B 0
  ## at(site, S4):rep 0.01442 0.02602 0.55 P 0
  ## site 10.45 8.604 1.2 P 0.1
  ## gen:site 0.1054 0.05905 1.8 P 0.1
  ## at(site, S1):rep:block 0.3312 0.3341 0.99 P 0
  ## at(site, S2):rep:block 0.4747 0.1633 2.9 P 0
  ## at(site, S3):rep:block 0 NA NA B 0
  ## at(site, S4):rep:block 0.06954 0.04264 1.6 P 0
  ## site_S1!R 1.346 0.3768 3.6 P 0
  ## site_S2!R 0.1936 0.06628 2.9 P 0
  ## site_S3!R 1.153 0.2349 4.9 P 0
  ## site_S4!R 0.1112 0.03665 3 P 0
}
## End(Not run)
```

---

**Description**

Darwin’s maize data of crossed/inbred plant heights.
Format

A data frame with 30 observations on the following 4 variables.

pot  Pot factor, 4 levels
pair Pair factor, 12 levels
type Type factor, self-pollinated, cross-pollinated
height Height, in inches (measured to 1/8 inch)

Details

Charles Darwin, in 1876, reported data from an experiment that he had conducted on the heights of corn plants. The seeds came from the same parents, but some seeds were produced from self-fertilized parents and some seeds were produced from cross-fertilized parents. Pairs of seeds were planted in pots. Darwin hypothesized that cross-fertilization produced produced more robust and vigorous offspring.

Darwin wrote, "I long doubted whether it was worth while to give the measurements of each separate plant, but have decided to do so, in order that it may be seen that the superiority of the crossed plants over the self-fertilised, does not commonly depend on the presence of two or three extra fine plants on the one side, or of a few very poor plants on the other side. Although several observers have insisted in general terms on the offspring from intercrossed varieties being superior to either parent-form, no precise measurements have been given;* and I have met with no observations on the effects of crossing and self-fertilising the individuals of the same variety. Moreover, experiments of this kind require so much time--mine having been continued during eleven years--that they are not likely soon to be repeated."

Darwin asked his cousin Francis Galton for help in understanding the data. Galton did not have modern statistical methods to approach the problem and said, "I doubt, after making many tests, whether it is possible to derive useful conclusions from these few observations. We ought to have at least 50 plants in each case, in order to be in a position to deduce fair results".

Later, R. A. Fisher used Darwin's data in a book about design of experiments and showed that a t-test exhibits a significant difference between the two groups.

Source


References


Examples

```r
library(agridat)
data(darwin.maize)
dat <- darwin.maize
```
# Compare self-pollination with cross-pollination
libs(lattice)
bwplot(height~type, dat, main="darwin.maize")

## Not run:
libs(reshape2)
dm <- melt(dat)
d2 <- dcast(dm, pot+pair~type)
d2$diff <- d2$cross-d2$self
t.test(d2$diff)
## One Sample t-test
## t = 2.148, df = 14, p-value = 0.0497
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 0.003899165 5.229434169
## End(Not run)

dasilva.maize

Multi-environment trial of maize

Description

Multi-environment trial of maize with 3 reps.

Usage

data("dasilva.maize")

Format

A data frame with 1485 observations on the following 4 variables.

env environment
rep replicate block, 3 per env
gen genotype
yield yield (tons/hectare)

Details

Each location had 3 blocks. Block numbers are unique across environments.

NOTE! The environment codes in the supplemental data file of da Silva 2015 do not quite match the environment codes of the paper, but are mostly off by 1.

DaSilva Table 1 has a footnote "Machado et al 2007". This reference appears to be:

In DaSilva Table 1, the mean of E1 is 10.803. This appears to be a copy of the mean from row 1 of Table 1 in Machado. Using the supplemental data from this paper, the correct mean is 8.685448.

Source


Used via Creative Commons Attribution License.

References


Examples

library(agridat)
data(dasilva.maize)
dat <- dasilva.maize

# Try to match Table 1 of da Silva 2015.
# aggregate(yield ~ env, data=dat, FUN=mean)
##   env  yield
## 1   E1 6.211817 # match E2 in Table 1
## 2   E2 4.549104 # E3
## 3   E3 5.152254 # E4
## 4   E4 6.245904 # E5
## 5   E5 8.084609 # E6
## 6   E6 13.191890 # E7
## 7   E7 8.895721 # E8
## 8   E8 8.685448 # E9
## 9   E9 8.737089

# Unable to match CVs in Table 2, but who knows what they used
# for residual variance.
# aggregate(yield ~ env, data=dat, FUN=function(x) 100*sd(x)/mean(x))

# Match DaSilva supplement 2, ANOVA
# m1 <- aov(yield ~ env + gen + rep:env + gen:env, dat)
# anova(m1)
## Response: yield
##    Df Sum Sq Mean Sq  F value    Pr(>F)
## env  8 8994.2 1124.28  964.1083 < 2.2e-16 ***
## gen 54  593.5  10.99   9.4247 < 2.2e-16 ***
## env:rep 18  57.5  3.19   2.7390 0.0001274 ***
## env:gen 432  938.1 2.17  1.8622 1.825e-15 ***
## Residuals 972 1133.5 1.17
Description

Uniformity trial of soybean in Brazil, 1970.

Usage

data("dasilva.soybean.uniformity")

Format

A data frame with 1152 observations on the following 3 variables.

row row
col column
yield yield, grams/plot

Details

Field length: 48 rows * .6 m = 28.8 m
Field width: 24 columns * .6 m = 14.4 m

Source


References

None

Examples

## Not run:

library(agridat)
data(dasilva.soybean.uniformity)
dat <- dasilva.soybean.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
flip=TRUE, aspect=28.8/14.4,
main="dasilva.soybean.uniformity")
## davidian.soybean

### Description

Growth of soybean varieties in 3 years

### Usage

```r
data("davidian.soybean")
```

### Format

A data frame with 412 observations on the following 5 variables.

- **plot**: plot code
- **variety**: variety, F or P
- **year**: 1988-1990
- **day**: days after planting
- **weight**: weight of soybean leaves

### Details

This experiment compared the growth patterns of two genotypes of soybean varieties: F=Forrest (commercial variety) and P=Plant Introduction number 416937 (experimental variety).

Data were collected in 3 consecutive years.

At the start of each growing season, 16 plots were seeded (8 for each variety). Data were collected approximately weekly. At each timepoint, six plants were randomly selected from each plot. The leaves from these 6 plants were weighed, and average leaf weight per plant was reported. (We assume that the data collection is destructive and different plants are sampled at each date).

Note: this data is the same as the "nlme::Soybean" data.

### Source


Electronic version retrieved from https://www4.stat.ncsu.edu/~davidian/data/soybean.dat

### References

Examples

```r
## Not run:

library(agridat)
data(davidian.soybean)
dat <- davidian.soybean
dat$year <- factor(dat$year)

library(lattice)
xyplot(weight ~ day|variety*year, dat,
      group=plot, type='l',
      main="davidian.soybean")

# The only way to keep your sanity with nlme is to use groupedData objects
# Well, maybe not. When I use "devtools::run_examples",
# the "groupedData" function creates a dataframe with/within(?) an
# environment, and then "nlsList" cannot find datg, even though
# ls() shows datg is visible and head(datg) is fine.
# Also works fine in interactive mode. It is driving me insane.
# reid.grasses has the same problem
# Use if(0){} to block this code from running.
if(0){
  library(nlme)
datg <- groupedData(weight ~ day|plot, dat)
  # separate fixed-effect model for each plot
  # 1988P6 gives unusual estimates
  m1 <- nlsList(SSlogis, data=datg,
                subset = plot != "1988P6")
  # plot(m1) # seems heterogeneous
  plot(intervals(m1), layout=c(3,1)) # clear year,variety effects in Asym

  # A = maximum, B = time of half A = steepness of curve
  # C = sharpness of curve (smaller = sharper curve)

  # switch to mixed effects
  m2 <- nlme(weight ~ A / (1+exp(-(day-B)/C)),
             data=datg,
             fixed=list(A ~ 1, B ~ 1, C ~ 1),
             random = A +B +C ~ 1,
             start=list(fixed = c(17,52,7.5))) # no list!

  # add covariates for A,B,C effects, correlation, weights
  # not necessarily best model, but it shows the syntax
  m3 <- nlme(weight ~ A / (1+exp(-(day-B)/C)),
             data=datg,
             fixed=list(A ~ variety + year,
                        B ~ year,
                        C ~ year),
             random = A +B +C ~ 1,
             start=list(fixed= c(19,0,0,0,
                                 55,0,0,
                                 7,0,0),
                        correlation = c(1, -0.5, 0, -0.5, 1, -0.5, 0, -0.5, 1),
                        weights = 1) # not list!
```
davies.pasture.uniformity

Uniformity trial of pasture.

Description
Uniformity trial of pasture in Australia.

Usage
data("davies.pasture.uniformity")

Format
A data frame with 760 observations on the following 3 variables.

row row
col column
yield yield per plot, grams

Details
Conducted at the Waite Agricultural Research Institute in 1928. A rectangle 250 x 200 links was selected, divided into 1000 plots measuring 10 x 5 links, that is 1/2000th acre. Plots were hand harvested for herbage and air-dried. Cutting began Tue, 25 Sep and ended Sat, 29 Sep, by which time 760 plots had been harvested. Rain fell, harvesting ceased.

The minimum recommended plot size is 150 square links. The optimum recommended plot size is 450 square links, 5 x 90 links in size.

Note, there were 4 digits that were hard to read in the original document. Best estimates of these digits were used for the yields of the affects plots. The yields were watermarked with an extra .01 added.

The botanical composition of species clearly influenced the total herbage.

Field length: 40 plots * 5 links = 200 links
Field width: 19 plots * 10 links = 190 links
day.wheat.uniformity

Source


References

None

Examples

```r
## Not run:
library(agridat)

data(davies.pasture.uniformity)
dat <- davies.pasture.uniformity

# range(dat$yield) # match Davies
# mean(dat$yield) # 227.77, Davies has 221.7
# sd(dat$yield)/mean(dat$yield) # 33.9, Davies has 32.5

# libs(lattice)
# qqmath(~ yield, dat) # clearly non-normal, skewed right

libs(desplot)
desplot(dat, yield ~ col*row,
    flip=TRUE, aspect=(40*5)/(19*10), # true aspect
    main="davies.pasture.uniformity")

## End(Not run)
```

day.wheat.uniformity  Uniformity trial of wheat

Description

Uniformity trial of wheat in 1903 in Missouri.

Usage

```r
data("day.wheat.uniformity")
```

Format

A data frame with 3090 observations on the following 4 variables.

row  row
col  col
Details

These data are from the Shelbina field of the Missouri Agricultural Experiment Station. The field (plat) was about 1/4 acre in area and apparently uniform throughout. In the fall of 1912, wheat was drilled in rows 8 inches apart, each row 155 feet long. The wheat was harvested in June, in 5-foot segments. The gross weight and the grain weight was measured, the straw weight was calculated by subtraction.

Field width: 31 series * 5 feet = 155 feet
Field length: 100 rows, 8 inches apart = 66.66 feet

Source

James Westbay Day (1916). The relation of size, shape, and number of replications of plats to probable error in field experimentation. Dissertation, University of Missouri. Table 1, page 22. https://hdl.handle.net/10355/56391

References

James W. Day (1920). The relation of size, shape, and number of replications of plats to probable error in field experimentation. Agronomy Journal, 12, 100-105. https://doi.org/10.2134/agronj1920.00021962001200030002x

Examples

```r
## Not run:

library(agridat)
data(day.wheat.uniformity)
dat <- day.wheat.uniformity

libs(desplot)
desplot(dat, grain~col*row,
    flip=TRUE, aspect=(100*8)/(155*12), # true aspect
    main="day.wheat.uniformity - grain yield")

# similar to Day table IV
libs(lattice)
xyplot(grain~straw, data=dat, main="day.wheat.uniformity", type=c('p','r'))
# cor(dat$grain, dat$straw) # .9498 # Day calculated 0.9416

libs(desplot)
desplot(dat, straw~col*row,
    flip=TRUE, aspect=(100*8)/(155*12), # true aspect
    main="day.wheat.uniformity - straw yield")

# Day fig 2
coldat <- aggregate(grain~col, dat, sum)
xyplot(grain ~ col, coldat, type='l', ylim=c(2500,6500))
```
dat$rowgroup <- round((dat$row +1)/3,0)
rowdat <- aggregate(grain~rowgroup, dat, sum)
xyplot(grain ~ rowgroup, rowdat, type='l', ylim=c(2500,6500))

## End(Not run)

denis.missing

Multi-environment trial with structured missing values

Description
Grain yield was measured on 5 genotypes in 26 environments. Missing values were non-random, but structured.

Format

- env  environment, 26 levels
- gen  genotype factor, 5 levels
- yield yield

Used with permission of Jean-Baptists Denis.

Source

References

Examples

## Not run:

library(agridat)
data(denis.missing)
dat <- denis.missing

# view missingness structure
libs(reshape2)
acast(dat, env~gen, value.var='yield')

libs(lattice)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(yield ~ gen*env, data=dat, 
    col.regions=redblue,)
main="denis.missing - incidence heatmap"

# stability variance (Table 3 in Piepho)
libs(nlme)
m1 <- lme(yield ~ -1 + gen, data=dat, random= ~ 1|env,
       weights = varIdent(form= ~ 1|gen),
       na.action=na.omit)
svar <- m1.sigma^2 * c(1, coef(m1$modelStruct$varStruct, unc = FALSE))^2
round(svar, 2)
## G5 G3 G1 G2
## 39.25 22.95 54.36 12.17 23.77

## End(Not run)

denis.ryegrass                  Multi-environment trial of perennial ryegrass in France

Description

Plant strength of perennial ryegrass in France for 21 genotypes at 7 locations.

Format

A data frame with 147 observations on the following 3 variables.

   gen genotype, 21 levels
   loc location, 7 levels
   strength average plant strength * 100

Details

INRA conducted a breeding trial in western France with 21 genotypes at 7 locations. The observed data is 'strength' averaged over 7-10 plants per plot and three plots per location (after adjusting for blocking effects). Each plant was scored on a scale 0-9.

The original data had a value of 86.0 for genotype G1 at location L4–this was replaced by an additive estimated value of 361.2 as in Gower and Hand (1996).

Source


References

Examples

```r
## Not run:
library(agridat)
data(denis.ryegrass)
dat <- denis.ryegrass

# biplots (without ellipses) similar to Denis figure 1
libs(gge)
m1 <- gge(dat, strength ~ gen*loc, scale=FALSE)
biplot(m1, main="denis.ryegrass biplot")

## End(Not run)
```

Description

Latin square of four breeds of sheep with four diets

Usage

```r
data("depalluel.sheep")
```

Format

A data frame with 32 observations on the following 5 variables.

food   diet
animal animal number
breed   sheep breed
weight  weight, pounds
date    months after start

Details

This may be the earliest known Latin Square experiment.
Four sheep from each of four breeds were randomized to four feeds and four slaughter dates.
Sheep that eat roots will eat more than sheep eating corn, but each acre of land produces more roots
than corn.

de Palleuel said: In short, by adopting the use of roots, instead of corn, for the fattening of all sorts
of cattle, the farmers in the neighborhood of the capital will not only gain great profit themselves,
but will also very much benefit the public by supplying this great city with resources, and preventing
the sudden rise of meat in her markets, which is often considerable.
Source


References

None

Examples

```r
library(agridat)
data(devries.pine)
dat <- devries.pine

# Not the best view...weight gain is large in the first month, then slows down
# and the linear line hides this fact
lib(lattice)
xyplot(weight ~ date|food, dat, group=animal, type='l', auto.key=list(columns=4),
xlab="Months since start",
main="devries.pine")
```

---

**devries.pine**  
*Graeco-Latin Square experiment in pine*

Description

Graeco-Latin Square experiment in pine

Usage

```r
data("devries.pine")
```

Format

A data frame with 36 observations on the following 6 variables.

- **block**  block
- **row**  row
- **col**  column
- **spacing**  spacing treatment
- **thinning**  thinning treatment
- **volume**  stem volume in m^3/ha
- **growth**  annual stem volume increment m^3/ha at age 11
Details

Experiment conducted on Caribbean Pine at Coebiti in Surinam (Long 55 28 30 W, Lat 5 18 5 N). Land was cleared in Jan 1965 and planted May 1965. Each experimental plot was 60m x 60m. Roads 10 m wide run between the rows. Each block is thus 180m wide and 200m deep. Data were collected only on 40m x 40m plots in the center of each experimental unit. Plots were thinned in 1972 and 1975. The two treatment factors (spacing, thinning) were assigned in a Graeco-Latin Square design.

Spacing: A=2.5, B=3, C=3.5. Thinning: Z=low, M=medium, S=heavy.

Field width: 4 blocks x 180 m = 720 m
Field length: 1 block x 200 m = 200 m.

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(devries.pine)
dat <- devries.pine
libs(desplot)
desplot(dat, volume ~ col*row,
       main="devries.pine - expt design and tree volume",
       col=spacing, num=thinning, cex=1, out1=block, aspect=200/720)

libs(HH)
HH::interaction2wt(volume ~ spacing+thinning, dat,
       main="devries.pine")

# ANOVA matches appendix 5 of DeVries
m1 <- aov(volume ~ block + spacing + thinning + block:factor(row) +
          block:factor(col), data=dat)
anova(m1)

## End(Not run)
```
**Description**

Yield of 10 spring wheat varieties for 17 locations in 1976.

**Format**

A data frame with 134 observations on the following 3 variables.

- `gen` genotype, 10 levels
- `env` environment, 17 levels
- `yield` yield (t/ha)

**Details**

Yield of 10 spring wheat varieties for 17 locations in 1976.

Used to illustrate modified joint regression.

**Source**


**References**


**Examples**

```r
## Not run:
library(agritat)
data(digby.jointregression)
dat <- digby.jointregression

# Simple gen means, ignoring unbalanced data.
# Matches Digby table 2, Unadjusted Mean
round(tapply(dat$yield, dat$gen, mean), 3)

# Two-way model. Matches Digby table 2, Fitting Constants
m00 <- lm(yield ~ 0 + gen + env, dat)
round(coef(m00)[1:10] - 2.756078 + 3.272, 3) # Adjust intercept
# genG01 genG02 genG03 genG04 genG05 genG06 genG07 genG08 genG09 genG10
```
n.gen <- nlevels(dat$gen)
n.env <- nlevels(dat$env)

# Estimate theta (env eff)
m0 <- lm(yield ~ -1 + env + gen, dat)
thetas <- coef(m0)[1:n.env]
# center env effects
thetas <- thetas-mean(thetas)

# Add env effects to the data
dat$theta <- thetas[match(paste("env",dat$env,sep=""), names(thetas))]

# Initialize beta (gen slopes) at 1
betas <- rep(1, n.gen)

done <- FALSE
while(!done){
  betas0 <- betas
  # M1: Fix thetas (env effects), estimate beta (gen slope)
m1 <- lm(yield ~ -1 + gen + gen:theta, data=dat)
betas <- coef(m1)[-c(1:n.gen)]
dat$beta <- betas[match(paste("gen",dat$gen,":theta",sep=""), names(betas))]
  # print(betas)

  # M2: Fix betas (gen slopes), estimate theta (env slope)
m2 <- lm(yield ~ env:beta + gen -1, data=dat)
theta[is.na(theta)] <- 0 # Change last coefficient from NA to 0
dat$theta <- thetas[match(paste("env",dat$env,":beta",sep=""), names(thetas))]
  # print(thetas)

  # Check convergence
  chg <- sum(((betas-betas0)/betas0)^2)
cat("Relative change in betas",chg,"\n")
  if(chg < .0001) done <- TRUE
}

libs(lattice)
xyplot(yield ~ theta|gen, data=dat, xlab="theta (environment effect)",
main="digby.jointregression - stability plot")

# Dibgy Table 2, modified joint regression
# Genotype sensitivities (slopes)
round(betas,3) # Match Digby table 2, Modified joint regression sensitivity
# genG01 genG02 genG03 genG04 genG05 genG06 genG07 genG08 genG09 genG10
# 0.953 0.739 1.082 1.024 1.142 0.877 1.089 0.914 1.196 0.947

# Env effects. Match Digby table 3, Modified joint reg
round(thetas,3)+1.164-.515 # Adjust intercept to match
# envE01 envE02 envE03 envE04 envE05 envE06 envE07 envE08 envE09 envE10
diggle.cow

Bodyweight of cows in a 2-by-2 factorial experiment

Description

Bodyweight of cows in a 2-by-2 factorial experiment.

Format

A data frame with 598 observations on the following 5 variables.

animal Animal factor, 26 levels
iron Factor with levels Iron, NoIron
infect Factor levels Infected, NonInfected
weight Weight in (rounded to nearest 5) kilograms
day Days after birth

Details

Diggle et al., 1994, pp. 100-101, consider an experiment that studied how iron dosing (none/standard) and micro-organism (infected or non-infected) influence the weight of cows.

Twenty-eight cows were allocated in a 2-by-2 factorial design with these factors. Some calves were inoculated with tuberculosis at six weeks of age. At six months, some calves were maintained on supplemental iron diet for a further 27 months.

The weight of each animal was measured at 23 times, unequally spaced. One cow died during the study and data for another cow was removed.
Examples

```r
## Not run:
library(agridat)
data(diggle.cow)
dat <- diggle.cow

# Figure 1 of Verbyla 1999
libs(latticeExtra)
useOuterStrips(xyplot(weight ~ day|iron*infect, dat, group=animal,
                      type='b', cex=.5,
                      main="diggle.cow")

# Scaling
dat <- transform(dat, time = (day-122)/10)

libs(asreml) # asreml4
## libs(latticeExtra)
## # Smooth for each animal. No treatment effects. Similar to SAS Output 38.6.9
m1 <- asreml(weight ~ 1 + lin(time) + animal + animal:lin(time), data=dat,
             random = ~ animal:spl(time))
p1 <- predict(m1, data=dat, classify="animal:time",
              design.points=list(time=seq(0,65.9, length=50)))
p1 <- p1$pvals
p1 <- merge(dat, p1, all=TRUE) # to get iron/infect merged in
foo1 <- xyplot(weight ~ day|iron*infect, dat, group=animal)
foo2 <- xyplot(predicted.value ~ day|iron*infect, p1, type='l', group=animal)
print(foo1+foo2)

## End(Not run)
```
Description
Uniformity trial of safflower in Arizona in 1958.

Usage
data("draper.safflower.uniformity")

Format
A data frame with 640 observations on the following 4 variables.

- `expt` experiment
- `row` row
- `col` column
- `yield` yield per plot (grams)

Details
Experiments were conducted at the Agricultural Experiment Station Farm at Eloy, Arizona. The crop was harvested in July 1958.
The crop was planted in two rows 12 inches apart on vegetable beds 40 inches center to center.
In each test, the end ranges and one row of plots on one side were next to alleys, and those plots gave estimates of border effects.

**Experiment E4 (four foot test).**
Sandy streaks were present in the field. Average yield was 1487 lb/ac. A diagonal fertility gradient was in this field. Widening the plot was equally effective as lengthening the plot to reduce variability. The optimum plot size was 1 bed wide, 24 feet long. Considering economic costs, the optimum size was 1 bed, 12 feet long.
Field width: 16 beds * 3.33 = 53 feet
Field length: 18 ranges * 4 feet = 72 feet

**Experiment E5 (five foot test).**
Average yield 2517 lb/ac, typical for this crop. Combining plots lengthwise was more effective than widening the plots, in order to reduce variability. The optimum plot size was 1 bed wide, 25 feet long. Considering economic costs, the optimum size was 1 bed, 18 feet long.
Field width: 14 beds * 3.33 = 46.6 feet.
Field length: 18 ranges * 5 feet = 90 feet.
Source


References

None

Examples

```r
## Not run:

library(agridat)

data(draper.safflower.uniformity)
dat4 <- subset(draper.safflower.uniformity, expt=="E4")
dat5 <- subset(draper.safflower.uniformity, expt=="E5")

libs(desplot)
desplot(dat4, yield~col*row,
        flip=TRUE, tick=TRUE, aspect=72/53, # true aspect
        main="draper.safflower.uniformity (four foot)")

desplot(dat5, yield~col*row,
        flip=TRUE, tick=TRUE, aspect=90/46, # true aspect
        main="draper.safflower.uniformity (five foot)")

libs(agricolae)
libs(reshape2)
# Remove border plots at either end and right side
dat4 <- subset(dat4, row>1 & row<20)
dat4 <- subset(dat4, col<17)
dat5 <- subset(dat5, row>1 & row<20)
dat5 <- subset(dat5, col<15)

s4 <- index.smith(acast(dat4, row~col, value.var='yield'),
                   main="draper.safflower.uniformity (four foot)",
                   col="red")$uni
s4 # match Draper table 2, p 22

## s5 <- index.smith(acast(dat5, row~col, value.var='yield'),
##                    main="draper.safflower.uniformity (five foot)",
##                    col="red")$uni
## s5 # match Draper table 1, p 21

## End(Not run)
```
durban.competition  

*Sugar beet yields with competition effects*

**Description**

Sugar beet yields with competition effects

**Format**

A data frame with 114 observations on the following 5 variables.

- `gen` Genotype factor, 36 levels plus Border
- `col` Column
- `block` Row/Block
- `wheel` Position relative to wheel tracks
- `yield` Root yields, kg/plot

**Details**

This sugar-beet trial was conducted in 1979.

Single-row plots, 12 m long, 0.5 m between rows. Each block is made up of all 36 genotypes laid out side by side. Guard/border plots are at each end. Root yields were collected.

Wheel tracks are located between columns 1 and 2, and between columns 5 and 6, for each set of six plots. Each genotype was randomly allocated once to each pair of plots (1,6), (2,5), (3,4) across the three reps. Wheel effect were not significant in _this_ trial.

Field width: 18m + 1m guard rows = 19m

Field length: 3 blocks * 12m + 2*0.5m spacing = 37m Retrieved from https://www.ma.hw.ac.uk/~iain/research/JAgSciData/data/Trial1.dat

Used with permission of Iain Currie.

**Source**


**Examples**

```r
## Not run:

library(agridat)

data(durban.competition)

dat <- durban.competition

# Check that genotypes were balanced across wheel tracks.
with(dat, table(gen,wheel))
```
libs(desplot)
desplot(dat, yield ~ col*block,
    out1=block, text=gen, col=wheel, aspect=37/19, # true aspect
    main="durban.competition")

# Calculate residual after removing block/genotype effects
m1 <- lm(yield ~ gen + block, data=dat)
dat$res <- resid(m1)

## desplot(dat, res ~ col*block, out1=block, text=gen, col=wheel,
## main="durban.competition - residuals")

# Calculate mean of neighboring plots
dat$comp <- NA

# Demonstrate the competition effect
# Competitor plots have low/high yield -> residuals are negative/positive
libs(lattice)
xyplot(res~comp, dat, type=c("p","r"), main="durban.competition",
       xlab="Average yield of neighboring plots", ylab="Residual")

## End(Not run)

durban.rowcol  

Description

Row-column experiment of spring barley, many varieties

Format

A data frame with 544 observations on the following 5 variables.

row row
bed bed (column)
rep rep, 2 levels
gen genotype, 272 levels
yield yield, tonnes/ha
Details

Spring barley variety trial of 272 entries (260 new varieties, 12 control). Grown at the Scottish Crop Research Institute in 1998. Row-column design with 2 reps, 16 rows (north/south) by 34 beds (east/west). The land sloped downward from row 16 to row 1. Plot yields were converted to tonnes per hectare.
Plot dimensions are not given.
Used with permission of Maria Durban.

Source

Durban, Maria and Hackett, Christine and McNicol, James and Newton, Adrian and Thomas, William and Currie, Iain. 2003. The practical use of semiparametric models in field trials, Journal of Agric Biological and Envir Stats, 8, 48-66. https://doi.org/10.1198/1085711031265

References


Examples

```r
# Not run:
library(agridat)
data(durban.rowcol)
dat <- durban.rowcol
libs(desplot)
desplot(dat, yield~bed*row,
       out1=rep, num=gen, # aspect unknown
       main="durban.rowcol")

# Durban 2003 Figure 1
m10 <- lm(yield~gen, data=dat)
dat$resid <- m10$resid
libs(lattice)
## xyplot(resid~row, dat, type=c(‘p’,‘smooth’), main="durban.rowcol")
## xyplot(resid~bed, dat, type=c(‘p’,‘smooth’), main="durban.rowcol")

# Figure 3
libs(lattice)
xyplot(resid ~ bed|factor(row), data=dat,
       main="durban.rowcol",
       type=c(‘p’,‘smooth’))

# Figure 5 - field trend
# note, Durban used gam package like this
```
## durban.splitplot

**Split-plot experiment of barley with fungicide treatments**

### Description

Split-plot experiment of barley with fungicide treatments

### Format

A data frame with 560 observations on the following 6 variables.

- **yield**: yield, tonnes/ha
block block, 4 levels
  gen genotype, 70 levels
  fung fungicide, 2 levels
  row row
  bed bed (column)

Details

Grown in 1995-1996 at the Scottish Crop Research Institute. Split-plot design with 4 blocks, 2 whole-plot fungicide treatments, and 70 barley varieties or variety mixes. Total area was 10 rows (north/south) by 56 beds (east/west).

Used with permission of Maria Durban.

Source


Examples

```r
## Not run:
library(agridat)
data(durban.splitplot)
dat <- durban.splitplot

libs(desplot)
dsplot(dat, yield~bed*row,
  out1=block, out2=fung, num=gen, # aspect unknown
  main="durban.splitplot")

# Durban 2003, Figure 2
m20 <- lm(yield~gen + fung + gen:fung, data=dat)
dat$resid <- m20$resid
# libs(lattice)
## xyplot(resid~row, dat, type=c(’p’,’smooth’), main="durban.splitplot")
## xyplot(resid~bed, dat, type=c(’p’,’smooth’), main="durban.splitplot")

# Figure 4 doesn't quite match due to different break points
# libs(lattice)
xyplot(resid ~ bed|factor(row), data=dat,
  main="durban.splitplot",
  type=c(’p’,’smooth’))

# Figure 6 - field trend
# note, Durban used gam package like this
# m2lo <- gam(yield ~ gen*fung + lo(row, bed, span=.082), data=dat)
```
eden.nonnormal

Height of barley plants in a study of non-normal data

Description

Height of barley plants in a study of non-normal data.

Usage

data("eden.nonnormal")

Format

A data frame with 256 observations on the following 3 variables.

pos position within block
block  block (numeric)
height  height of wheat plant

Details

This data was used in a very early example of a permutation test.

Eden & Yates used this data to consider the impact of non-normal data on the validity of a hypothesis test that assumes normality. They concluded that the skew data did not negatively affect the analysis of variance.

Grown at Rothamsted. Eight blocks of Yeoman II wheat. Sampling of the blocks was quarter-meter rows, four times in each row. Rows were selected at random. Position within the rows was partly controlled to make use of the whole length of the block. Plants at both ends of the sub-unit were measured. Shoot height is measured from ground level to the auricle of the last expanded leaf.

Source


References

Kenneth J. Berry, Paul W. Mielke, Jr., Janis E. Johnston *Permutation Statistical Methods: An Integrated Approach.*

Examples

```r
library(agridat)
data(eden.nonnormal)
dat <- eden.nonnormal
mean(dat$height) # 55.23 matches Eden table 1

# Eden figure 2
libs(dplyr, lattice)
# Blocks had different means, so subtract block mean from each datum
dat <- group_by(dat, block)
dat <- mutate(dat, blkmn=mean(height))
dat <- transform(dat, dev=height-blkmn)

histogram(~ dev, data=dat, breaks=seq(from=-40, to=30, by=2.5),
  xlab="Deviations from block means",
  main="eden.nonnormal - heights skewed left")

## Not run:

# calculate skewness, permutation
libs(dplyr, lattice, latticeExtra)
```
# Eden table 1
# anova(aov(height ~ factor(block), data=dat))

# Eden table 2,3. Note, this may be a different definition of skewness
# than is commonly used today (e.g. e1071::skewness).
skew <- function(x){
  n <- length(x)
  x <- x - mean(x)
  s1 = sum(x)
  s2 = sum(x^2)
  s3 = sum(x^3)
  k3=n/((n-1)*(n-2)) * s3 -3/n*s2*s1 + 2/n^2 * s1^3
  return(k3)
}
# Negative values indicate data are skewed left
dat <- group_by(dat, block)
summarize(dat, s1=sum(height), s2=sum(height^2), mean2=var(height), k3=skew(height))
## block s1 s2 mean2 k3
## <int> <dbl> <dbl> <dbl> <dbl>
## 1 1 1682.0 95929.5 242.56048 -1268.5210
## 2 2 1858.0 111661.5 121.97984 -1751.9919
## 3 3 1809.5 108966.8 214.36064 -3172.5284
## 4 4 1912.0 121748.5 242.14516 -2548.2194
## 5 5 1722.0 99026.5 205.20565 -559.0629
## 6 6 1339.0 63077.0 227.36190 -801.2740
## 7 7 1963.0 123052.5 84.99093 -713.2595
## 8 8 1854.0 112366.0 159.67339 -1061.9919

# Another way to view skewness with qq plot. Panel 3 most skewed.
qqmath(~ dev|factor(block), data=dat,
  as.table=TRUE,
  ylab="Deviations from block means",
  panel = function(x, ...){
    panel.qqmathline(x, ...)
    panel.qqmath(x, ...)
  })

# Now, permutation test.
# Eden: "By a process of amalgamation the eight sets of 32 observations were
# reduced to eight sets of four and the data treated as a potential
# layout for a 32-plot trial".
dat2 <- transform(dat, grp = rep(1:4, each=8))
dat2 <- aggregate(height ~ grp+block, dat2, sum)
dat2$trt <- rep(letters[1:4], 8)
dat2$block <- factor(dat2$block)

# Treatments were assigned at random 1000 times
set.seed(54323)
fobs <- rep(NA, 1000)
for(i in 1:1000){
  # randomize treatments within each block
  # trick from https://stackoverflow.com/questions/25085537
  dat2$trt <- with(dat2, ave(trt, block, FUN = sample))
```r
fobs[i] <- anova(aov(height ~ block + trt, dat2))['trt', 'F value']
```

# F distribution with 3,21 deg freedom
# Similar to Eden's figure 4, but on a different horizontal scale
xval <- seq(from = 0, to = max(fobs), length = 50)
yval <- df(xval, df1 = 3, df2 = 21)

# Re-scale, 10 = max of histogram, 0.7 = max of density
histogram(~ fobs, breaks = xval,
         xlab = "F value",
         main = "Observed (histogram) & theoretical (line) F values") +
         xyplot((10/.7) * yval ~ xval, type = "l", lwd = 2)

## End(Not run)

---

**eden.potato**

*Potato yields in response to potash and nitrogen fertilizer*

**Description**

Potato yields in response to potash and nitrogen fertilizer. Data from Fisher’s 1929 paper *Studies in Crop Variation 6*. A different design was used each year.

**Format**

A data frame with 225 observations on the following 9 variables.

- `year` year/type factor
- `yield` yield, pounds per plot
- `block` block
- `row` row
- `col` column
- `trt` treatment factor
- `nitro` nitrogen fertilizer, cwt/acre
- `potash` potash fertilizer, cwt/acre
- `ptype` potash type

**Details**

The data is of interest to show the gradual development of experimental designs in agriculture.

In 1925/1926 the potato variety was Kerr’s Pink. In 1927 Arran Comrade.

In the 1925a/1926a qualitative experiments, the treatments are O=None, S=Sulfate, M=Muriate, P=Potash manure salts. The design was a Latin Square.

The 1925/1926b/1927 experiments were RCB designs with treatment codes defining the amount and type of fertilizer used. Note: the ‘t’ treatment was not defined in the original paper.
Source


References


Examples

```r
library(agridat)
data(eden.potato)
dat <- eden.potato

# 1925 qualitative
d5a <- subset(dat, year=='1925a')
libs(desplot)
desplot(d5a, trt=col*row,
  text=yield, cex=1, shorten='no', # aspect unknown
  main="eden.potato: 1925 qualitative")
anova(m5a <- aov(yield~trt+factor(row)+factor(col), d5a)) # table 2

# 1926 qualitative
d6a <- subset(dat, year=='1926a')
libs(desplot)
desplot(d6a, trt=col*row,
  text=yield, cex=1, shorten='no', # aspect unknown
  main="eden.potato: 1926 qualitative")
anova(m6a <- aov(yield~trt+factor(row)+factor(col), d6a)) # table 4

# 1925 quantitative
d5 <- subset(dat, year=='1925b')
libs(desplot)
desplot(d5, yield ~ col*row,
  out1=block, text=trt, cex=1, # aspect unknown
  main="eden.potato: 1925 quantitative")

# Trt 't' not defined, seems to be the same as 'a'
libs(lattice)
dotplot(trt-yield|block, d5,
  # aspect unknown
  main="eden.potato: 1925 quantitative")
anova(m5 <- aov(yield-trt+block, d5)) # table 6

# 1926 quantitative
d6 <- subset(dat, year=='1926b')
libs(desplot)
desplot(d6, yield ~ col*row,
```
out1=block, text=trt, cex=1, # aspect unknown
main="eden.potato: 1926 quantitative")
anova(m6 <- aov(yield~trt+block, d6)) # table 7

# 1927 qualitative + quantitative
d7 <- droplevels(subset(dat, year==1927))
libs(desplot)
desplot(d7, yield ~ col*row,
       out1=block, text=trt, cex=1, col=pctype, # aspect unknown
       main="eden.potato: 1927 qualitative + quantitative")

# Table 8. Anova, mean yield tons / acre
anova(m7 <- aov(yield~trt+block+pctype + ptype:potash, d7))
libs(reshape2)
me7 <- melt(d7, measure.vars='yield')
acast(me7, potash~nitro, fun=mean) * 40/2240 # English ton = 2240 pounds
acast(me7, potash~ptype, fun=mean) * 40/2240

---

**eden.tea.uniformity**  
*Uniformity trial of tea*

**Description**

Uniformity trial of tea in Ceylon.

**Usage**

```r
data("eden.tea.uniformity")
```

**Format**

A data frame with 144 observations on the following 4 variables.

- entry entry number
- yield yield
- row row
- col column

**Details**

Tea plucking in Ceylon extended from 20 Apr 1928 to 10 Dec 1929. There were 42 pluckings. The field was divided into 144 plots of 1/72 acre = 605 sq ft. Each plot contained 6 rows of bushes, approximately 42 bushes. (Each row was thus about 7 bushes). Plots in row 12 were at high on a hillside, plots in row 1 were low on the hill. Note: We will assume the plots are roughly square: 6 rows of 7 bushes.

Field width: 12 plots * 24.6 feet = 295 feet
Field length: 12 plots * 24.6 feet = 295 feet
edwards.oats

Source

References
None

Examples

```r
## Not run:

library(agridat)

data(eden.tea.uniformity)
dat <- eden.tea.uniformity

# sum(dat$yield) # 140050.6 matches total yield in appendix A
# mean(dat$yield) # 972.574 match page 5554

m1 <- aov(yield ~ factor(entry) + factor(row) + factor(col), data=dat)
summary(m1)

libs(desplot)
desplot(dat, yield ~ col*row,
       aspect=1,
       main="eden.tea.uniformity")

## End(Not run)
```

---

edwards.oats  Multi-environment trial of oats in United States, 5 locations, 7 years.

Description
Multi-environment trial of oats in 5 locations, 7 years, with 3 replicates in each trial.

Usage
data("edwards.oats")

Format
A data frame with 3694 observations on the following 7 variables.

eid Environment identification (factor)
year Year
Details

This data comes from a breeding program, but does not have the usual pattern of (1) genotypes entering/leaving the program (2) check genotypes that remain throughout the duration of the program. Experiments were conducted by the Iowa State University Oat Variety Trial in the years 1997 to 2003.

In each year there were 40 genotypes, with about 30 released checks and 10 experimental lines. Each genotype appeared in a range of 3 to 34 of the year-loc combinations.

The trials were grown in five locations in Iowa: Ames, Nashua, Crawfordsville, Lewis, Sutherland. In 1998 there was no trial grown at Sutherland. There were 3 blocks in each trial.

Five genotypes were removed from the data because of low yields (and are not included here).

The environment identification values are the same as in Edwards (2006) table 1.

Electronic data supplied by Jode Edwards.

Source


References

None

Examples

## Not run:

```r
library(agridat)
libs(dplyr,lattice, reshape2, stringr)
data(edwards.oats)
dat <- edwards.oats
dat$env <- paste0(dat$year, ".", dat$loc)
dat$eid <- factor(dat$eid)
mat <- reshape2::acast(dat, env ~ gen,
  fun.aggregate=mean, value.var="yield", na.rm=TRUE)
lattice::levelplot(mat, aspect="m",
  main="edwards.oats",
  xlab="environment", ylab="genotype",
  scales=list(x=list(rot=90)))

# Calculate BLUEs of gen/env effects
m1 <- lm(yield ~ gen+eid, dat)
```
```r
gg <- coef(m1)[2:80]
names(gg) <- stringr::str_replace(names(gg), "gen", "")
gg <- c(0, gg)
names(gg)[1] <- "ACStewart"

ee <- coef(m1)[81:113]
names(ee) <- stringr::str_replace(names(ee), "eid", "")
ee <- c(0, ee)
names(ee)[1] <- "1"

# Subtract gen/env coefs from yield values
dat2 <- dat
dat2$gencoef <- gg[match(dat2$gen, names(gg))]
dat2$envcoef <- ee[match(dat2$eid, names(ee))]
dat2 <- dplyr::mutate(dat2, y = yield - gencoef - envcoef)

# Calculate variance for each gen*env. Shape of the graph is vaguely
# similar to Fig 2 of Edwards et al (2006), who used a Bayesian model
dat2 <- group_by(dat2, gen, eid)
dat2sum <- summarize(dat2, stddev = sd(y))
bwplot(stddev ~ eid, dat2sum)

## End(Not run)
```

Description

Corn yield response to nitrogen fertilizer for a single variety of corn at two locations over five years

Format

A data frame with 60 observations on the following 4 variables.

- loc  location, 2 levels
- year  year, 1962-1966
- nitro nitrogen fertilizer kg/ha
- yield  yield, quintals/ha

Details

Corn yield response to nitrogen fertilizer for a single variety of corn at two locations in Tennessee over five years. The yield data is the mean of 9 replicates. The original paper fits quadratic curves to the data. Schabenberger and Pierce fit multiple models including linear plateau. The example below fits a quadratic plateau for one year/loc. In the original paper, the 1965 and 1966 data for the Knoxville location was not used as it appeared that the response due to nitrogen was minimal in
1965 and nonexistent in 1966. The economic optimum can be found by setting the tangent equal to the ratio of (fertilizer price)/(grain price).

**Source**


**References**


**Examples**

```r
## Not run:
library(agridat)
data(engelstad.nitro)
dat <- engelstad.nitro

libs(latticeExtra)
useOuterStrips(xyplot(yield ~ nitro | factor(year)*loc, dat,
    main="engelstad.nitro"))

# Fit a quadratic plateau model to one year/loc
j62 <- droplevels(subset(dat, loc=="Jackson" & year==1962))
# ymax is maximum yield, M is the change point, k affects curvature
m1 <- nls(yield ~ ymax*(nitro > M) +
    (ymax - (k/2) * (M-nitro)^2) * (nitro < M),
    data= j62,
    start=list(ymax=80, M=150, k=.01))

# Plot the raw data and model
newdat <- data.frame(nitro=seq(0,max(dat$nitro)))
p1 <- predict(m1, new=newdat)
plot(yield ~ nitro, j62)
lines(p1 ~ newdat$nitro, col="blue")
    title("engelstad.nitro: quadratic plateau at Jackson 1962")

# Optimum nitro level ignoring prices = 225
coef(m1)["M"]

# Optimum nitro level using $0.11 for N cost, $1.15 for grain price = 140
# Set the first derivative equal to N/corn price, k(M-nitro)=.11/1.15
tmp <- .11/1.15/coef(m1)["k"]

## End(Not run)
```
Description

Yield of 13 hybrids, grown in 10 locations across 2 years. Conducted in Yunnan, China.

Format

A data frame with 260 observations on the following 5 variables.

- gen  genotype
- maturity  maturity, days
- year  year
- loc  location
- yield  yield, Mg/ha

Details

Data are the mean of 3 reps.
These data were used to conduct a stability analysis of yield.
Used with permission of Manjit Kang.

Source


Examples

```r
## Not run:
library(agridat)
data(fan.stability)
dat <- fan.stability
dat$env <- factor(paste(dat$loc, dat$year, sep=""))
libs(lattice)
dotplot(gen~yield|env, dat, main="fan.stability")

libs(reshape2, agricolae)
dm <- acast(dat, gen~env, value.var='yield')
# Use 0.464 as pooled error from ANOVA. Calculate yield mean/stability.
stability.par(dm, rep=3, MSerror=0.464) # Table 5 of Fan et al.

## End(Not run)
```
Description

Wheat experiment augmented with two check varieties in diagonal strips.

Format

A data frame with 180 observations on the following 4 variables.

- row
- col
- gen
- yield

Details

This experiment was conducted by Matthew Reynolds, CIMMYT. There are 180 plots in the field, 60 for the diagonal checks (G121 and G122) and 120 for new varieties.

Federer used this data in multiple papers to illustrate the use of orthogonal polynomials to model field trends that are not related to the genetic effects.

Note: Federer and Wolfinger (2003) provide a SAS program for analysis of this data. However, when the SAS program is used to analyze this data, the results do not match the results given in Federer (1998) nor Federer and Wolfinger (2003). The differences are slight, which suggests a typographical error in the presentation of the data.

The R code below provides results that are consistent with the SAS code of Federer & Wolfinger (2003) when both are applied to this version of the data.

Plot dimensions are not given.

Source


References

Examples

## Not run:

```r
library(agridat)
data(federer.diagcheck)
dat <- federer.diagcheck
dat$check <- ifelse(dat$gen == "G121" | dat$gen="G122", "C","N")

# Show the layout as in Federer 1998.
libs(desplot)
desplot(dat, yield ~ col*row,
  text=gen, show.key=FALSE, # aspect unknown
  shorten="no", col=check, cex=.8, col.text=c("yellow","gray"),
  main="federer.diagcheck")

# Now reproduce the analysis of Federer 2003.

# Only to match SAS results
dat$row <- 16 - dat$row
dat <- dat[order(dat$col, dat$row), ]

# Add row / column polynomials to the data.
# The scaling factors sqrt() are arbitrary, but used to match SAS
nr <- length(unique(dat$row))
nc <- length(unique(dat$col))
rpoly <- poly(dat$row, degree=10) * sqrt(nc)
cpoly <- poly(dat$col, degree=10) * sqrt(nr)
dat <- transform(dat,
  c1 = cpoly[,1], c2 = cpoly[,2], c3 = cpoly[,3],
  c4 = cpoly[,4], c6 = cpoly[,6], c8 = cpoly[,8],
  r1 = rpoly[,1], r2 = rpoly[,2], r3 = rpoly[,3],
  r4 = rpoly[,4], r8 = rpoly[,8], r10 = rpoly[,10])
dat$trtn <- ifelse(dat$gen == "G121" | dat$gen="G122", dat$gen, "G999")
dat$new <- ifelse(dat$gen == "G121" | dat$gen="G122", "N", "Y")
dat <- transform(dat, trtn=factor(trtn), new=factor(new))

m1 <- lm(yield ~ c1 + c2 + c3 + c4 + c6 + c8
  + r1 + r2 + r4 + r8 + r10
  + c1:r1 + c2:r1 + c3:r1 + gen, data = dat)

# To get Type III SS use the following
# libs(car)
# car::Anova(m1, type=3) # Matches PROC GLM output
```

```
#   Sum Sq Df F value  Pr(>F)
# (Intercept) 538948  1 159.580 3.103e-16 ***
# c1          13781  1  4.080  0.0495  0.0494940 *
# c2          51102  1 15.131  0.0003  0.0003354 ***
# c3          45735  1 13.542  0.0006  0.0006332 ***
# c4          24670  1  7.305  0.0097  0.0097349 **
# ...
# lmer
lib(lme4,lucid)

# "group" for all data
dat$one <- factor(rep(1, nrow(dat)))

# lmer with bobyqa (default)
m2b <- lmer(yield ~ trtn + (0 + r1 + r2 + r4 + r8 + r10 +
c1 + c2 + c3 + c4 + c6 + c8 + r1:c1 + r1:c2 + r1:c3 || one) +
    (1|new:gen),
    data = dat,
    control=lmerControl(check.nlev.gtr.1="ignore"))

vc(m2b)
## grp var1 var2 vcov sdcor
## new.gen (Intercept) <NA> 2869 53.57
## one r1:c3 <NA> 5532 74.37
## one.1 r1:c2 <NA> 58230 241.3
## one.2 r1:c1 <NA> 128000 357.8
## one.3 c8 <NA> 6456 80.35
## one.4 c6 <NA> 1400 37.41
## one.5 c4 <NA> 1792 42.33
## one.6 c3 <NA> 2549 50.49
## one.7 c2 <NA> 5942 77.08
## one.8 c1 <NA> 0 0
## one.9 r10 <NA> 1133 33.66
## one.10 r8 <NA> 1355 36.81
## one.11 r4 <NA> 2269 47.63
## one.12 r2 <NA> 241.8 15.55
## one.13 r1 <NA> 9200 95.92
## Residual <NA> <NA> 4412 66.42

# lmer with Nelder_Mead gives 'wrong' results
## m2n <- lmer(yield ~ trtn + (0 + r1 + r2 + r4 + r8 + r10 +
c1 + c2 + c3 + c4 + c6 + c8 + r1:c1 + r1:c2 + r1:c3 || one) +
    (1|new:gen),
    data = dat,
    control=lmerControl(optimizer="Nelder_Mead",
    check.nlev.gtr.1="ignore"))

vc(m2n)
## groups name variance stddev
## new.gen (Intercept) 3228 56.82
## one r1:c3 7688 87.68
## one.1 r1:c2 69750 264.1
## one.2 r1:c1 107400 327.8
## one.3 c8 6787 82.38
## one.4 c6 1636 40.45
## one.5 c4 12270 110.8
## one.6 c3 2686 51.83
## one.7 c2 7645 87.43
## one.8 c1 0 0.0351
## one.9 r10 1976 44.45
## one.10 r8 1241 35.23
## one.11 r4 2811 53.02
## one.12 r2 928.2 30.47
## one.13 r1 10360 101.8
## Residual 4127 64.24

libs(asreml,lucid) # asreml4

m3 <- asreml(yield ~ -1 + trtn, data=dat,
    random = ~ r1 + r2 + r4 + r8 + r10 +
    c1 + c2 + c3 + c4 + c6 + c8 +
    r1:c1 + r1:c2 + r1:c3 + new:gen)

## coef(m3)
## # REML cultivar means. Very similar to Federer table 2.
## rev(sort(round(coef(m3)$fixed[3] + coef(m3)$random[137:256],0)))
## ## gen_G060 gen_G021 gen_G011 gen_G099 gen_G002
## ## 974 949 945 944 942
## ## gen_G118 gen_G058 gen_G035 gen_G111 gen_G120
## ## 938 937 937 933 932
## ## gen_G046 gen_G061 gen_G082 gen_G038 gen_G090
## ## 932 931 927 927 926

## vc(m3)
## ## effect component std.error z.ratio constr
## ## r1!r1.var 9201 13720 0.67 pos
## ## r2!r2.var 241.7 1059 0.23 pos
## ## r4!r4.var 2269 3915 0.58 pos
## ## r8!r8.var 1355 2627 0.52 pos
## ## r10!r10.var 1133 2312 0.49 pos
## ## c1!c1.var 0.01 0 4.8 bound
## ## c2!c2.var 5942 8969 0.66 pos
## ## c3!c3.var 2549 4177 0.61 pos
## ## c4!c4.var 1792 3106 0.58 pos
## ## c6!c6.var 1400 2551 0.55 pos
## ## c8!c8.var 6456 9702 0.67 pos
## ## r1:c1!r1.var 128000 189700 0.67 pos
## ## r1:c2!r1.var 58230 90820 0.64 pos
## ## r1:c3!r1.var 5531 16550 0.33 pos
## ## new:gen!new.var 2869 1367 2.1 pos
## ## R!variance 4412 915 4.8 pos

## End(Not run)

---

federer.tobacco | RCB of tobacco, height plants exposed to radiation

**Description**

RCB of tobacco, height plants exposed to radiation
Format

A data frame with 56 observations on the following 4 variables.

row row
block block, numeric
dose radiation dose, roentgens
height height of 20 plants, cm

Details

An experiment conducted in 1951 and described in Federer (1954). The treatment involved exposing tobacco seeds to seven different doses of radiation. The seedlings were transplanted to the field in an RCB experiment with 7 treatments in 8 blocks. The physical layout of the experiment was in 8 rows and 7 columns.

Shortly after the plants were transplanted to the field it became apparent that an environmental gradient existed. The response variable was the total height (centimeters) of 20 plants.

Source


References


Examples

```r
## Not run:

library(agridat)
data(federer.tobacco)
dat <- federer.tobacco

# RCB analysis. Treatment factor not significant.
dat <- transform(dat, dosef=factor(dose), rowf=factor(row),
                 blockf=factor(block))
m1 <- lm(height ~ blockf + dosef, data=dat)
anova(m1)

# RCB residuals show strong spatial trends
library(desplot)
dat$resid <- resid(m1)
```

```r
desplot(dat, resid ~ row * block,
    # aspect unknown
    main="federer.tobacco")

# Row-column analysis. Treatment now significant
m2 <- lm(height ~ rowf + blockf + dosef, data=dat)
anova(m2)

## End(Not run)
```

**fisher.barley**

Multi-environment trial of 5 barley varieties, 6 locations, 2 years

**Description**

Multi-environment trial of 5 barley varieties, 6 locations, 2 years

**Usage**

```r
data("fisher.barley")
```

**Format**

A data frame with 60 observations on the following 4 variables.

- **yield** yield, bu/ac
- **gen** genotype/variety, 5 levels
- **env** environment/location, 2 levels
- **year** year, 1931/1932

**Details**

Trials of 5 varieties of barley were conducted at 6 stations in Minnesota during the years 1931-1932. This is a subset of Immer’s barley data. The yield values here are totals of 3 reps (Immer gave the average yield of 3 reps).

**Source**

References


Examples

```r
## Not run:

library(agridat)
data(fisher.barley)
dat <- fisher.barley

libs(dplyr,lattice)
# Yates 1938 figure 1. Regression on env mean
# Sum years within loc
dat2 <- aggregate(yield ~ gen + env, data=dat, FUN=sum)
# Avg within env
emn <- aggregate(yield ~ env, data=dat2, FUN=mean)
dat2$envmn <- emn$yield[match(dat2$env, emn$env)]
xyplot(yield ~ envmn, dat2, group=gen, type=c('p','r'),
main="fisher.barley - stability regression",
xlab="Environment total", ylab="Variety mean",
auto.key=list(columns=3))

if(0){
  # calculate stability according to the sum-of-squares approach used by
  # Shukla (1972), eqn 11. match to Shukla, Table 4, M.S. column
  # also matches fernandez, table 3, stabvar column
  libs(dplyr)
dat2 <- dat
dat2 <- group_by(dat2, gen,env)
dat2 <- summarize(dat2, yield=sum(yield)) # means across years
dat2 <- group_by(dat2, env)
dat2 <- mutate(dat2, envmn=mean(yield)) # env means
dat2 <- group_by(dat2, gen)
dat2 <- mutate(dat2, genmn=mean(yield)) # gen means
dat2 <- ungroup(dat2)
dat2 <- mutate(dat2, grandmn=mean(yield)) # grand mean
# correction factor overall
dat2 <- mutate(dat2, cf = sum((yield - genmn - envmn + grandmn)^2))
t=5; s=6 # t genotypes, s environments
dat2 <- group_by(dat2, gen)
dat2 <- mutate(dat2, ss=sum((yield-genmn-envmn+grandmn)^2))
# divide by 6 to scale down to plot-level
dat2 <- mutate(dat2, sig2i = 1/((s-1)*(t-1)*(t-2)) * (t*(t-1)*ss-cf)/6)
dat2[!duplicated(dat2$gen),c('gen','sig2i')]
```
libs(asreml,lucid) # asreml3
# mixed model approach gives similar results (but not identical)

dat2 <- dat
dat2 <- dplyr::group_by(dat2, gen, env)
dat2 <- dplyr::summarize(dat2, yield=sum(yield)) # means across years
dat2 <- dat2[order(dat2$gen),]

# G-side
m1g <- asreml(yield ~ gen, data=dat2,
random = ~ env + at(gen):units,
family=asr_gaussian(dispersion=1.0))
m1g <- update(m1g)
summary(m1g)$varcomp[-1,1:2]/6
  component     std.error
at(gen, Manchuria):units 33.81450 3.27221
at(gen, Peatland):units 70.44891 5.05268
at(gen, Svansota):units 25.27286 2.18920
at(gen, Trebi):units 231.69817 15.07856
at(gen, Velvet):units 13.93264 1.65558
units!R 0.16666 NA

# R-side estimates = G-side estimate + 0.1666 (resid variance)
m1r <- asreml(yield ~ gen, data=dat2,
random = ~ env,
residual = ~ dsum(~ units|gen))
m1r <- update(m1r)
summary(m1r)$varcomp[-1,1:2]/6
  component     std.error
gen_Manchuria!R 34.00058 2.72487
gen_Peatland!R 70.65501 5.05892
gen_Svansota!R 25.42022 2.18860
gen_Trebi!R 231.85846 15.07856
gen_Velvet!R 14.08405 1.65558

## End(Not run)
Description
Latin square experiment on mangolds. Used by R. A. Fisher.

Usage
data("fisher.latin")

Format
A data frame with 25 observations on the following 4 variables.

trt treatment factor, 5 levels
yield yield
row row
col column

Details
Yields are root weights. Data originally collected by Mercer and Hall as part of a uniformity trial. This data is the same as the data from columns 1-5, rows 16-20, of the mercer.mangold.uniformity data in this package.
Unsurprisingly, there are no significant treatment differences.

Source

Examples
library(agridat)
data(fisher.latin)
dat <- fisher.latin

# Standard latin-square analysis
m1 <- lm(yield ~ trt + factor(row) + factor(col), data=dat)
anova(m1)
Calving difficulty by calf sex and age of dam

Description

Calving difficulty by calf sex and age of dam

Usage

data("foulley.calving")

Format

A data frame with 54 observations on the following 4 variables.

- **sex**: calf gender
- **age**: dam age factor, 9 levels
- **score**: score for birthing difficulty, S1 < S2 < S3
- **count**: count of births for each category

Details

These data are calving difficulty scores for purebred US Simmental cows.

The raw data show that the greatest calving difficulty is for young dams with male calves. Differences between male/female calves decreased with age of the dam.

The goodness of fit can be improved by using a scaling effect for age of dam.

Note: The paper by Foulley and Gianola has '21943' as the count for score 1, F, >8. This data uses '20943' so that the marginal totals from this data match the marginal totals given in the paper.

Used with permission of Jean-Louis Foulley.

Source


Examples

```r
## Not run:
library(agridat)
data(foulley.calving)
dat <- foulley.calving

## Plot
d2 <- transform(dat,
    age=ordered(age, levels=c("0.0-2.0","2.0-2.5","2.5-3.0"),
```
foulley.calving

"3.0-3.5","3.5-4.0",
"4.0-4.5","4.5-5.0","5.0-8.0","8.0+"),
score=ordered(score, levels=c('S1','S2','S3'))

libs(reshape2)
d2 <- acast(dat, sex+age~score, value.var='count')
d2 <- prop.table(d2, margin=1)

libs(lattice)

thm <- simpleTheme(col=c('skyblue','gray','pink'))
barchart(d2, par.settings=thm, main="foulley.calving",
 xlab="Frequency of calving difficulty", ylab="Calf gender and dam age",
 auto.key=list(columns=3, text=c("Easy","Assited","Difficult")))

# Ordinal multinomial model

libs(ordinal)
m2 <- clm(score ~ sex*age, data=dat, weights=count, link='probit')

summary(m2)

# Coefficients:
#
# sexM 0.500605 0.015178 32.982 < 2e-16 ***
# age2.0-2.5 -0.237643 0.013846 -17.163 < 2e-16 ***
# age2.5-3.0 -0.681648 0.018894 -36.077 < 2e-16 ***
# age3.0-3.5 -0.957138 0.018322 -52.241 < 2e-16 ***
# age3.5-4.0 -1.082520 0.024356 -44.446 < 2e-16 ***
# age4.0-4.5 -1.146834 0.022496 -50.981 < 2e-16 ***
# age4.5-5.0 -1.175312 0.028257 -41.594 < 2e-16 ***
# age5.0-8.0 -1.280587 0.016948 -75.559 < 2e-16 ***
# age8.0+ -1.323749 0.024079 -54.974 < 2e-16 ***
# sexM:age2.0-2.5 0.003035 0.019333 0.157 0.87527
# sexM:age2.5-3.0 -0.076677 0.026106 -2.937 0.00331 **
# sexM:age3.0-3.5 -0.080657 0.024635 -3.274 0.00106 **
# sexM:age3.5-4.0 -0.135774 0.032927 -4.124 3.73e-05 ***
# sexM:age4.0-4.5 -0.124303 0.029819 -4.169 3.07e-05 ***
# sexM:age4.5-5.0 -0.198897 0.038309 -5.192 2.08e-07 ***
# sexM:age5.0-8.0 -0.135524 0.022804 -5.943 2.80e-09 ***
# sexM:age8.0+ -0.131033 0.031852 -4.114 3.89e-05 ***
# ---
# Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Threshold coefficients:
#
# S1|S2 0.82504 0.01083 76.15
# S2|S3 1.52017 0.01138 133.62

# Note 1.52017 - 0.82504 = 0.695 matches Foulley's '2-3' threshold estimate

predict(m2) # probability of each category
### Description
Wheat yields of 22 varieties at 14 sites in Australia

### Usage
```r
data("fox.wheat")
```

### Format
A data frame with 308 observations on the following 4 variables.
- `gen`: genotype/variety factor, 22 levels
- `site`: site factor, 14 levels
- `yield`: yield, tonnes/ha
- `state`: state in Australia

### Details
The 1975 Interstate Wheat Variety trial in Australia used RCB design with 4 blocks, 22 varieties in 14 sites. Wagga is represented twice, by trials sown in May and June.

The 22 varieties were a highly selected and represent considerable genetic diversity with four different groups. (i) from the University of Sydney: Timson, Songlen, Gamenya. (ii) widely grown on Mallee soils: Heron and Halberd. (iii) late maturing varieties from Victoria: Pinnacle, KL-21, JL-157. (iv) with Mexican parentage: WW-15 and Oxley.

### Source

Electronic version supplied by Jonathan Godfrey.

### Examples
```r
library(agridat)
data(fox.wheat)
dat <- fox.wheat

# Means of varieties. Slight differences from Fox and Rathjen suggest
# they had more decimals of precision than shown.
```
tapply(dat$yield, dat$gen, mean)

# Calculate genotype means, merge into the data
genm <- tapply(dat$yield, dat$gen, mean)
dat$genm <- genm[match(dat$gen, names(genm))]

# Calculate slopes for each site. Matches Fox, Table 2, Col A.
m1 <- lm(yield~site+site:genm, data=dat)
sort(round(coef(m1)[15:28],2), dec=TRUE)

# Figure 1 of Fox
libs(lattice)
xypplot(yield~genm|state, data=dat, type=c('p','r'), group=site,
auto.key=list(columns=4),
main="fox.wheat", xlab="Variety mean across all sites",
ylab="Variety yield at each site within states")

---

garber.multi.uniformity

Uniformity trials of oat hay and wheat grain

---

Description

Uniformity trials of oat hay and wheat grain, at West Virginia Agricultural Experiment Station, 1923-1924, on the same land.

Format

A data frame with 270 observations on the following 4 variables.

row row
col column
oats yield of oat hay
wheat yield of wheat grain

Details

Each plot was 68 feet x 21 feet. After discarding a 3.5 foot border on all sides, the harvested area was 61 feet x 14 feet. The plots were laid out in doubles with a 14-foot roadway between the plots.

Note: The orientation of the plots is a guess.

Field width: 6 plots * 68 feet + 14 ft/roadway * 2 = 436 feet
Field length: 45 plots * 21 feet/plot = 945 feet

In 1923 the field was planted to oats and harvested as hay.
In 1924 the field was planted to wheat and harvested for grain.
Garber: "Plots 211 to 214, and 261 to 264, [note, these are rows 11-14, columns 5-6] inclusive, were eliminated from this study because of the fact that a few years ago a straw stack had stood on or in the vicinity...which undoubtedly accounts for the relatively high yields on plots 261 to 264, inclusive."

Note: The source document contained mean-subtracted yields. The oat yield in row 22, column 5 was given as +59.7. This is obviously incorrect, since the negative yields all end in ".7" and positive yields all ended in ".3". We used -59.7 as the centered yield value and added the mean of 1883.7 to all centered yields to obtain absolute yields.

Note: Another paper by these authors, "A Method of Laying Out Experimental Plats", contains three additional years of uniformity trials, with each crop expressed as a percentage. https://archive.org/details/in.ernet.dli.2015.229

Source

Examples
## Not run:
library(agridat)
data(garber.multi.uniformity)
dat <- garber.multi.uniformity

libs(desplot)
desplot(dat, oats ~ col*row,
  flip=TRUE, tick=TRUE, aspect=945/436, # true aspect
  main="garber.multi.uniformity oats")

desplot(dat, wheat ~ col*row,
  flip=TRUE, tick=TRUE, aspect=945/436, # true aspect
  main="garber.multi.uniformity wheat")

with(dat, cor(oats, wheat)) # = .37 matches Garber

## End(Not run)

gartner.corn  Yield monitor data from a corn field in Minnesota

Description
Yield monitor data from a corn field in Minnesota

Usage
data("gartner.corn")
Format

A data frame with 4949 observations on the following 8 variables.

- long: longitude
- lat: latitude
- mass: grain mass flow per second, pounds
- time: GPS time, in seconds
- seconds: seconds elapsed for each datum
- dist: distance traveled for each datum, in inches
- moist: grain moisture, percent
- elev: elevation, feet

Details

The data was collected 5 Nov 2011 from a corn field south of Mankato, Minnesota, using a combine-mounted yield monitor. https://www.google.com/maps/place/43.9237575,-93.9750632

Each harvested swath was 12 rows wide = 360 inches.

Time 0 is 5 Nov 2011, 12:38:03 Central Time. Time 16359 = 4.54 hours later.

Yield is calculated as total dry weight (corrected to 15.5 percent moisture), divided by 56 pounds (to get bushels), divided by the harvested area. drygrain = \[\text{massflow \times seconds \times (100-\text{moisture}) / (100-15.5)}\] / 56 harvested area = (distance \times swath width) / 6272640 yield = drygrain / area

Source

Originally from University of Minnesota Precision Agriculture Center. https://www.soils.umn.edu/academics/classes/soil411


Used under Creative Commons BY-SA 3.0 license.

References


Examples

```r
## Not run:
library(agridat)
data(gartner.corn)
dat <- gartner.corn

# Calculate yield
dat <- transform(dat, yield=(mass*seconds*(100-moist)/(100-15.5))/56/(dist*360/6272640))
# Delete yield outliers
```
dat <- subset(dat, yield > 50)

# Colors for yield
medy <- median(dat$yield)
cols <- 20
wwidth <- 150
brks <- seq(from = -wwidth/2, to=wwidth/2, length=ncols-1)
brks <- c(-250, brks, 250) # 250 is safe..we cleaned data outside (?50,450)?
yldbrks <- brks + medy
dat <- transform(dat, yldbin = as.numeric(cut(yield, breaks= yldbrks)))

# Add polygons for soil map units
# Go to: https://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx
# Click: Lat and Long. 43.924, -93.975
# Click the little AOI rectangle icon. Drag around the field
# In the AOI Properties, enter the Name: Gartner
# Click the tab Soil Map to see map unit symbols, names
# Click: Download Soils Data. Click: Create Download Link.
# Download the zip file and find the soilmu_a_aoi files.

# Read shape files
libs(rgdal)
shp <- readOGR(system.file(package="agridat", "files"), "gartner.corn")

# Plot the shapefiles first to set up the coordinate system
plot(shp, xlim=range(dat$long), ylim=range(dat$lat))
box() # Add the yield points
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
with(dat, points(long,lat, main="yield heat map",
    col=redblue(ncols)[yldbin], cex=.75, pch=16))
plot(shp, add=TRUE, lwd=2) # Overlay soil polygons on top
title("gartner.corn - yield heatmap with soil map unit symbols")

# Manual annotation of soil map units
text(x = c(-93.97641, -93.97787, -93.97550, -93.97693, -93.97654, -93.97480,
    -93.97375, -93.972163, 43.922163, 43.926427, 43.926993, 43.926631),
    y = c(43.92185, 43.92290, 43.92358, 43.92445, 43.92532, 43.92553,
    43.92568, 43.922163, 43.926427, 43.926993, 43.926631),
    lab=c("110","319","319","230","105C","110","211","211","211","230","105C"))

# Trim off the ends of the field & re-do image above
dat <- subset(dat, lat < 43.925850 & lat > 43.921178)

# Identify the soil type for each yield point
dat$ix <- over(SpatialPoints(dat[, c('long','lat')]),
    SpatialPolygons(shp@polygons))
dat$mu <- shp$data[, "MUSYM"][dat$ix]
# Check the points are properly identified
# with(dat, points(long,lat, col=redblue(ncols)[ix], cex=.75, pch=16))

# Aggregate points by soil type and analyze
tapply(dat$yield, dat$mu, mean)
tapply(dat$yield, dat$mu, sd)
Impact of Bt corn on non-target species

Description
Impact of Bt corn on non-target species

Format
A data frame with 16 observations on the following 3 variables.

- gen: genotype/maize, Bt ISO
- thysan: thysan abundance
- aranei: aranei abundance

Details
The experiment involved comparing a Bt maize and a near-isogenic control variety. Species abundances were measured for Thysanoptera (thrips) and Araneida (spiders) in 8 different plots. Each response is probably a mean across repeated measurements. Used with permission of Achim Gathmann.

Source
Examples

```r
## Not run:

library(agridat)
data(gathmann.bt)
dat <- gathmann.bt

# EDA suggests Bt vs ISO is significant for thysan, not for aranei
libs(lattice)
libs(reshape2)
d2 <- melt(dat, id.var='gen')
bwplot(value ~ gen|variable, d2,
     main="gathmann.bt", ylab="Insect abundance",
     panel=function(x,y,...){
       panel.xyplot(jitter(as.numeric(x)),y,...)
       panel.bwplot(x,y,...)
     },
     scales=list(relation="free"))

if(0){

# ----- Parametric CI. Thysan significant, aranei not significant.
libs(equivalence)

th0 <- with(dat, tost(thysan[1:8], thysan[9:16], alpha=.05, paired=FALSE))
lapply(th0[c("estimate","tost.interval")], round, 2)
# 14.28-8.72=5.56, (2.51, 8.59) # match Gathmann p. 11

ar0 <- with(dat, tost(aranei[1:8], aranei[9:16], alpha=.05, epsilon=.4))
lapply(ar0[c("estimate","tost.interval")], round, 2)
# .57-.47=.10, (-.19, .40) # match Gathmann p. 11

# ----- Non-parametric exact CI. Same result.
libs(coin)

th1 <- wilcox_test(thysan ~ gen, data=dat, conf.int=TRUE, conf.level=0.90)
lapply(confint(th1), round, 2)
# 6.36, (2.8, 9.2) # Match Gathmann p. 11

ar1 <- wilcox_test(aranei ~ gen, data=dat, conf.int=TRUE, conf.level=0.90)
lapply(confint(ar1), round, 2)
# .05 (-.2, .4)

# ----- Log-transformed exact CI. Same result.
th2 <- wilcox_test(log(thysan) ~ gen, data=dat, alternative=c("two.sided"),
                   conf.int=TRUE, conf.level=0.9)
lapply(confint(th2), function(x) round(exp(x),2))
# 1.66, (1.38, 2.31) # Match Gathmann p 11
```
gauch.soy

### Multi-environment trial of soybeans in New York, 1977 to 1988

**Description**

New York soybean yields, 1977 to 1988, for 7 genotypes, 55 environments (9 loc, 12 years), 2-3 reps.

**Format**

A data frame with 1454 observations on the following 4 variables.

- **yield** yield, kg/ha
- **rep** repeated measurement
- **gen** genotype, 7 levels
- **env** environment, 55 levels
- **year** year, 77-88
- **loc** location, 10 levels

**Details**

Soybean yields at 13 percent moisture for 7 genotypes in 55 environments with 4 replicates. Some environments had only 2 or 3 replicates. The experiment was an RCB design, but some plots were missing and there were many other soybean varieties in the experiment. The replications appear in random order and do _NOT_ define blocks. Environment names are a combination of the first letter of the location name and the last two digits of the year. The location codes are: A=Aurora, C=Chazy, D=Riverhead, E=Etna, G=Geneseo, I=Ithica, L=Lockport, N=Canton, R=Romulus, V=Valatie. Plots were 7.6 m long, four rows wide (middle two rows were harvested).

This data has been widely used (in various subsets) to promote the benefits of AMMI (Additive Main Effects Multiplicative Interactions) analyses.

The gen x env means of Table 1 (Zobel et al 1998) are least-squares means (personal communication).


Used with permission of Hugh Gauch.
Source


References

None

Examples

```r
## Not run:
library(agridat)
data(gauch.soy)
dat <- gauch.soy

## dat <- transform(dat,
## # year = substring(env, 2),
## loc = substring(env, 1, 1))

# AMMI biplot
libs(agricolae)
# Figure 1 of Zobel et al 1988, means vs PC1 score
dat2 <- droplevels(subset(dat, is.element(env, c("A77","C77","V77","V78","A79","C79","G79","R79","V79","A80","C80","G80","L80","D80","R80","V80","A81","C81","G81","L81","D81","R81","V81","A82","L82","G82","V82","A83","I83","G83","A84","N84","C84","I84","G84"))))
m2 <- with(dat2, AMMI(env, gen, rep, yield))
bip <- m2$biplot
with(bip, plot(yield, PC1, type="n", main="gauch.soy -- AMMI biplot"))
with(bip, text(yield, PC1, rownames(bip),
col=ifelse(bip$type=="GEN", "darkgreen", "blue"),
cex=ifelse(bip$type=="GEN", 1.5, .75)))

## End(Not run)
```

Description

Multi-location/year breeding trial in California

Usage

data("george.wheat")
Format

A data frame with 13996 observations on the following 5 variables.

- gen: genotype number
- year: year
- loc: location
- block: block
- yield: yield per plot

Details

This is a nice example of data from a breeding trial, in which some check genotypes are kept during the whole experiment, while other genotypes enter and leave the breeding program. The data is highly unbalanced with respect to genotypes-by-environments.

Results of late-stage small-trials of 211 genotypes of wheat in California, conducted at 9 locations during the years 2004-2018.

Each trial was an RCB with 4 blocks.

The authors used this data to look at GGE biplots across years and concluded that repeatable genotype-by-location patterns were weak, and therefore the California cereal production region is a large, unstable, mega-environment.

Data downloaded 2019-10-29 from Dryad, https://doi.org/10.5061/dryad.bf8rt6b. Data are public domain.

Source


References

None

Examples

```r
## Not run:
library(agridat)
library(lattice)
library(reshape2)
data(george.wheat)
dat <- george.wheat

dat$env <- paste0(dat$year, ".", dat$loc)
# average reps, cast to matrix
mat <- reshape2::acast(dat, gen ~ env, value.var="yield", fun=mean, na.rm=TRUE)
lattice::levelplot(mat, aspect="m",
main="george.wheat", xlab="genotype", ylab="environment",
scales=list(x=list(cex=.3,rot=90),y=list(cex=.5)))
```
**giles.wheat**

Straw length and ear emergence for wheat genotypes. Data are unbalanced with respect to experiment year and genotype.

**Usage**

```r
data("giles.wheat")
```

**Format**

A data frame with 247 observations on the following 4 variables.

- gen: genotype. Note, this is numeric!
- env: environment
- straw: straw length
- emergence: ear emergence, Julian date

**Details**

Highly unbalanced data of straw length and ear emergence date for wheat genotypes.

The 'genotype' column is called 'Accession number' in original data. The genotypes were chosen to represent the range of variation in the trait.

The Julian date was found to be preferable to other methods (such as days from sowing).

Piepho (2003) fit a bilinear model to the straw emergence data. This is similar to Finlay-Wilkinson regression.

**Source**


**References**

Examples

### Not run:

```r
library(agridat)
data(giles.wheat)
dat <- giles.wheat
dat <- transform(dat, gen=factor(gen))
dat <- transform(dat, env=factor(env))
dat1 <- subset(dat, !is.na(straw))
dat2 <- subset(dat, !is.na(emergence))

# Traits are not related
# with(dat, plot(straw~emergence))

# Show unbalancedness of data
libs(lattice, reshape2)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(acast(dat, gen ~ env, value.var='emergence'),
  col.regions=redblue,
  xlab="genotype", ylab="year",
  main="giles.wheat - straw length")

# ----- Analysis of straw length ----- 

libs(emmeans)

# Mean across years. Matches Piepho Table 7 'Simple'
m1 = lm(straw ~ gen, data=dat1)
emmeans(m1, 'gen')

# Simple two-way model. NOT the bi-additive model of Piepho.
m2 = lm(straw ~ gen + env, data=dat1)
emmeans(m2, 'gen')

# Bi-additive model. Matches Piepho Table 6, rows (c)
libs(gnm)
m3 <- gnm(straw ~ env + Mult(gen,env), data=dat1)
cbind(adjusted=round(fitted(m3),0), dat1)

# ----- Analysis of Ear emergence ----- 

# Simple two-way model.
m4 = lm(emergence ~ 1 + gen + env, data=dat2)
emmeans(m4, c('gen','env')) # Matches Piepho Table 9. rpws (c)
emmeans(m4, 'gen') # Match Piepho table 10, Least Squares column
```

## End(Not run)
Description

An RCB experiment of wheat in South Australia, with strong spatial variation and serpentine row/column effects.

Format

A data frame with 330 observations on the following 5 variables.

col column
row row
rep replicate factor, 3 levels
gen wheat variety, 108 levels
yield yield

Details

A randomized complete block experiment. There are 108 varieties in 3 reps. Plots are 6 meters long, 0.75 meters wide, trimmed to 4.2 meters lengths before harvest.Trimming was done by spraying the wheat with herbicide. The sprayer travelled in a serpentine pattern up and down columns. The trial was sown in a serpentine manner with a planter that seeds three rows at a time (Left, Middle, Right).

Field width 15 columns * 6 m = 90 m
Field length 22 plots * .75 m = 16.5 m

Used with permission of Arthur Gilmour, in turn with permission from Gil Hollamby.

Source


References

N. W. Galwey. 2014. Introduction to Mixed Modelling: Beyond Regression and Analysis of Variance. Table 10.9
Examples

```r
## Not run:

library(agridat)
data(gilmour.serpentine)
dat <- gilmour.serpentine

libs(desplot)
desplot(dat, yield ~ col*row,
    num=gen, show.key=FALSE, out1=rep,
    aspect = 16.5/90, # true aspect
    main="gilmour.serpentine")

# Extreme field trend. Blocking insufficient--needs a spline/smoother
# xyplot(yield~col, data=dat, main="gilmour.serpentine")

# -------------------------------

libs(asreml,lucid) # asreml4

dat <- transform(dat, rowf=factor(row),
    colf=factor(10*(col-8)))
dat <- dat[order(dat$rowf, dat$colf), ] # Sort order needed by asreml

# RCB
m0 <- asreml(yield ~ gen, data=dat, random=~rep)

# Add AR1 x AR1
m1 <- asreml(yield ~ gen, data=dat,
    resid = ~ar1(rowf):ar1(colf))

# Add spline
m2 <- asreml(yield ~ gen + col, data=dat,
    random= ~ spl(col) + colf,
    resid = ~ar1(rowf):ar1(colf))

# Figure 4 shows serpentine spraying
p2 <- predict(m2, data=dat, classify="colf")$pvals
plot(p2$predicted, type="b", xlab="column number", ylab="BLUP")

# Define column code (due to serpentine spraying)
# Rhelp doesn't like double-percent modulus symbol, so compute by hand
dat <- transform(dat, colcode = factor(dat$col-floor((dat$col-1)/4)*4 -1))

m3 <- asreml(yield ~ gen + lin(colf) + colcode, data=dat,
    random= ~ colf + rowf + spl(colf),
    resid = ~ar1(rowf):ar1(colf))

# Figure 6 shows serpentine row effects
p3 <- predict(m3, data=dat, classify="rowf")$pvals
```
Description

Yields for a trial at Slate Hall Farm in 1978.

Format

A data frame with 150 observations on the following 5 variables.

- row
- col
- yield
- gen
- rep

Details

The trial was of spring wheat at Slate Hall Farm in 1978. The experiment was a balanced lattice with 25 varieties in 6 replicates. The 'rep' labels are arbitrary (no rep labels appeared in the source data). Each row within a rep is an incomplete block. The plot size was 1.5 meters by 4 meters.

Field width: 10 plots x 4 m = 40 m
Field length: 15 plots x 1.5 meters = 22.5 m

Source

References

None.

Examples

```r
## Not run:

library(agridat)
data(gilmour.slatehall)
dat <- gilmour.slatehall

libs(desplot)
desplot(dat, yield ~ col * row,
       aspect=22.5/40, num=gen, out1=rep, cex=1,
       main="gilmour.slatehall")

# libs(asreml,lucid) # asreml4

# Model 4 of Gilmour et al 1997
dat <- transform(dat, xf=factor(col), yf=factor(row))
dat <- dat[order(dat$xf, dat$yf), ]
m4 <- asreml(yield ~ gen + lin(row), data=dat,
             random = ~ dev(row) + dev(col),
             resid = ~ ar1(xf):ar1(yf))
# coef(m4)$fixed[1] # linear row
# [1] 31.72252 # (sign switch due to row ordering)

vc(m4)
##  effect component std.error z.ratio bound
## dev(col) 2519 1959 1.3 P 0
## dev(row) 20290 10260 2 P 0
## xf:yf(R) 23950 4616 5.2 P 0
## xf:yf!xf!cor 0.439 0.113 3.9 U 0
## xf:yf!yf!cor 0.125 0.117 1.1 U 0

plot(varioGram(m4), main="gilmour.slatehall")

## End(Not run)
```

---

gomez.fractionalfactorial

*Rice fractional factorial experiment 1/2 2^6.*

**Description**

Rice fractional factorial experiment 1/2 2^6. Two reps with 2 blocks in each rep.
**Format**

A data frame with 64 observations on the following 6 variables.

- **yield**: grain yield in tons/ha
- **rep**: replicate, 2 levels
- **block**: block within rep, 2 levels
- **trt**: treatment, levels (1) to abcdef
- **col**: column position in the field
- **row**: row position in the field

**Details**

Grain yield from a $2^6$ fractional factorial experiment in blocks of 16 plots each, with two replications.

Gomez has some inconsistencies. One example:

Page 171: treatment (1) in rep 1, block 2 and rep 2, block 1.
Page 172: treatment (1) in Rep 1, block 1 and rep 2, block 1.

This data uses the layout shown on page 171.

Used with permission of Kwanchai Gomez.

**Source**


**Examples**

```r
## Not run:

library(agridat)
data(gomez.fractionalfactorial)
dat <- gomez.fractionalfactorial

# trt abcdef has the highest yield
# Gomez, Figure 4.8
libs(desplot)
desplot(dat, yield~col*row,
   # aspect unknown
text=trt, shorten="none", show.key=FALSE, cex=1,
   main="gomez.fractionalfactorial - treatment & yield")

# Split treatment into individual factors
dat <- transform(dat,
   a = -1 + 2 * grepl("a",trt),
   b = -1 + 2 * grepl("b",trt),
   c = -1 + 2 * grepl("c",trt),
)```
\[
\begin{align*}
    d &= -1 + 2 \times \text{grepl('d',trt)}, \\
    e &= -1 + 2 \times \text{grepl('e',trt)}, \\
    f &= -1 + 2 \times \text{grepl('f',trt)}
\end{align*}
\]

# Gomez table 4.24, trt SS totalled together.
# Why didn't Gomez nest block within rep?
\[
\text{m0} \leftarrow \text{lm(yield} \sim \text{rep} \times \text{block + trt, dat)}
\]
\[
\text{anova(m0)}
\]

# Gomez table 4.24, trt SS split apart
\[
\text{m1} \leftarrow \text{lm(yield} \sim \text{rep} \times \text{block + (a+b+c+d+e+f)^3, dat)}
\]
\[
\text{anova(m1)}
\]

\text{libs(FrF2)}
\text{aliases(m1)}
\text{MEPlot(m1, select=3:8,}
\text{main="gomez.fractionalfactorial - main effects plot")}

## End(Not run)

---

## gomez.groupsplit

*Group balanced split-plot design in rice*

### Description

Group balanced split-plot design in rice

### Format

A data frame with 270 observations on the following 7 variables.

- **col** column
- **row** row
- **rep** replicate factor, 3 levels
- **fert** fertilizer factor, 2 levels
- **gen** genotype factor, 45 levels
- **group** grouping (genotype) factor, 3 levels
- **yield** yield of rice

### Details

Genotype group S1 is less than 105 days growth duration, S2 is 105-115 days growth duration, S3 is more than 115 days.

Used with permission of Kwanchai Gomez.
### Examples

```r
## Not run:
library(agridat)
data(gomez.groupsplit)
dat <- gomez.groupsplit

# Gomez figure 3.10. Obvious fert and group effects
libs(desplot)
desplot(dat, group ~ col*row,
       out1=rep, col=fert, text=gen, # aspect unknown
       main="gomez.groupsplit")

# Gomez table 3.19 (not partitioned by group)
m1 <- aov(yield ~ fert*group + gen:group + fert:gen:group +
           Error(rep/fert/group), data=dat)
summary(m1)

## End(Not run)
```

---

**gomez.heterogeneity**  
*RCB experiment of rice, heterogeneity of regressions*

### Description

RCB experiment of rice, heterogeneity of regressions

### Usage

```r
data("gomez.heterogeneity")
```

### Format

- **gen**: genotype
- **yield**: yield kg/ha
- **tillers**: tillers no/hill

### Details

An experiment with 3 genotypes to examine the relationship of yield to number of tillers.  
Used with permission of Kwanchai Gomez.
gomez.heteroskedastic

Source


References

None.

Examples

```r
## Not run:

library(agridat)
data(gomez.heterogeneity)
dat <- gomez.heterogeneity

libs(lattice)
xyplot(yield ~ tillers, dat, groups=gen,
type=c("p","r"),
main="gomez.heterogeneity")

## End(Not run)
```

gomez.heteroskedastic  RCB experiment of rice, heteroskedastic varieties

Description

RCB experiment of rice, heteroskedastic varieties

Usage

data("gomez.heteroskedastic")

Format

A data frame with 105 observations on the following 4 variables.

- gen  genotype
- group group of genotypes
- rep  replicate
- yield yield

Details

RCB design with three reps. Genotypes 1-15 are hybrids, 16-32 are parents, 33-35 are checks. Used with permission of Kwanchai Gomez.
Source


References

None.

Examples

```r
## Not run:
library(agridat)
data(gomez.heteroskedastic)
dat <- gomez.heteroskedastic

# Fix the outlier as reported by Gomez p. 311
dat[dat$gen=="G17" & dat$rep=="R2","yield"] <- 7.58

library(lattice)
bwplot(gen ~ yield, dat, group=as.numeric(dat$group),
       ylab="genotype", main="gomez.heterogeneous")

# Match Gomez table 7.28
m1 <- lm(yield ~ rep + gen, data=dat)
anova(m1)
## Response: yield
## Df Sum Sq Mean Sq F value Pr(>F)         
## rep   2 3.306 1.65304  5.6164 0.00553  **
## gen  34 40.020 1.17705  3.9992 5.806e-07 ***
## Residuals 68 20.014 0.29432

## End(Not run)
```

gomez.multilocsplitplot

Multi-environment trial of rice, split-plot design

Description

Grain yield was measured at 3 locations with 2 reps per location. Within each rep, the main plot was 6 nitrogen fertilizer treatments and the sub plot was 2 rice varieties.

Format

A data frame with 108 observations on the following 5 variables.

toc  location, 3 levels
nitro  nitrogen in kg/ha
rep  replicate, 2 levels
gen  genotype, 2 levels
yield  yield, kg/ha

Used with permission of Kwanchai Gomez.

Source

Examples

library(agridat)
data(gomez.multilocsplitplot)
dat <- gomez.multilocsplitplot
dat$nf <- factor(dat$nitro)

# Gomez figure 8.3
libs(lattice)
xyplot(yield~nitro, dat, group=loc, type=c('p','smooth'), auto.key=TRUE,
   main="gomez.multilocsplitplot")

# AOV
# Be careful to use the right stratum, 'nf' appears in both strata.
# Still not quite the same as Gomez table 8.21

## Not run:
# F values are somewhat similar to Gomez Table 8.21
libs(lme4)
m2 <- lmer(yield ~ loc * nf * gen + (1|loc/rep/nf),
   data=dat)
anova(m2)

## Analysis of Variance Table
## Df Sum Sq Mean Sq F value
## loc 2 117942 58971 0.1525
## nf 5 72841432 14568286 37.6777
## gen 1 7557570 7557570 19.5460
## loc:nf 10 10137188 1013719 2.6218
## loc:gen 2 4270469 2135235 5.5223
## nf:gen 5 1501767 300353 0.7768
## loc:nf:gen 10 1502273 150227 0.3885

## End(Not run)
Description

Soil nitrogen at three times for eight fertilizer treatments

Format

A data frame with 96 observations on the following 4 variables.

trt  nitrogen treatment factor
nitro  soil nitrogen content, percent
rep  replicate
stage  growth stage, three periods

Details

Eight fertilizer treatments were tested.
Soil nitrogen content was measured at three times. P1 = 15 days post transplanting. P2 = 40 days post transplanting. P3 = panicle initiation.
Used with permission of Kwan chai Gomez.

Source


References


Examples

## Not run:

library(agridat)
data(gomez.nitrogen)
dat <- gomez.nitrogen

# Note the depletion of nitrogen over time (stage)
interaction2wt(nitro ~ rep/trt + trt*stage, data=dat,
x.between=0, y.between=0,
main="gomez.nitrogen")
# Just the fertilizer profiles
with(dat, interaction.plot(stage, trt, nitro,
    col=1:4, lty=1:3, main="gomez.nitrogen",
    xlab="Soil nitrogen at three times for each treatment"))

# Gomez table 6.16
m1 <- aov(nitro ~ Error(rep/trt) + trt*stage, data=dat)
summary(m1)

# Gomez table 6.18
# Treatment  1 2 3 4 5 6 7 8
cont <- cbind("T7 vs others" = c(1, 1, 1, 1, 1, 1, -7, 1),
    "T8 vs others" = c(1, 1, 1, 1, 1, 1, 0, -6),
    "T2,T5 vs others" = c(-1, 2,-1,-1, 2,-1, 0, 0),
    "T2 vs T5" = c(0, 1, 0, 0,-1, 0, 0, 0))
contrasts(dat$trt) <- cont
contrasts(dat$trt)

m2 <- aov(nitro ~ Error(rep/trt) + trt*stage, data=dat)
summary(m2, expand.split=FALSE,
    split=list(trt=list(
    "T7 vs others"=1,
    "T8 vs others"=2,
    "T2,T5 vs others"=3,
    "T2 vs T5"=4,
    rest=c(5,6,7)),
    "trt:stage"=list(
    "(T7 vs others):P"=c(1,8),
    "(T8 vs others):P"=c(2,9),
    "(T2,T5 vs others):P"=c(3,10),
    "(T2 vs T5):P"=c(4,11),
    "rest:P"=c(5,6,7,12,13,14))
))

## End(Not run)

gomez.nonnormal1

Insecticide treatment effectiveness

Description
Insecticide treatment effectiveness

Usage
data("gomez.nonnormal1")
**Format**

A data frame with 36 observations on the following 3 variables.

- `trt` insecticidal treatment
- `rep` replicate
- `larvae` number of larvae

**Details**

Nine treatments (including the control, T9) were used on four replicates. The number of living insect larvae were recorded.

The data show signs of non-normality, and a log transform was used by Gomez.

Used with permission of Kwanchai Gomez.

**Source**


**References**

None.

**Examples**

```r
## Not run:
library(agridat)
data(gomez.nonnormal1)
dat <- gomez.nonnormal1

# Gomez figure 7.3
# libs(dplyr)
# dat2 <- dat
# dat2 <- summarize(dat2, mn=mean(larvae), rng=diff(range(larvae)))
# plot(rng ~ mn, data=dat2,
# xlab="mean number of larvae", ylab="range of number of larvae",
# main="gomez.nonnormal1")

# Because some of the original values are less than 10,
# the transform used is log10(x+1) instead of log10(x).
dat <- transform(dat, tlarvae=log10(larvae+1))

# QQ plots for raw/transformed data
libs(reshape2, lattice)
qqmath(~ variable, data=melt(dat),
main="gomez.nonnormal1 - raw/transformed QQ plot",
scales=list(relation="free"))
```
```r
# Gomez table 7.16
m1 <- lm(tlarvae ~ rep + trt, data=dat)
anova(m1)
## Response: tlarvae
## Df Sum Sq Mean Sq F value Pr(>F)
## rep  3 0.9567 0.31889 3.6511 0.0267223 *
## trt  8 3.9823 0.49779 5.6995 0.0004092 ***
## Residuals 24 2.0961 0.08734
##
## End(Not run)
```

### Description

RCB experiment of rice, measuring white heads

### Usage

```r
data("gomez.nonnormal2")
```

### Format

A data frame with 42 observations on the following 3 variables.

- **gen**: genotype
- **rep**: replicate
- **white**: percentage of white heads

### Details

The data are the percent of white heads from a rice variety trial of 14 varieties with 3 reps. Because many of the values are less than 10, the suggested data transformation is \( \sqrt{x+0.5} \).

Used with permission of Kwanchai Gomez.

### Source


### References

None.
Examples

```r
## Not run:

library(agridat)

data(gomez.nonnormal2)
dat <- gomez.nonnormal2

# Gomez suggested sqrt transform
dat <- transform(dat, twhite = sqrt(white+.5))

# QQ plots for raw/transformed data
libs(reshape2, lattice)
qqmath(~ value|variable, data=melt(dat),
       main="gomez.nonnormal2 - raw/transformed QQ plot",
       scales=list(relation="free"))

# Gomez anova table 7.21
m1 <- lm(twhite ~ rep + gen, data=dat)
anova(m1)
```

```r
## Response: twhite2
## Df  Sum Sq Mean Sq  F value Pr(>F)
## rep  2 2.401 1.2004 1.9137  0.1678
## gen 13 48.011 3.6931 5.8877  6.366e-05 ***
## Residuals 26 16.309 0.6273
```

```r
## End(Not run)
```

Description

RCB experiment of rice, 12 varieties with leafhopper survival

Usage

```r
data("gomez.nonnormal3")
```

Format

A data frame with 36 observations on the following 3 variables.

- `gen`: genotype/variety of rice
- `rep`: replicate
- `hoppers`: percentage of surviving leafhoppers
Details

For each rice variety, 75 leafhoppers were caged and the percentage of surviving insects was determined.

Gomez suggest replacing 0 values by $1/(4*75)$ and replacing 100 by $1-1/(4*75)$ where 75 is the number of insects.

In effect, this means, for example, that $(1/4)$th of an insect survived.

Because the data are percents, Gomez suggested using the arcsin transformation.

Used with permission of Kwanchai Gomez.

Source


References

None.

Examples

```r
## Not run:
library(agridat)
data(gomez.nonnormal3)
dat <- gomez.nonnormal3

# First, replace 0, 100 values
dat$hoppers <- dat$hoppers
dat <- transform(dat, hoppers = ifelse(hoppers == 0, 1/(4*75), hoppers))
dat <- transform(dat, hoppers = ifelse(hoppers == 100, 100-1/(4*75), hoppers))
# Arcsin transformation of percentage p converted to degrees
# is arcsin(sqrt(p))/(pi/2)*90
dat <- transform(dat, hoppers = asin(sqrt(hoppers/100))/(pi/2)*90)

# QQ plots for raw/transformed data
libs(reshape2, lattice)
qqmath(~ value|variable, data=melt(dat),
    main="gomez.nonnormal3 - raw/transformed QQ plot",
    scales=list(relation="free"))

m1 <- lm(hoppers ~ gen, data=dat)
anova(m1) # Match Gomez table 7.25
## Response: hoppers
##     Df Sum Sq Mean Sq F value Pr(>F)
## gen 11 16838.7 1530.79  16.502 1.316e-08 ***
## Residuals 24 2226.4  92.77

## End(Not run)
```
gomez.rice.uniformity  Uniformity trial of rice

Description
Uniformity trial of rice in Philippines.

Format
A data frame with 648 observations on the following 3 variables.

row  row
col  column
yield  grain yield, grams/m^2

Details
An area 20 meters by 38 meters was planted to rice variety IR8. At harvest, a 1-meter border was removed around the field and discarded. Each square meter (1 meter by 1 meter) was harvested and weighed.

Field width: 18 plots x 1 m = 18 m
Field length: 38 plots x 1 m = 38 m
Used with permission of Kwanchai Gomez.

Source

Examples
## Not run:

library(agridat)
data(gomez.rice.uniformity)
dat <- gomez.rice.uniformity

libs(desplot)
# Raw data plot
desplot(dat, yield ~ col*row,
       aspect=38/18, # true aspect
       main="gomez.rice.uniformity")

libs(desplot, reshape2)
# 3x3 moving average. Gomez figure 12.1
dmat <- melt(dat, id.var=c('col','row'))
dmat <- acast(dmat, row=col)
gomez.seedrate

RCB experiment of rice, 6 densities

Description

RCB experiment of rice, 6 densities

Format

A data frame with 24 observations on the following 3 variables.

rate kg seeds per hectare
rep rep (block), four levels
yield yield, kg/ha

Details

Rice yield at six different densities in an RCB design.
Used with permission of Kwanchai Gomez.
Split-plot experiment of rice, with subsamples

## Description

Split-plot experiment of rice, with subsamples

## Format

A data frame with 186 observations on the following 5 variables.

- **time**: time factor, T1-T4
- **manage**: management, M1-M6
- **rep**: rep/block, R1-R3
- **sample**: subsample, S1-S2
- **height**: plant height (cm)
Details

A split-plot experiment in three blocks. Whole-plot is 'management', sub-plot is 'time' of application, with two subsamples. The data are the heights, measured on two single-hill sampling units in each plot.

Used with permission of Kwanchai Gomez.

Source


Examples

```r
## Not run:
library(agridat)
data(gomez.splitplot.subsample)
dat <- gomez.splitplot.subsample

libs(HH)
interaction2wt(height ~ rep + time + manage, data=dat,
  x.between=0, y.between=0,
  main="gomez.splitplot.subsample - plant height")

# Management totals, Gomez table 6.8
# tapply(dat$height, dat$manage, sum)

# Gomez table 6.11 analysis of variance
m1 <- aov(height ~ rep + manage + time + manage:time +
  Error(rep/manage/time), data=dat)
summary(m1)
## Error: rep
## Df Sum Sq Mean Sq F value Pr(>F)
## rep 2 2632 1316
## Error: rep:manage
## Df Sum Sq Mean Sq F value Pr(>F)
## manage 7 1482 211.77 2.239 0.0944 .
## Residuals 14 1324 94.59
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Error: rep:manage:time
## Df Sum Sq Mean Sq F value Pr(>F)
## time 3 820.8 273.61 7.945 0.000211 ***
## manage:time 21 475.3 22.63 0.657 0.851793
## Residuals 48 1653.1 34.44
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
## Error: Within
## Df Sum Sq Mean Sq F value Pr(>F)
## Residuals 96 167.4 1.744

## End(Not run)

gomez.splitsplit  
**Split-split-plot experiment of rice**

### Description
Grain yield of three varieties of rice grown in a split-split plot arrangement with 3 reps, nitrogen level as the main plot, management practice as the sub-plot, and rice variety as the sub-sub plot.

### Format
A data frame with 135 observations on the following 7 variables.

- **rep**  block, 3 levels
- **nitro**  nitrogen fertilizer, in kilograms/hectare
- **management**  plot management
- **gen**  genotype/variety of rice
- **yield**  yield
- **col**  column position in the field
- **row**  row position in the field

Used with permission of Kwanchai Gomez.

### Source

### References
https://doi.org/10.1111/jac.12267
Examples

## Not run:

library(agridat)

data(gomez.splitsplit)
dat <- gomez.splitsplit
dat$nf <- factor(dat$nitro)

libs(desplot)
desplot(dat, nf ~ col*row,
    # aspect unknown
    out1=rep, col=management, num=gen, cex=1,
    main="gomez.splitsplit")
desplot(dat, yield ~ col*row,
    # aspect unknown
    out1=rep, main="gomez.splitsplit")

libs(HH)
position(dat$nf) <- c(0,50,80,110,140)
interaction2wt(yield~rep+nf+management+gen, data=dat,
    main="gomez.splitsplit",
    x.between=0, y.between=0,
    relation=list(x="free", y="same"),
    rot=c(90,0), xlab="",
    par.strip.text.input=list(cex=.7))

# AOV. Gomez page 144-153
m0 <- aov(yield~ nf * management * gen + Error(rep/nf/management),
    data=dat)
summary(m0) # Similar to Gomez, p. 153.

## End(Not run)
nitro  Nitrogen fertilizer in kg/ha
  gen  Rice variety
  col  column
  row  row

Details
Note, this is a subset of the 'gomez.stripplot' data.
Used with permission of Kwanchai Gomez.

Source

References

Examples
```r
## Not run:
library(agridat)
data(gomez.stripplot)
dat <- gomez.stripplot

# Gomez figure 3.7
libs(desplot)
desplot(dat, gen ~ col*row,
  # aspect unknown
  out1=rep, out2=nitro, num=nitro, cex=1,
  main="gomez.stripplot")

# Gertheiss figure 1
# library(lattice)
# dotplot(factor(nitro) ~ yield|gen, data=dat)

# Gomez table 3.12
# tapply(dat$yield, dat$rep, sum)
# tapply(dat$yield, dat$gen, sum)
# tapply(dat$yield, dat$nitro, sum)

# Gomez table 3.15. Anova table for strip-plot
dat <- transform(dat, nf=factor(nitro))
m1 <- aov(yield ~ gen * nf + Error(rep + rep:gen + rep:nf), data=dat)
summary(m1)
## Error: rep
##       Df  Sum Sq Mean Sq F value Pr(>F)
```
## Residuals

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gen</td>
<td>5</td>
<td>57100201</td>
<td>11420040</td>
<td>7.653</td>
</tr>
<tr>
<td>Residuals</td>
<td>10</td>
<td>14922619</td>
<td>1492262</td>
<td></td>
</tr>
</tbody>
</table>

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Error: rep:nf

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>nf</td>
<td>2</td>
<td>50676061</td>
<td>25338031</td>
<td>34.07</td>
</tr>
<tr>
<td>Residuals</td>
<td>4</td>
<td>2974908</td>
<td>743727</td>
<td></td>
</tr>
</tbody>
</table>

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Error: Within

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gen:nf</td>
<td>10</td>
<td>23877979</td>
<td>2387798</td>
<td>5.801</td>
</tr>
<tr>
<td>Residuals</td>
<td>20</td>
<td>8232917</td>
<td>411646</td>
<td></td>
</tr>
</tbody>
</table>

# More compact view

```r
## libs(agricolae)
## with(dat, strip.plot(rep, nf, gen, yield))
```

## Analysis of Variance Table

<table>
<thead>
<tr>
<th>Response: yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Df</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>rep</td>
</tr>
<tr>
<td>nf</td>
</tr>
<tr>
<td>Ea</td>
</tr>
<tr>
<td>gen</td>
</tr>
<tr>
<td>Eb</td>
</tr>
<tr>
<td>gen:nf</td>
</tr>
<tr>
<td>Ec</td>
</tr>
</tbody>
</table>

# Mixed-model version

```r
## libs(lme4)
## m3 <- lmer(yield ~ gen * nf + (1|rep) + (1|rep:nf) + (1|rep:gen), data=dat)
## anova(m3)
```

## Analysis of Variance Table

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gen</td>
<td>5</td>
<td>15751300</td>
<td>3150260</td>
</tr>
<tr>
<td>nf</td>
<td>2</td>
<td>28048730</td>
<td>14024365</td>
</tr>
<tr>
<td>gen:nf</td>
<td>10</td>
<td>23877979</td>
<td>2387798</td>
</tr>
</tbody>
</table>

# End(Not run)
Description

A strip-split-plot experiment with three reps, genotype as the horizontal strip, nitrogen fertilizer as the vertical strip, and planting method as the subplot factor.

Format

- yield: grain yield in kg/ha
- planting: planting factor, P1=broadcast, P2=transplanted
- rep: rep, 3 levels
- nitro: nitrogen fertilizer, kg/ha
- gen: genotype, G1 to G6
- col: column
- row: row

Details

Note, this is a superset of the 'gomez.stripplot' data.

Used with permission of Kwanchai Gomez.

Source


Examples

```r
## Not run:
library(agridat)
data(gomez.stripsplitplot)
dat <- gomez.stripsplitplot

# Layout
libs(desplot)
desplot(dat, gen ~ col*row,          
    out1=rep, col=nitro, text=planting, cex=1,  
    main="gomez.stripsplitplot")

# Gomez table 4.19, ANOVA of strip-split-plot design
dat <- transform(dat, nf=factor(nitro))
m1 <- aov(yield ~ nf * gen * planting +      
    Error(rep + rep:nf + rep:gen + rep:nf:gen), data=dat)
```
There is a noticeable linear trend along the y coordinate which may be an artifact that blocking will remove, or may need to be modeled. Note the outside values in the high-nitro boxplot.

```r
interaction2wt(yield ~ nitro + gen + planting + row, dat,
             x.between=0, y.between=0,
             x.relation="free")
```

## End(Not run)

---

**Rice yield in wet & dry seasons with nitrogen fertilizer treatments**

### Description

Rice yield in wet & dry seasons with nitrogen fertilizer treatments

### Format

A data frame with 96 observations on the following 4 variables.

- **season**: season = wet/dry
- **nitrogen**: nitrogen fertilizer kg/ha
- **rep**: replicate
- **yield**: grain yield, t/ha

### Details

Five nitrogen fertilizer treatments were tested in 2 seasons using 3 reps. Used with permission of Kwanchai Gomez.

### Source


### References

Examples

```r
## Not run:
library(agridat)
data(gomez.wetdry)
dat <- gomez.wetdry

library(lattice)
foo1 <- xyplot(yield ~ nitrogen|season, data=dat,
               group=rep,type='l',auto.key=list(columns=3),
               ylab="yield in each season",
               main="gomez.wetdry raw data & model")

# Yang & Juskiw fit a quadratic model with linear and quadratic
# contrasts using non-equal intervals of nitrogen levels.
# This example below omits the tedious contrasts

library(latticeExtra, lme4)
m1 <-lmer(yield ~ season*poly(nitrogen, 2) + (1|season:rep), data=dat)
pdat <- expand.grid(season=c('dry','wet'),
                    nitrogen=seq(from=0,to=150,by=5))
pdat$pred <- predict(m1, newdata=pdat, re.form= ~ 0)
foo1 +
   xyplot(pred ~ nitrogen|season, data=pdat, type='l',lwd=2,col="black")

# m2 <-lmer(yield ~ poly(nitrogen, 2) + (1|season:rep), data=dat)
#anova(m1,m2)
## m2: yield ~ poly(nitrogen, 2) + (1 | season:rep)
## m1: yield ~ season * poly(nitrogen, 2) + (1 | season:rep)
## Df  AIC  BIC logLik deviance Chisq Chi Df Pr(>Chisq)
## m2  5 86.418 93.424 -38.209 76.418
## m1  8 64.216 75.425 -24.108 48.216 28.202 3  3.295e-06 ***
```

### End(Not run)

---

gateway.hessianfly

**Hessian fly damage to wheat varieties**

Description

Hessian fly damage to wheat varieties

Format

- **block**: block factor, 4 levels
- **genotype**: factor, 16 wheat varieties
- **lat**: latitude, numeric
gotway.hessianfly

long longitude, numeric
y number of damaged plants
n number of total plants

Details
The response is binomial.
Each plot was square.

Source
https://doi.org/10.2307/1400401

References

Examples
## Not run:
library(agridat)
data(gotway.hessianfly)
dat <- gotway.hessianfly
dat$prop <- dat$y / dat$n
libs(desplot)
desplot(dat, prop~long*lat,
aspect=1, # true aspect
out1=block, num=gen, cex=.75,
main="gotway.hessianfly")

# spaMM package example
libs(spaMM)
m1 = HLCor(cbind(y, n-y) ~ 1 + gen + (1|block) + Matern(1|long+lat),
data=dat, family=binomial(), ranPars=list(nu=0.5, rho=1/7))
summary(m1)
fixef(m1)
# The following line fails with "Invalid graphics state"
# when trying to use pkgdown::build_site
# filled.mapMM(m1)

# Block random. See Glimmix manual, output 1.18.
# Note: (Different parameterization)

```r
libs(lme4)
l2 <- glmer(cbind(y, n-y) ~ gen + (1|block), data=dat, family=binomial, control=glmerControl(check.nlev.gtr.1="ignore"))
coef(l2)
```

## End(Not run)

---

**goulden.barley.uniformity**

*Uniformity trial of barley*

---

**Description**

Uniformity trial of barley

**Format**

A data frame with 400 observations on the following 3 variables.

- `row` row
- `col` column
- `yield` yield, grams per plot

**Details**

Yield in grams for 400 square-yard barley plots.

- Field width: 20 plots x 3 feet = 60 feet
- Field length: 20 plots x 3 feet = 60 feet

**Source**


**References**

Examples

### Not run:

```r
library(agridat)
data(goulden.barley.uniformity)
dat <- goulden.barley.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
       aspect=20/20, # true aspect
       main="goulden.barley.uniformity")

# Left skewed distribution. See LeClerg, Leonard, Clark
hist(dat$yield, main="goulden.barley.uniformity",
     breaks=c(21,40,59,78,97,116,135,154,173,192,211,230,249)+.5)
```

### End(Not run)

goulden.eggs

Sample of egg weights on 24 consecutive days

Description

Sample of egg weights on 24 consecutive days

Usage

```r
data("goulden.eggs")
```

Format

A data frame with 240 observations on the following 2 variables.

- day
- weight

Details

Data are the weights of 10 eggs taken at random on each day for 24 days. Day 1 was Dec 10, and Day 24 was Jan 2.

The control chart for standard deviations shows 4 values beyond the upper limits. The data reveals a single, unusually large egg on each of these days. These are almost surely double-yolk eggs.

Source

goulden.latin

References

None.

Examples

## Not run:

library(agridat)
data(goulden.eggs)
dat <- goulden.eggs

libs(qicharts)
# Figure 19-4 of Goulden. (Goulden uses 1/n when calculating std dev)
op <- par(mfrow=c(2,1))
qic(weight, x = day, data = dat, chart = 'xbar',
    main = 'goulden.eggs - Xbar chart',
    xlab = 'Date', ylab = 'Avg egg weight' )
qic(weight, x = day, data = dat, chart = 's',
    main = 'goulden.eggs - S chart',
    xlab = 'Date', ylab = 'Std dev egg weight' )
par(op)

## End(Not run)

goulden.latin  

Latin square experiment for testing fungicide

Description

Latin square experiment for testing fungicide

Usage

data("goulden.latin")

Format

A data frame with 25 observations on the following 4 variables.

trt  treatment factor, 5 levels
yield  yield
row  row
col  column
Details

Five treatments were tested to control stem rust in wheat. Treatment codes and descriptions: A = Dusted before rains. B = Dusted after rains. C = Dusted once each week. D = Drifting, once each week. E = Not dusted.

Source


Examples

```r
## Not run:
library(agridat)
data(goulden.latin)
dat <- goulden.latin

desplot(dat, yield ~ col*row,
       text=trt, cex=1, # aspect unknown
       main="goulden.latin")

# Matches Goulden.
m1 <- lm(yield ~ trt + factor(row) + factor(col), data=dat)
anova(m1)

## End(Not run)
```

goulden.splitsplit

Split-split-plot experiment of wheat

Description

Split-split-plot experiment of wheat

Usage

data("goulden.splitsplit")

Format

A data frame with 160 observations on the following 9 variables.

row row
col column
An interesting split-split plot experiment in which the sub-plot treatments have a 2*5 factorial structure.

An experiment was conducted in 1932 on the experimental field of the Dominion Rust Research Laboratory. The study was designed to determine the effect on the incidence of root rot, of variety of wheat, kinds of dust for seed treatment, method of application of the dust, and efficacy of soil inoculation with the root-rot organism.

The field had 4 blocks.

Each block has 2 whole plots for the genotypes.

Each whole-plot had 10 sub-plots for the 5 different kinds of dust and 2 methods of application.

Each sub-plot had 2 sub-sub-plots, one for inoculated soil and the other one for uninoculated soil.

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(goulden.splitsplit)
dat <- goulden.splitsplit
libs(desplot)

## Experiment design. Goulden p. 152-153
## desplot(gen ~ col*row, data=dat,
##         out1=block, out2=trt, text=dust, col=inoc, cex=1,
##         main="goulden.splitsplit")

desplot(dat, yield ~ col*row,
        out1=block, out2=gen,
        col=inoc, num=trt, cex=1,
```
graybill.heteroskedastic

Multi-environment trial of wheat varieties with heteroskedastic yields

Description
Wheat varieties with heteroskedastic yields

Format
A data frame with 52 observations on the following 3 variables.

env  environment, 13 levels
gen  genotype, 4 levels
yield yield

Details
Yield of 4 varieties of wheat at 13 locations in Oklahoma, USA.
The data was used to explore variability between varieties.

Source

References

```r
# Match Goulden table 40
m1 <- aov(yield ~ gen + dust + dry + dust:dry + gen:dust + gen:dry + gen:dust:dry
          + inoc + inoc:gen + inoc:dust + inoc:dry
          + inoc:dust:dry +inoc:gen:dust + inoc:gen:dry
          + Error(block/(gen+gen:dust+gen:inoc:dry)), data=dat)
summary(m1)

## End(Not run)
```
Examples

```r
## Not run:
library(agridat)
data(graybill.heteroskedastic)
dat <- graybill.heteroskedastic

# Genotypes are obviously not homoscedastic
boxplot(yield ~ gen, dat, main="graybill.heteroskedastic")

# Shukla stability variance of each genotype, same as Grubbs' estimate
# Matches Piepho 1994 page 143.
# Do not do this! Nowadays, use mixed models instead.
library("reshape2")
datm <- acast(dat, gen~env)
w <- datm
w <- sweep(w, 1, rowMeans(datm))
w <- sweep(w, 2, colMeans(datm))
w <- w + mean(datm)
w <- rowSums(w^2)
k=4; n=13
sig2 <- k*w/((k-2)*(n-1)) - sum(w)/((k-1)*(k-2)*(n-1))

var.shukla <- function(x,N){
  # Estimate variance of shukla stability statistics
  # Piepho 1994 equation (5)
  K <- length(x) # num genotypes
  S <- outer(x,x)
  S1 <- diag(S)
  S2 <- rowSums(S) - S1
  S[!upper.tri(S)] <- 0 # Make S upper triangular
  # The ith element of S3 is the sum of the upper triangular elements of S,
  # excluding the ith row and ith column
  S3 <- sum(S) - rowSums(S) - colSums(S)
  var.si2 <- 2*S1/(N-1) + 4/( (N-1)*(K-1)^2 ) * ( S2 + S3/(K-2)^2 )
  return(var.si2)
}

# Set negative estimates to zero
sig2[sig2<0] <- 0

# Variance of shukla stat. Match Piepho 1994, table 5, example 1
var.shukla(sig2,13)

## G1  G2  G3  G4
## 145.98 14.14 75.15 18.25

# End(Not run)
```
Description
Factorial experiment of cotton in Sudan.

Usage
data("gregory.cotton")

Format
A data frame with 144 observations on the following 6 variables.

- yield: a numeric vector
- year: year
- nitrogen: nitrogen level
- date: sowing date
- water: irrigation amount
- spacing: spacing between plants

Details
Experiment conducted in Sudan at the Gezira Research Farm in 1929-1930 and 1930-1931. The effects on yield of four factors was studied in all possible combinations.

- Sowing dates in 1929: D1 = Jul 24, D2 = Aug 11, D3 = Sep 2, D4 = Sep 25.
- Spacing: S1 = 25 cm between holes, S2 = 50 cm, S3 = 75 cm. The usual spacing is 50-70 cm.
- Irrigation: I1 = Light, I2 = Medium, I3 = Heavy.
- Nitrogen: N0 = None/Control, N1 = 600 rotls/feddan.

In each year there were 4*3*2*2=72 treatments, each replicated four times. The means are given here.

Gregory (1932) has two interesting graphics: 1. radial bar plot 2. photographs of 3D model of treatment means.

Source
Gregory, FG and Crowther, F and Lambert, AR (1932). The interrelation of factors controlling the production of cotton under irrigation in the Sudan. The Journal of Agricultural Science, 22, 617-638. Table 1, 10. https://doi.org/10.1017/S0021859600054137

References
Paterson, D. Statistical Technique in Agricultural Research, p. 211.
## Not run:

library(agridat)
data(gregory.cotton)
dat <- gregory.cotton

# Main effect means, Gregory table 2
## libs(dplyr)
## dat
dat
dat
dat
dat

Figure 2 of Gregory. Not recommended, but an interesting exercise.

https://stackoverflow.com/questions/13887365

if(FALSE){
    libs(ggplot2)
    d1 <- subset(dat, year=="Y1")
    d1 <- transform(d1, grp=factor(paste(date,nitrogen,water,spacing)))
    d1 <- d1[order(d1$grp),] # for angles
    # Rotate labels on the left half 180 deg. First 18, last 18 labels
    d1$ang <- 90+seq(from=(360/nrow(d1))/1.5, to=(1.5*(360/nrow(d1)))-360,
                     length.out=nrow(d1)+80
    # Lables on left half to right-adjusted
    d1$hjust <- 0
    d1$hjust[1:18] <- d1$hjust[55:72] <- 1

    gg <- ggplot(d1, aes(x=grp, y=yield, fill=factor(spacing))) +
        geom_col() +
        guides(fill=FALSE) + # no legend for 'spacing'
        coord_polar(start=-pi/2) + # default is to start at top
        labs(title="gregory.cotton 1929", x="", y="", label="") +
        # The bar columns are centered on 1:72, subtract 0.5 to add radial axes
        geom_vline(xintercept = seq(1, 72, by=3)-0.5, color="gray", size=.25) +
        geom_vline(xintercept = seq(1, 72, by=18)-0.5, size=1) +
        geom_vline(xintercept = seq(1, 72, by=9)-0.5, size=.5) +
        geom_hline(yintercept=c(1,2,3)) +
        geom_text(data=d1, aes(x=grp, y=max(yield), label=grp, angle=ang, hjust=hjust),
                  size=2) +
        theme(panel.background=element_blank(),
              axis.title=element_blank(),
              panel.grid=element_blank(),
              axis.text.x=element_blank(),
              axis.text.y=element_blank(),
              axis.ticks=element_blank() )
    print(gg)
}
## End(Not run)

```r
grover.diallel

### Diallel 6x6

#### Description

Diallel 6x6 in 4 blocks.

#### Usage

```r
data("grover.diallel")
```

#### Format

A data frame with 144 observations on the following 5 variables.

- `yield` a numeric vector
- `rep` a character vector
- `parent1` a character vector
- `parent2` a character vector
- `cross` a character vector

#### Details

Yield for a 6x6 diallel with 4 reps.

Note: The mean for the 2x2 cross is slightly different than Grover p. 252. There appears to be an unknown error in the one of the 4 reps in the data on page 250.

#### Source


#### References

None
Examples

```R
## Not run:
data(grover.diallel)
dat <- grover.diallel

anova(aov(yield ~ rep + cross, data=dat))

# These effects match the GCA and SCA values in Grover table 3, page 253.

library(lmDiallel)
m2 <- lm.diallel(yield ~ parent1 + parent2, Block=rep,
data=dat, fct="GRIFFING1")

library(multcomp)
summary(glht(linfct=diallel.eff(m2), test=adjusted(type="none")))

## Linear Hypotheses:
## Estimate Std. Error t value Pr(>|t|)
## Intercept == 0  93.0774  0.9050 102.851  <0.01 ***
## g_P1 == 0   -0.9911  1.4309  -0.693 1.0000
## g_P2 == 0   -2.2631  1.4309  -1.582 0.9748
## g_P3 == 0    5.4247  1.4309   3.791 0.0302 *
## g_P4 == 0   -4.2490  1.4309  -2.969 0.1972
## g_P5 == 0   -3.9328  1.4309  -2.748 0.3008
## ts_P1:P1 == 0 -10.4026  4.5249  -2.299 0.6014
## ts_P1:P2 == 0  -9.7214  3.2629  -2.979 0.1928
## ts_P1:P3 == 0  -0.4581  3.2629  -0.140 1.0000
## ts_P1:P4 == 0  17.0428  3.2629   5.223  <0.01 ***
## ts_P1:P5 == 0  25.4765  3.2629   7.808  <0.01 ***
## ts_P1:P6 == 0 -21.9372  3.2629  -6.723  <0.01 ***
## ts_P2:P1 == 0  -9.7214  3.2629  -2.979  0.1928
## ts_P2:P2 == 0   7.0899  4.5249   1.567  0.9773
```

### End(Not run)

---

**grover.rcb.subsample**  
**Rice RCB with subsamples**

**Description**

An experiment on rice with 9 fertilizer treatments in 4 blocks, 4 hills per plot.

**Usage**

```R
data("grover.rcb.subsample")
```

**Format**

A data frame with 144 observations on the following 4 variables.
tiller  number of tillers  
trt   treatment factor  
block block factor  
unit subsample unit  

Details  
An experiment on rice with 9 fertilizer treatments in 4 blocks, 4 hills per plot. The response variable is tiller count (per hill). The hills are sampling units.  

Source  

References  
None.  

Examples  
```r  
## Not run:  
data(grover.rcb.subsample)  
# Fixed-effects ANOVA. Matches Grover page 86.  
anova(aov(tiller ~ block + trt + block:trt, data=grover.rcb.subsample))  
## Response: tiller  
## Df Sum Sq Mean Sq F value Pr(>F)  
## block 3 930 310.01 3.6918 0.01415 *  
## trt 8 11816 1477.00 17.5891 < 2e-16 ***  
## block:trt 24 4721 196.71 2.3425 0.00158 **  
## Residuals 108 9069 83.97  
## End(Not run)  
```

---

**gumpertz.pepper**  
*Phytophtera disease incidence in a pepper field*

Description  
Phytophtera disease incidence in a pepper field  

Format  
A data frame with 800 observations on the following 6 variables.  

field  field factor, 2 levels  
row  x ordinate  
quadrat  y ordinate  

disease presence (Y) or absence (N) of disease
water soil moisture percent
leaf leaf assay count

Details

Each field is 20 rows by 20 quadrates, with 2 to 3 bell pepper plants per plot. If any plant was wilted, dead, or had lesions, the Phytophthora disease was considered to be present in the plot. The soil pathogen load was assayed as the number of leaf disks colonized out of five. In field 2, the pattern of disease presence appears to follow soil water content. In field 1, no obvious trends were present.

Gumpertz et al. model the presence of disease using soil moisture and leaf assay as covariates, and using disease presence of neighboring plots as covariates in an autologistic model.

Used with permission of Marcia Gumpertz. Research funded by USDA.

Source


Examples

```r
## Not run:

library(agridat)
data(gumpertz.pepper)
dat <- gumpertz.pepper

# Gumpertz deletes two outliers
dat[ dat$field == "F1" & dat$row==20 & dat$quadrat==10, 'water' ] <- NA
dat[ dat$field == "F2" & dat$row==5 & dat$quadrat==4, 'water' ] <- NA

# Horizontal flip
dat <- transform(dat, row=21-row)

# Disease presence. Gumpertz fig 1a, 2a.
lib('desplot')
grays <- colorRampPalette(c("#d9d9d9","#252525"))
desplot(dat, disease ~ row*quadrat|field,
    col.regions=c('white','black'), aspect=1, # uncertain aspect
    main="gumpertz.pepper disease presence", )

# Soil water. Gumpertz fig 1b, 2b
desplot(dat, water ~ row*quadrat|field,
    col.regions=grays(5), aspect=1, # uncertain aspect
    at=c(5,7.5,10,12.5,15,18),
    main="gumpertz.pepper soil moisture")

# Leaf assay. Gumpertz fig 1c, 2c
```
hadach.lettuce

Lettuce resistance to downy mildew resistance (with marker data)

Description

Lettuce resistance to downy mildew resistance (with marker data).

Usage

data("hadach.lettuce")

Format

A data frame with 703 observations on the following 4 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>loc</td>
<td>locations</td>
</tr>
<tr>
<td>gen</td>
<td>genotype</td>
</tr>
<tr>
<td>rep</td>
<td>replicate</td>
</tr>
<tr>
<td>dmr</td>
<td>downy mildew resistance</td>
</tr>
</tbody>
</table>
Details

A biparental cross of 95 recombinant inbred lines of "Salinas 88" (susceptible) and "La Brillante" (highly resistant to downy mildew). The 89 RILs were evaluated in field experiments performed in 2010 and 2011 near Salinas, California. Each loc had a 2 or 3 rep RCB design. There were approximately 30 plants per plot. Plots were scored 0 (no disease) to 5 (severe disease).

The authors used the following model in a first-stage analysis to compute adjusted means for each genotype:

\[ y = \text{loc} + \text{gen} + \text{gen:loc} + \text{block:loc} + \text{error} \]

where gen was fixed and all other terms random. The adjusted means were used as the response in a second stage:

\[ mn = 1 + Zu + \text{error} \]

where \( Z \) is the design matrix of marker effects. The error term is fixed to have covariance matrix \( R \) be the same as from the first stage.

Genotyping was performed with 95 SNPs and 205 amplified fragment length polymorphism markers so that a marker matrix \( M \) (89×300) was provided. The biallelic marker \( M(iw) \) for the \( i \)th genotype and the \( w \)th marker with alleles A1 (i.e. the reference allele) and A2 was coded as 1 for A1,A1, -1 for A2,A2 and 0 for A1,A2 and A2,A2.

The electronic version of the lettuce data are licensed CC-BY 4 and were downloaded 20 Feb 2021. https://figshare.com/articles/dataset/Lettuce_trial_phenotypic_and_marker_data_/8299493

Source


References


Examples

```r
## Not run:
library(agridat)
data(hadasch.lettuce)
dat <- hadasch.lettuce
datm <- hadasch.lettuce.markers
libs(agridat)
# loc 1 has 2 reps, loc 3 has higher dmr
dotplot(dmr ~ factor(gen)|factor(loc), dat, 
group=rep, layout=c(1,3), 
main="hadasch.lettuce")

# kinship matrix
```
tcrossprod(as.matrix(datm[, -1]))

libs(asreml)
dat <- transform(dat, loc=factor(loc), gen=factor(gen), rep=factor(rep))
m1 <- asreml(dmr ~ 1 + gen, data=dat,
               random = ~ loc + gen:loc + rep
p1 <- predict(m1, classify="gen")$pvals

libs(sommer)
m2 <- mmer(dmr ~ 0 + gen, data=dat,
               random = ~ loc + gen:loc + rep
p2 <- coef(m2)
head(p1)
head(p2)

## End(Not run)

---

**hanks.sprinkler**

Wheat yields in a line-source sprinkler experiment

### Description

Three wheat varieties planted in 3 blocks, with a line sprinkler crossing all whole plots.

### Format

A data frame with 108 observations on the following 7 variables.

- **block** block
- **row** row
- **subplot** column
- **gen** genotype, 3 levels
- **yield** yield (tons/ha)
- **irr** irrigation level, 1..6
- **dir** direction from sprinkler, N/S

### Details

A line-source sprinkler is placed through the middle of the experiment (between subplots 6 and 7). Subplots closest to the sprinkler receive the most irrigation. Subplots far from the sprinkler (near the edges) have the lowest yields.

One data value was modified from the original (following the example of other authors).

### Source

References


Examples

```r
## Not run:

library(agridat)
data(hanks.sprinkler)
dat <- hanks.sprinkler

# The line sprinkler is vertical between subplots 6 & 7
libs(desplot)
desplot(dat, yield~subplot*row,
   out1=block, out2=irr, cex=1, # aspect unknown
   num=gen, main="hanks.sprinkler")

libs(lattice)
xyplot(yield~subplot|block, dat, type=c("l", "s"), group=gen,
   layout=c(1,3), auto.key=TRUE,
   main="hanks.sprinkler",
   panel=function(x,y,...){
     panel.xyplot(x,y,...)
     panel.abline(v=6.5, col='wheat')
   })

# This is the model from the SAS documentation
## proc mixed;
## class block gen dir irr;
## model yield = gen|dir|irr@2;
## random block block|dir block|irr;
## repeated / type=toep(4) sub=block|gen r;

# asreml 3
libs(asreml,lucid)
if( utils::packageVersion("asreml") < "4") {
  # asreml3
  dat <- transform(dat, subf=factor(subplot),
       irrf=factor(irr))
  ```
dat <- dat[order(dat$block, dat$gen, dat$subplot),]

m1 <- asreml(yield ~ gen + dir + irrf + gen:dir + gen:irrf + dir:irrf, 
data=dat, 
  random= ~ block + block:dir + block:irrf, 
rcoV= ~ block:gen:corb(subf, k=4))

libs(lucid)
vc(m1)
## effect component std.error z.ratio constr
## block!block.var 0.2194 0.2393 0.92 pos
## block:dir!block.var 0.01768 0.03154 0.56 pos
## block:irrf!block.var 0.03539 0.03617 0.98 pos
## R!variance 0.285 0.05086 5.6 pos
## R!cor1 0.02802 0.1143 0.25 uncon
## R!cor2 0.005095 0.1278 0.04 uncon
## R!cor3 -0.3246 0.0905 -3.6 uncon

## # convert asreml correlations to SAS covariances
## round(.2850 * c(1, .02802, .005095, -.3246),4) # res var * (cor1, cor2, cor3)
## [1] 0.2850 0.0080 0.0015 -0.0925

# asreml 4
libs(asreml,lucid)
if( utils::packageVersion("asreml") > "4") {

dat <- transform(dat, subf=factor(subplot), 
  irrf=factor(irr))
dat <- dat[order(dat$block, dat$gen, dat$subplot),]

# In asreml3, we can specify corb(subf, 3)
# In asreml4, only corb(subf, 1) runs. corb(subf, 3) says:
# Correlation structure is not positive definite
m1 <- asreml(yield ~ gen + dir + irrf + gen:dir + gen:irrf + dir:irrf, 
data=dat, 
  random= ~ block + block:dir + block:irrf, 
  resid = ~ block:gen:corb(subf, 2))
lucid::vc(m1)
## effect component std.error z.ratio bound
## block 0.194 0.2231 0.87 P 0.5
## block:dir 0.02729 0.04959 0.55 P 0
## block:irrf 0.02275 0.0347 0.66 P 0.1
## block:gen:subf!R 0.3234 0.05921 5.5 P 0
## block:gen:subf!subf!cor1 0.169 0.09906 1.7 P 0.1

}

## End(Not run)
Mating crosses of white pine trees

Usage

data("hanover.whitepine")

Format

A data frame with 112 observations on the following 4 variables.

- rep: replicate
- female: female parent
- male: male parent
- length: epicotyl length, cm

Details

Four male (pollen parent) White Pine trees were mated to seven female trees and 2654 progeny were grown in four replications, one plot per mating in each replication. Parent trees were sourced from Idaho, USA. The data are plot means of epicotyl length.

Becker (1984) used these data to demonstrate the calculation of heritability.

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(hanover.whitepine)
dat <- hanover.whitepine

# Relatively high male-female interaction in growth comared
```
# to additive gene action. Response is more consistent within
# male progeny than female progeny.
# with(dat, interaction.plot(female, male, length))
# with(dat, interaction.plot(male, female, length))
bwplot(length ~ male|female, data=dat,  
main="hanover.whitepine - length for male:female crosses",  
xlab="Male parent", ylab="Epicotyl length")

# Progeny sums match Becker p 83
sum(dat$length) # 380.58
aggregate(length ~ female + male, data=dat, FUN=sum)

# Sum of squares matches Becker p 85
m1 <- aov(length ~ rep + male + female + male:female, data=dat)
anova(m1)

# Variance components match Becker p. 85
libs(lme4)
libs(lucid)
m2 <- lmer(length ~ (1|rep) + (1|male) + (1|female) + (1|male:female), data=dat)
#as.data.frame(lme4::VarCorr(m2))
vc(m2)
## grp      var1   var2   vcov   sdcor
## male:female (Intercept) <NA> 0.1369 0.3699
## female (Intercept) <NA>  0.02094 0.1447
## male (Intercept) <NA>  0.1204 0.3469
## rep (Intercept) <NA>  0.01453 0.1205
## Residual     <NA>  0.2004 0.4477

# Becker used this value for variability between individuals, within plot
s2w <- 1.109

# Calculating heritability for individual trees
s2m <- .120
s2f <- .0209
s2mf <- .137
vp <- s2m + s2f + s2mf + s2w  # variability of phenotypes = 1.3869
4*s2m / vp  # heritability male 0.346
4*s2f / vp  # heritability female 0.06
2*(s2m+s2f)/vp  # heritability male+female 0.203
# As shown in the boxplot, heritability is stronger through the
# males than through the females.

## End(Not run)

harris.multi.uniformity

Uniformity trials with multiple crops, 15 years on the same land
**Description**

Uniformity trials with multiple crops, at Huntley Field Station, Montana, 1911-1925.

**Format**

A data frame with 1058 observations on the following 5 variables.

- `series` series (field coordinate)
- `plot` plot number (field coordinate)
- `year` year, 1911-1925
- `crop` crop factor
- `yield` yield, numeric

**Details**

Field width: 2 plots * 317 ft + 5 feet alley = 639 feet
Field length: 23 plots * 23.3 feet = 536 feet

All yields are given in pound per plot, except sugar beets, which are given in tons per acre. Harris (1928) shows a map of the location.

Harris 1920: In the spring of 1911 this field was laid out into 46 plots, each measuring 23.5 by 317 feet and containing 0.17 acre, arranged in two parallel series of 23 plots each. The two series of plots were separated merely by a temporary irrigation ditch. In 1911 it was planted to sugar beets, and in the spring of 1912 it was seeded to alfalfa, and one cutting was harvested that year. This stand remained on the ground during 1913 and 1914, when the entire field was fall-plowed. In 1913 three cuttings were made, but the third cutting was lost in a heavy wind which scattered and mixed the crop before weighings from the various plots could be made. The first cutting, designated as alfalfa I, was made on plots one-half the original size. The second cutting was harvested from plots one-quarter the original size. The first and second cuttings in 1914 were weighed for plots one-quarter the original size—that is, 0.0425-acre plots—while the third cutting was recorded for plots one-third the original size. These furnish the data for alfalfa I, II, and III for 1914. Total yields for the first and second cuttings in 1913 and 1914 and for the first, second, and third cuttings in 1914 are also considered. In 1915 and 1916 ear corn was grown. In 1917 the fields were planted to oats, and records were made of grain, straw, and total yield. In 1918 silage corn was grown. In 1919 the land produced a crop of barley.

Harris 1928: The southeast corner of Series II, the east series, is about 80 feet from the main canal, and the southwest corner of Series III is about 50 feet from Ouster Coulee. The main project canal carries normally during the irrigation season about 400 second-feet of water. The water surface in the canal is about 4 feet above the high corner of the field. It is evident from surface conditions, as well as from borings made between the canal and the field, that there is extensive seepage from the canal into the subsoil of the field. The volume of this seepage has been larger in recent years than it was in the earlier years of the cropping experiments, probably because the canal bank has been worn away by internal erosion, exposing a stratum of sandy subsoil that underlies the canal and part of the field.

Whereas in the earlier crops Series II was better for alfalfa, Series III was better for alfalfa in the later period. The writers feel inclined to suggest that in the earlier experiments the height of the water table had no harmful effect upon a deep-rooted crop such as alfalfa. It is quite possible that
during drier periods the higher water table actually favored alfalfa growth on Series II. The higher water tables of recent years have probably had a deleterious influence, which has been especially marked on Series II, where the water apparently comes nearer to the surface than in Series III.

Source


Examples

```r
# Not run:
library(agridat)

data(harris.multi.uniformity)
dat <- harris.multi.uniformity
# Combine year/crop into 'harvest'
dat <- transform(dat, harv = factor(paste0(year, ".", crop)))
# Convert 1911 from tons to pounds
dat$yield[dat$year==1911] <- 340 * dat$yield[dat$year==1911]

# Average yields. Harris 1928, table 2.
aggregate(yield~harv, dat, mean)

# Corrgram
libs(reshape2,corrgram)
mat <- acast(dat, series+plot~harv, value.var="yield")
corrgram(mat, main="harris.multi.uniformity - correlation of crop yields")

# Compare to Harris 1928, table 4. More positive than negative correlations.
# densityplot(as.vector(cor(mat)), xlab="correlations",
# main="harris.multi.uniformity")

# Standardize yields for each year
mats <- scale(mat)
# Melt and re-name columns so we can make field maps. Obvious spatial # patterns that persist over years
d2 <- melt(mats)
names(d2) <- c('ord','harv','yield')
d2$series <- as.numeric(substring(d2$ord,1,1))
d2$plot <- as.numeric(substring(d2$ord,3))

# Series 2 is on the east side, so switch 2 and 3 for correct plotting
d2$xord <- 5 - dat$series
# Note that for alfalfa, higher-yielding plots in 1912-1914 were
# lower-yielding in 1922-1923.
libs(desplot)
desplot(d2, yield ~ xord*plot|harv,
```
# Crude fertility map by averaging across years shows probable sub-surface water effects

```r
agg <- aggregate(yield ~ xord + plot, data=d2, mean)
desplot(agg, yield ~ xord + plot,
   aspect=536/639, # true aspect
   main="harris.multi.uniformity fertility")
```

## End(Not run)

---

**harris.wateruse**  
*Water use by horticultural trees*

### Description

Water use by horticultural trees

### Format

A data frame with 1040 observations on the following 6 variables.

- **species**: species factor, 2 levels
- **age**: age factor, 2 levels
- **tree**: tree factor, 40 (non-consecutive) levels
- **day**: day, numeric
- **water**: water use, numeric

### Details

Ten trees in each of four groups (two species, by two ages) were assessed for water usage, approximately every five days.

Missing values are included for the benefit of asreml, which needs a 'balanced' data set due to the kronecker-like syntax of the R matrix.

Used with permission of Roger Harris at Virginia Polytechnic.

### Source

## Examples

```r
## Not run:
library(agridat)
data(harris.wateruse)
dat <- harris.wateruse

# Compare to Schabenberger & Pierce, fig 7.23
libs(latticeExtra)
useOuterStrips(xyplot(water ~ day|species*age, dat, as.table=TRUE,
group=tree, type=c('p','smooth'),
main="harris.wateruse 2 species, 2 ages (10 trees each)")

# Note that measurements on day 268 are all below the trend line and
# thus considered outliers. Delete them.
dat <- subset(dat, day!=268)

# Schabenberger figure 7.24
xyplot(water ~ day|tree, dat, subset=age=="A2" & species=="S2",
as.table=TRUE, type=c('p','smooth'),
ylab="Water use profiles of individual trees",
main="harris.wateruse (Age 2, Species 2)")

# Rescale day for nicer output, and convergence issues, add quadratic term
dat <- transform(dat, ti=day/100)
dat <- transform(dat, ti2=ti*ti)

d22 <- droplevels(subset(dat, age=="A2" & species=="S2")

# ----- Model 1, for subgroup A2,S2

# First, a fixed quadratic that is common to all trees, plus
# a random quadratic deviation for each tree.

## Schabenberger, Output 7.26
## proc mixed;
## class tree;
## model water = ti ti*ti / s;
## random intercept ti ti*ti/subject=tree;

libs(nlme, lucid)

## We use pdDiag() to get uncorrelated random effects
m1n <- lme(water ~ 1 + ti + ti2, data=d22, na.action=na.omit,
random = list(tree=pdDiag(~1+ti+ti2)))

# vc(m1n)
## effect variance  stddev
## (Intercept)  0.2691  0.5188
##    ti    0     0.0000144
##    ti2   0     0.0000039
```
if(0)
  # Various other models with lme4 & asreml

libs(lme4, lucid)

m1l <- lmer(water ~ 1 + ti + ti2 + (1|tree) +
            (0+ti|tree) + (0+ti2|tree), data=d22)
  # vc(m1l)
  ## grp  var1  var2  vcov  sdcor
  ## tree (Intercept) <NA> 0.2691 0.5188
  ## tree.1   ti <NA>   0   0
  ## tree.2  ti2 <NA>   0   0
  ## Residual <NA> <NA> 0.1472 0.3837

# Once the overall quadratic trend has been removed, there is not
# too much evidence for consecutive observations being correlated
## d22r <- subset(d22, !is.na(water))
## d22r$res <- resid(m1n)
## xypplot(res ~ day|tree,d22r,
## as.table=TRUE, type=c('p','smooth'),
## ylab="residual",
## main="harris.wateruse - Residuals of individual trees")
## op <- par(mfrow=c(4,3))
## tapply(d22r$res, d22r$tree, acf)
## par(op)

# ----- Model 2, add correlation of consecutive measurements

## Schabenberger (page 516) adds correlation.
## Note how the fixed quadratic model is on the "ti = day/100" scale
## and the correlated observations are on the "day" scale. The
## only impact this has on the fitted model is to increase the
## correlation parameter by a factor of 100, which was likely
## done to get better convergence.

## proc mixed data=age2sp2;
##   class tree;
##   model water = ti ti*ti / s ;
##   random intercept /subject=tree s;
##   repeated /subject=tree type=sp(exp)(day);

## Same as SAS, use ti for quadratic, day for correlation
m2l <- lme(water ~ 1 + ti + ti2, data=d22,
          random =~-1|tree,
          cor = corExp(form=~ day|tree),
          na.action=na.omit)

m2l # Match output 7.27. Same fixef, ranef, variances, exp corr
  # vc(m2l)
  ## effect variance stddev
  ## (Intercept)    0.2656 0.5154
## Residual 0.1541 0.3926

# ---

## Now use asreml. When I tried rcov=-tree:exp(ti),
## the estimated parameter value was on the 'boundary', i.e. 0.
## Changing rcov to the 'day' scale produced a sensible estimate
## that matched SAS.
## Note: SAS and asreml use different parameterizations for the correlation
## SAS uses exp(-d/phi) and asreml uses phi^d.
## SAS reports 3.79, asreml reports 0.77, and exp(-1/3.7945) = 0.7683274
## Note: normally a quadratic would be included as 'pol(day,2)'

libs(asreml)
d22 <- d22[order(d22$tree, d22$day),]
m2a <- asreml(water ~ 1 + ti + ti2,
data=d22,
random = ~ tree,
rcov=~tree:exp(day))

# code below is for asremlr 4
# m2a <- asreml(water ~ 1 + ti + ti2,
#     data=d22,
#     random = ~ tree,
#     resid= -tree:exp(day))

vc(m2a)
#
## effect component std.error z.ratio constr
## tree!tree.var 0.2656 0.1301 2 pos
## R!variance 0.1541 0.01611 9.6 pos
## R!day.pow 0.7683 0.04191 18 uncon

# ----- Model 3. Full model for all species/ages. Schabenberger p. 518

## /* Continuous AR(1) autocorrelations included */
## proc mixed data=wateruse;
## class age species tree;
## model water = age*species age*species*ti age*species*ti*ti / noint s;
## random intercept ti / subject=age*species*tree s;
## repeated / subject=age*species*tree type=sp(exp)(day);

m3l <- lme(water ~ 0 + age:species + age:species:ti + age:species:ti2,
data=dat, na.action=na.omit,
random = list(tree=pdDiag(~1+ti)),
    cor = corExp(form=-day|tree) )
m3l # Match Schabenberger output 7.27. Same fixef, ranef, variances, exp corr

# vc(m3l)
## effect variance stddev
## (Intercept) 0.1549 0.3936
## ti 0.02785 0.1669
## Residual 0.16 0.4

# --- asreml

dat <- dat[order(dat$tree,dat$day),]
m3a <- asreml(water ~ 0 + age:species + age:species:ti + age:species:ti2,
data=dat,
random = ~ age:species:tree + age:species:tree:ti,
rcov = ~ tree:exp(day) )

# code for asremlr 4
## m3a <- asreml(water ~ 0 + age:species + age:species:ti + age:species:ti2,
data=dat,
random = ~ age:species:tree + age:species:tree:ti,
resid = ~ tree:exp(day) )

# vc(m3a) # Note: day.pow = .8091 = exp(-1/4.7217)
##  effect component std.error z.ratio constr
## age:species:tree!age.var 0.1549 0.07192 2.2 pos
## age:species:tree:ti!age.var 0.02785 0.01343 2.1 pos
## R!variance 0.16 0.008917 18 pos
## R!day.pow 0.8091 0.01581 51 uncon

} 

## End(Not run)

### Description

Ranges of analytes in soybean from other authors

### Format

A data frame with 80 observations on the following 5 variables.

source  Source document
substance Analyte substance
min  minimum amount (numeric)
max  maximum analyte amount (numeric)
number  number of substances
Details

Harrison et al. show how to construct an informative Bayesian prior from previously-published ranges of concentration for several analytes.

The units for daidzein, genistein, and glycitein are micrograms per gram.

The raffinose and stachyose units were converted to a common ‘percent’ scale.

The author names in the ‘source’ variable are shortened forms of the citations in the supplemental information of Harrison et al.

Source


Data retrieved from the Supplemental Information of this source.

References


Examples

```r
## Not run:
library(agridat)
data(harrison.priors)
dat <- harrison.priors
d1 <- subset(dat, substance=="daidzein")

# Stack the data to 'tall' format and calculate empirical cdf
d1t <- with(d1, data.frame(xx = c(min, max), yy=c(1/(number+1), number/(number+1))))

# Harrison 2012 Example 4: Common prior distribution
# Harrison uses the minimum and maximum levels of daidzein from previous
# studies as the first and last order statistics of a lognormal
# distribution, and finds the best-fit lognormal distribution.

m0 <- mean(log(d1t$xx)) # 6.37
s0 <- sd(log(d1t$xx)) # .833
mod <- nls(yy ~ plnorm(xx, meanlog, sdlog), data=d1t,
          start=list(meanlog=m0, sdlog=s0))
coef(mod) # Matches Harrison 2012
## meanlog      sdlog
## 6.4187829  0.6081558

plot(yy~xx, data=d1t, xlab="", ylab="", xaxt="n", yaxt="n",
     xaxs="i", yaxs="i",
     ylim=c(0,1),
     xlim=c(0,2000),
     xlab="", ylab="")
```
```r
mlog <- coef(mod)[1] # 6.4
slog <- coef(mod)[2] # .61
xvals <- seq(0, 2000, length=100)
lines(xvals, plnorm(xvals, meanlog=mlog, sdlog=slog))

d1a <- d1

d1a$source <- as.character(d1a$source)
d1a[19,'source'] <- "(All)" # Add a blank row for the densitystrip
d1

libs(latticeExtra)
# Plot the range for each source, a density curve (with arbitary
# vertical scale) for the common prior distribution, and a density
# strip by stacking the individual bands and using transparency
segplot(factor(source) ~ min+max, d1a,
  main="harrison.priors",xlab="daidzein level",ylab="source") +
  xyplot(5000*dlnorm(xvals, mlog, slog)~xvals, type='l') +
  segplot(factor(rep(1,18)) ~ min+max, d1, 4, level=d1$number,
  col.regions="gray20", alpha=.1)

## End(Not run)
```

---

**hartman.tomato.uniformity**

*Uniformity trial of tomato*

---

**Description**

Uniformity trial of tomato in Indiana

**Usage**

```r
data("hartman.tomato.uniformity")
```

**Format**

A data frame with 384 observations on the following 3 variables.

- row row
- col column
- yield yield, pounds per plot
Details

Grown in Indiana in 1941. Plants were spaced 3 feet apart in rows 6 feet apart, 330 feet long. Each row was divided into 3 sections of 34 plants separated by strips 12 feet long to provide roadways for vehicles.

Each row was divided into 4-plant plots, with 8 plots in each section of row and with one plant left as a guard at the end of each section.

There were 49 plants missing out of 3072 total plants, but these have been ignored.

Note, the data given in Table 1 of Hartman are for 8-plant plots!

Field width: 3 sections (34 plants * 3 feet) + 2 roads * 12 feet = 330 feet.

Field length: 32 rows * 6 feet = 192 feet

As oriented on the page, plots were, on average, 330/12=27.5. feet wide, 6 feet tall.

Discussion notes from Hartman.

Total yield is 26001 pounds. Hartman says the yield of the field was 10.24 tons per acre, which we can verify:

\[
\frac{26001 \text{ lb/field} \times \left(\frac{1}{384} \text{ field/plot}\right) \times \left(\frac{1}{(24*6) \text{ plot/ft}^2}\right) \times \left(\frac{43560 \text{ ft}^2/\text{acre}}{}\right) \times \left(\frac{1}{2000 \text{ tons/lb}}\right)}{} = 10.24 \text{ tons/acre}
\]

The rows on the top/bottom (north/south) were intended as guard rows, and had yields similar to the other rows, suggesting that competition between rows did not exist. For comparing varieties, 96*6 foot plots work well.

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(hartman.tomato.uniformity)
libs(desplot)
desplot(hartman.tomato.uniformity, yield ~ col*row,
       flip=TRUE, aspect=192/330, # true aspect
       main="hartman.tomato.uniformity")

## End(Not run)
```
**Description**

Average daily gain of 65 steers for 3 lines, 9 sires.

**Usage**

data("harvey.lsmeans")

**Format**

A data frame with 65 observations on the following 7 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>line</td>
<td>line of the dam</td>
</tr>
<tr>
<td>sire</td>
<td>sire</td>
</tr>
<tr>
<td>damage</td>
<td>age class of the dam</td>
</tr>
<tr>
<td>calf</td>
<td>calf number</td>
</tr>
<tr>
<td>weanage</td>
<td>calf age at weaning</td>
</tr>
<tr>
<td>weight</td>
<td>calf weight at start of feeding</td>
</tr>
<tr>
<td>adg</td>
<td>average daily gain</td>
</tr>
</tbody>
</table>

**Details**

The average daily gain 'adg' for each of 65 Hereford steers.
The calf age at weaning and initial weight at the beginning of the test feeding is also given.
The steers were fed for the same length of time in the feed lot.
It is assumed that each calf has a unique dam and there are no twins or repeat matings.
Harvey (1960) is one of the earliest papers presenting least squares means (lsmeans).

**Source**


**References**

Also appears in the 'dmm' package as 'harv101.df' See that package vignette for a complete analysis of the data.
Examples

```r
## Not run:

library(agridat)
data(harvey.lsmeans)
dat = harvey.lsmeans

libs(lattice)
dotplot(adg ~ sire|line,dat,
    main="harvey.lsmeans", xlab="sire", ylab="average daily gain")

# Model suggested by Harvey on page 103
m0 <- lm(adg ~ 1 + line + sire + damage + line:damage + weanage +
    weight, data=dat)

# Due to contrast settings, it can be hard to compare model coefficients to Harvey,
# but note the slopes of the continuous covariates match Harvey p. 107, where his
# b is weanage, d is weight
# coef(m0)
# weanage weight
# -0.008154879 0.001970446

# A quick attempt to reproduce table 4 of Harvey, p. 109. Not right.
# libs(emmeans)
# emmeans(m0,c('line','sire','damage'))

## End(Not run)
```

---

**harville.lamb**

*Birth weight of lambs from different lines/sires*

**Description**

Birth weight of lambs from different lines/sires

**Usage**

```r
data("harville.lamb")
```

**Format**

A data frame with 62 observations on the following 4 variables.

- `line` genotype line number
- `sire` sire number
- `damage` dam age, class 1,2,3
- `weight` lamb birth weight
Details

Weight at birth of 62 lambs. There were 5 distinct lines.

Some sires had multiple lambs. Each dam had one lamb.

The age of the dam is a category: 1 (1-2 years), 2 (2-3 years) or 3 (over 3 years).

Note: Jiang gives the data in table 1.2, but there is a small error. Jiang has a weight 9.0 for sire 31, line 3, age 3. The correct value is 9.5.

Source


References

Jiming Jiang, Linear and Generalized Linear Mixed Models and Their Applications. Table 1.2.


Examples

```r
## Not run:

library(agridat)
data(harville.lamb)
dat <- harville.lamb
dat <- transform(dat, line=factor(line), sire=factor(sire), damage=factor(damage))

library(lattice)
bwplot(weight ~ line, dat, 
      main="harville.lamb",
      xlab="line", ylab="birth weights")

if(0){
  libs(lme4, lucid)

  m1 <- lmer(weight ~ -1 + line + damage + (1|sire), data=dat)
  summary(m1)
  vc(m1) # Khuri reports variances 0.5171, 2.9616

  ## grp  var1  var2  vcov  sdcor
  ## sire (Intercept) <NA> 0.5171 0.7191
  ## Residual <NA> <NA> 2.962 1.721
}

## End(Not run)
```
Diallel cross of Aztec tobacco in 2 reps

Format

year year
block block factor, 2 levels
male male parent, 8 levels
female female parent
day mean flowering time (days)

Details

Data was collected in 1951 (Hayman 1954a) and 1952 (Hayman 1954b).

In each year there were 8 varieties of Aztec tobacco, *Nicotiana rustica* L.

Each cross/self was represented by 10 progeny, in two plots of 5 plants each. The data are the mean flowering time per plot.

Note, the 1951 data as published in Hayman (1954a) Table 5 contain "10 times the mean flowering time". The data here have been divided by 10 so as to be comparable with the 1952 data.

Hayman (1954b) says "Table 2 lists...three characters from a diallel cross of Nicotiana rustica varieties which was repeated for three years." This seems to indicate that the varieties are the same in 1951 and 1952. Calculating the GCA effects separately for 1951 and 1952 and then comparing these estimates shows that they are highly correlated.

Source


References

# For 1951 data

# For 1952 data

Examples

## Not run:

```r
library(agridat)

# 1951 data. Fit the first REML model of Mohring 2011 Supplement.
data(hayman.tobacco)
dat1 <- subset(hayman.tobacco, year==1951)

# Hayman's model
# dat1 <- subset(hayman.tobacco, year==1951)
# libs(lmDiallel)
# m1 <- lm.diallel(day ~ male+female, Block=block, data=dat1, fct="HAYMAN2")
# anova(m1) # Similar to table 7 of Hayman 1954a
## Response: day
## Df Sum Sq Mean Sq F value Pr(>F)
## Block 1 1.42 1.42 0.3416 0.56100
## Mean Dom. Dev. 1 307.97 307.97 73.8840 3.259e-12 ***
## GCA 7 2777.17 396.74 95.1805 < 2.2e-16 ***
## Dom. Dev. 7 341.53 48.79 11.7050 1.957e-09 ***
## SCA 20 372.89 18.64 4.4729 2.560e-06 ***
## RGCA 7 67.39 9.63 2.3097 0.03671 *
## RSCA 21 123.73 5.89 1.4135 0.14668
## Residuals 63 262.60

# Griffing's model
# https://www.statforbiology.com/2021/stat_met_diallel_griffing/
data(hayman.tobacco)
dat1 <- subset(hayman.tobacco, year==1951)
# libs(lmDiallel)
# contrasts(dat1$block) <- “contr.sum”
# dmod1 and dmod2 are the same model with different syntax
# dmod1 <- lm(day ~ block + GCA(male, female) + tSCA(male, female) +
# REC(male, female), data = dat1)
# dmod2 <- lm.diallel(day ~ male + female, Block=block,
# data = dat1, fct = "GRIFFING1")
# anova(dmod1)
# anova(dmod2)
## Response: day
## Df Sum Sq Mean Sq F value Pr(>F)
## Block 1 1.42 1.42 0.3416 0.56100
## GCA 7 2777.17 396.74 95.1805 < 2.2e-16 ***
## SCA 28 1022.38 36.51 8.7599 6.656e-13 ***
## Reciprocals 28 191.12 6.83 1.6375 0.05369 .
## Residuals 63 262.60

# Make a factor ‘comb’ in which G1xG2 is the same cross as G2xG1
dat1 <- transform(dat1, comb =
  ifelse(as.character(male) < as.character(female),
  "G1xG2", "G2xG1")
```
```r
paste0(male, female), paste0(female, male))
# 'dr' is the direction of the cross, 0 for self
dat1$dr <- 1
dat1 <- transform(dat1,
  dr = ifelse(as.character(male) < as.character(female), -1, dr))

# asreml r version 3 & 4 code for Mixed Griffing.
# Mohring Table 2, column 2 (after dividing by 10^2) gives variances:
# GCA 12.77, SCA 11.09, RSCA .65, Error 4.23.
# Mohring Supplement ASREML code part1 model is:
# y ~ mu r !r mother and(father) combination combination.dr
# Note that the levels of 'male' and 'female' are the same, so the
# and(female) term tells asreml to use the same levels (or, equivalently,
# fix the correlation of the male/female levels to be 1.
# The block effect is minimal and therefore ignored.
## libs(asreml, lucid)
## m1 <- asreml(day~1, data=dat1,
## random = ~ male + and(female) + comb + comb:dr)
## vc(m1)
## effect component std.error z.ratio con
## male!male.var 12.77 7.502 1.7 Positive
## comb!comb.var 11.11 3.353 3.3 Positive
## comb:dr!comb.var 0.6603 0.4926 1.3 Positive
## R!variance 4.185 0.7449 5.6 Positive
# ----------
# 1952 data. Reproduce table 3 and figure 2 of Hayman 1954b.

dat2 <- subset(hayman.tobacco, year==1952)
# Does flowering date follow a gamma dist? Maybe.
libs(lattice)
densityplot(~day, data=dat2, main="hayman.tobacco",
  xlab="flowering date")
d1 <- subset(dat2, block=='B1')
d2 <- subset(dat2, block=='B2')

libs(reshape2)
m1 <- acast(d1, male~female, value.var="day")
m2 <- acast(d2, male~female, value.var="day")

mn1 <- (m1+t(m1))/2
mn2 <- (m2+t(m2))/2

# Variance and covariance of 'rth' offspring
vr1 <- apply(mn1, 1, var)
v2 <- apply(mn2, 1, var)
wr1 <- apply(mn1, 1, cov, diag(mn1))
```
hazell.vegetables

Gross profit for 4 vegetable crops in 6 years

Description

Gross profit for 4 vegetable crops in 6 years

Usage

data("hazell.vegetables")

Format

A data frame with 6 observations on the following 5 variables.

year year factor, 6 levels
carrot Carrot profit, dollars/acre
celery Celery profit, dollars/acre
cucumber Cucumber profit, dollars/acre
pepper Pepper profit, dollars/acre

Details

The values in the table are gross profits (loss) in dollars per acre. The criteria in the example below are (1) total acres < 200, (2) total labor < 10000, (3) crop rotation.

The example shows how to use linear programming to maximize expected profit.

Source

References


Examples

```r
## Not run:

library(agridat)
data(hazell.vegetables)
dat <- hazell.vegetables

library(lattice)
xplot(carrot+celery+cucumber+pepper ~ year, dat, 
     ylab="yearly profit by crop", 
     type='b', auto.key=list(columns=4), 
     panel.hline=0)

# optimal strategy for planting crops (calculated below)
dat2 <- apply(dat[, -1], 1, function(x) x*c(0, 27.5, 100, 72.5))/1000
colnames(dat2) <- rownames(dat)

barplot(dat2, legend.text=c(" 0 carrot", "27.5 celery", " 100 cucumber", "72.5 pepper"),
        xlim=c(0,7), ylim=c(-5,120),
        col=c("orange", "green", "forestgreen", "red"),
        xlab="year", ylab="Gross profit, $1000",
        main="hazell.vegetables - retrospective profit from optimal strategy",
        args.legend=list(title="acres, crop"))

library(linprog)
# colMeans(dat[, -1])
# 252.8333 442.6667 283.8333 515.8333

# cvec = avg across-years profit per acre for each crop
cvec <- c(253, 443, 284, 516)

# Maximize c'x for Ax=b
A <- rbind(c(1,1,1,1), c(25,36,27,87), c(-1,1,-1,1))
colnames(A) <- names(cvec) <- c("carrot","celery","cucumber","pepper")
rownames(A) <- c("land","labor","rotation")

# bvec criteria = (1) total acres < 200, (2) total labor < 10000, 
# (3) crop rotation.
bvec <- c(200, 10000, 0)
const.dir <- c("<", "<", "=")

m1 <- solveLP(cvec, bvec, A, maximum=TRUE, const.dir=const.dir, lpSolve=TRUE)
# m1$solution # optimal number of acres for each crop
# carrot celery cucumber pepper
# 0.00000 27.45098 100.00000 72.54902

# Average income for this plan
## sum(cvec * m1$solution)
## [1] 77996.08

# Year-to-year income for this plan
## as.matrix(dat[, -1])
## [,1]
## [1,] 80492.16
## [2,] 80431.37
## [3,] 81884.31
## [4,] 106868.63
## [5,] 37558.82
## [6,] 80513.73

# optimum allocation that minimizes year-to-year income variability.
# brute-force search

# For generality, assume we have unequal probabilities for each year.
probs <- c(.15, .20, .20, .15, .15, .15)
# Randomly allocate crops to 200 acres, 100,000 times
#set.seed(1)
mat <- matrix(runif(4*100000), ncol=4)
mat <- 200*sweep(mat, 1, rowSums(mat), "/")
# each row is one strategy, showing profit for each of the six years
# profit <- mat
profit <- tcrossprod(mat, as.matrix(dat[, -1])) # Each row is profit, columns are years
# calculate weighted variance using year probabilities
wtvar <- apply(profit, 1, function(x) cov.wt(as.data.frame(x), wt=probs)$cov)
# five best planting allocations that minimizes the weighted variance
ix <- order(wtvar)[1:5]
mat[ix]
## carrot celery cucumber pepper
## [1,] 71.26439 28.09259 85.04644 15.59657
## [2,] 72.04428 27.53299 84.29760 16.12512
## [3,] 72.16332 27.35147 84.16669 16.31853
## [4,] 72.14622 29.24590 84.12452 14.48335
## [5,] 68.95226 27.39246 88.61828 15.03700

## End(Not run)

#### heady.fertilizer

**Yield of corn, alfalfa, clover with two fertilizers**

**Description**

Yield of corn, alfalfa, clover with two fertilizers

**Usage**

`data("heady.fertilizer")`
Format

A data frame with 81 observations on the following 3 variables.

crop  crop
rep  replicate (not block)
P  phosphorous, pounds/acre
K  potassium, pounds/acre
N  nitrogen, pounds/acre
yield  yield

Details

Heady et al. fit two-variable semi-polynomial response surfaces for each crop.
Clover and alfalfa yields are in tons/acre. The clover and alfalfa experiments were grown in 1952.
Corn yields are given as bu/acre. The corn experiments were grown in 1952 and 1953. The same
test plots were used in 1953 and in 1952, but no fertilizer was applied in 1953–any response in yield
is due to residual fertilizer from 1952.
All experiments used an incomplete factorial design. Not all treatment combinations were present.

Source

Optima in Fertilizer Use. Agricultural Experiment Station, Iowa State College. Research bulletin
424. Pages 330-332. https://lib.dr.iastate.edu/cgi/viewcontent.cgi?filename=12&article=1032&context=ag_researchbulletins

References

Pesek, John and Heady, Earl O. 1956. A two nutrient-response function with determination of
economic optima for the rate and grade of fertilizer for alfalfa. Soil Science Society of America

Examples

library(agridat)
data(heady.fertilizer)
dat <- heady.fertilizer

libs(lattice)
xyplot(yield ~ P|crop, data=dat, scales=list(relation="free"),
groups=factor(paste(dat$N,dat$K)), auto.key=list(columns=5),
main="heady.fertilizer", xlab="Phosphorous")

# Corn. Matches Heady, p. 292
d1 <- subset(dat, crop="corn")
m1 <- lm(yield ~ N + P + sqrt(N) + sqrt(P) + sqrt(N*P), data=d1)
summary(m1)
heath.raddish.uniformity

# Alalfa. Matches Heady, p. 292. Also Peck equation 3, p. 241

d2 <- subset(dat, crop=="alfalfa")
m2 <- lm(yield ~ K + P + sqrt(K) + sqrt(P) + sqrt(K*P), data=d2)
summary(m2)

## Coefficients:
##         Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8735521  0.1222501 15.326  < 2e-16 ***
## K         -0.0013943  0.0007371  -1.891   0.061237 .
## P         -0.0050195  0.0007371  -6.810   5.74e-10 ***
## sqrt(K)    0.0617458  0.0160142   3.856   0.000196 ***
## sqrt(P)    0.1735383  0.0160142  10.837   < 2e-16 ***
## sqrt(K * P) -0.0014402  0.0007109  -2.026   0.045237 *

# Clover. Matches Heady, p. 292.
d3 <- subset(dat, crop=="clover")
m3 <- lm(yield ~ P + sqrt(K) + sqrt(P) + sqrt(K*P), data=d3)
summary(m3)

# Corn with residual fertilizer. Matches Heady eq, 56, p. 322.
d4 <- subset(dat, crop=="corn2")
m4 <- lm(yield ~ N + P + sqrt(N) + sqrt(P) + sqrt(N*P), data=d4)
summary(m4)

## Not run:
  libs(rgl)
  with(d1, plot3d(N,P,yield))
  with(d2, plot3d(K,P,yield))
  with(d3, plot3d(K,P,yield))
  with(d4, plot3d(N,P,yield)) # Mostly linear in both N and P
  rgl.close()

## End(Not run)

---

**heath.raddish.uniformity**

*Uniformity trial of raddish*

---

**Description**

Uniformity trial of raddish in four containers.

**Usage**

`data("heath.raddish.uniformity")`
Format

A data frame with 400 observations on the following 4 variables.

row row
col column
block block
yield weight per plant

Details

Weight of 399 radish plants grown at 1in x 1in spacing in four plastic basins. Seed wetted 1968-02-15, planted 1968-02-17, harvested 1968-03-26.

Heath said, Most of the large plants were round the edges...one important source of variation might have been competition for light.

Source


References

None

Examples

```
## Not run:
require(agridat)
data(heath.raddish.uniformity)
dat <- heath.raddish.uniformity
libs(desplot, dplyr)
desplot(dat, yield ~ col*row|block,
       aspect=1,
       main="heath.raddish.uniformity")

# Indicator for border/interior plants
dat <- mutate(dat,
              inner = row > 1 & row < 10 & col > 1 & col < 10)
# Heath has 5.80 and 9.63 (we assume this is a typo of 9.36)
dat <- group_by(dat, inner)
summarize(dat, mean=mean(yield, na.rm=TRUE))
# Interior plots are significantly lower yielding
anova(aov(yield ~ block + inner, dat))
# lattice::bwplot(yield ~ inner, dat, horiz=0)
# similar to Heath fig 2-2
# lattice::histogram(~ yield|inner, dat, layout=c(1,2), n=20)
```

## End(Not run)
**henderson.milkfat**  
*Milk fat yields for a single cow*

**Description**  
Average daily fat yields (kg/day) from milk from a single cow for each of 35 weeks.

**Format**  
A data frame with 35 observations on the following 2 variables.

- **week**  
  week, numeric
- **yield**  
  yield, kg/day

**Source**  
Charles McCulloch. Workshop on Generalized Linear Mixed Models.  
Used with permission of Charles McCulloch and Harold Henderson.

**References**  
None.

**Examples**

```r
## Not run:  
library(agridat)  
data(henderson.milkfat)  
dat <- henderson.milkfat

plot(yield~week, data=dat, cex = 0.8, ylim=c(0,.9),  
     main="henderson.milkfat", xlab = "Week",  
     ylab = "Fat yield (kg/day)"

# Yield ~ a * t^b * exp(g*t) # where t is time
m1 <- nls(yield ~ alpha * week^beta * exp(gamma * week),  
          data=dat,  
          start=list(alpha=.1, beta=.1, gamma=.1))

# Or, take logs and fit a linear model
# log(yield) ~ log(alpha) + beta*log(t) + gamma*t
m2 <- lm(log(yield) ~ 1 + log(week) + week, dat)

# Or, use glm and a link to do the transform
m3 <- glm(yield ~ 1 + log(week) + week, quasi(link = "log"), dat)

# Note: m2 has E[log(y)] = log(alpha) + beta*log(t) + gamma*t  
# and m3 has log(E[y]) = log(alpha) + beta*log(t) + gamma*t
```
# Generalized additive models

```r
libs("mgcv")
m4 <- gam(log(yield) ~ s(week), gaussian, dat)
m5 <- gam(yield ~ s(week), quasi(link = "log"), dat)
```

# Model predictions

```r
pdat <- data.frame(week = seq(1, 35, by = 0.1))
pdat <- transform(pdat, p1 = predict(m1, pdat),
                  p2 = exp(predict(m2, pdat)), # back transform
                  p3 = predict(m3, pdat, type="resp"), # response scale
                  p4 = exp(predict(m4, pdat)),
                  p5 = predict(m5, pdat, type="response"))
```

# Compare fits

```r
with(pdat, {
  lines(week, p1)
  lines(week, p2, col = "red", lty="dotted")
  lines(week, p3, col = "red", lty="dashed")
  lines(week, p4, col = "blue", lty = "dashed")
  lines(week, p5, col = "blue")
})
```

legend("topright",
c("obs", "lm, log-transformed", "glm, log-link",
  "gam, log-transformed", "gam, log-link"),
lty = c("solid", "dotted", "dashed", "dashed", "solid"),
col = c("black", "red", "red", "blue", "blue"),
cex = 0.8, bty = "n")
```
Details

Experiment was conducted in 2006 at 5 sites in Minnesota.

Source


Electronic data kindly supplied by Jose Hernandez.

Examples

```r
## Not run:

library(agridat)
data(hernandez.nitrogen)
dat <- hernandez.nitrogen
cprice <- 118.1 # $118.1/Mg or $3/bu
nprice <- 0.6615 # $0.66/kg N or $0.30/lb N

# Hernandez optimized yield with a constraint on the ratio of the prices.
# Simpler to just calculate the income and optimize that.
dat <- transform(dat, inc = yield * cprice - nitro * nprice)
libs(lattice)
xyplot(inc ~ nitro|site, dat, groups=rep, auto.key=list(columns=4),
xlab="nitrogen", ylab="income", main="hernandez.nitrogen")

# Site 5 only
dat1 <- subset(dat, site=="S5")

# When we optimize on income, a simple quadratic model works just fine,
# and matches the results of the nls model below.
# Note, `poly(nitro)` gives weird coefs
lm1 <- lm(inc ~ 1 + nitro + I(nitro^2), data=dat1)
c1 <- coef(lm1)
-c1[2] / (2*c1[3])
## nitro
## 191.7198 # Optimum nitrogen is 192 for site 5

# Use the delta method to get a conf int
libs("car")
del1 <- deltaMethod(lm1, "-b1/(2*b2)", parameterNames= paste("b", 0:2, sep=""))
# Simple Wald-type conf int for optimum
del1$Est + c(-1,1) * del1$SE * qt(1-.1/2, nrow(dat1)-length(coef(lm1)))
## 118.9329 264.5067

# Nonlinear regression
# Reparameterize $b0 + b1*x + b2*x^2$ using $th2 = -b1/2b2$ so that $th2$ is optimum
nls1 <- nls(inc - th1- (2*th2+th12)*nitro + th12*nitro^2,
data = dat1, start = list(th11 = 5, th2 = 150, th12 =-.1),)
```

```
```r

wald <- function(object, alpha=0.1){
  nobs <- length(resid(object))
  npar <- length(coef(object))
  est <- coef(object)
  stderr <- summary(object)$parameters[,2]
  tval <- qt(1-alpha/2, nobs-npar)
  ci <- cbind(est - tval * stderr, est + tval * stderr)
  colnames(ci) <- paste(round(100*c(alpha/2, 1-alpha/2), 1), "pct", sep = "")
  return(ci)
}

round(wald(nls1),2)
##
## th11  936.44 1081.93
## th2  118.93  264.51 # th2 is the optimum
## th12  -0.03  -0.01

# Likelihood conf int
libs(MASS)
round(confint(nls1, "th2", level = 0.9),2)
##
## 147.96 401.65

# Bootstrap conf int
libs(boot)
dat1$fit <- fitted(nls1)
bootfun <- function(rs, i) { # bootstrap the residuals
  dat1$y <- dat1$fit + rs[i]
  coef(nls(y ~ th11 - (2*th2*th12)*nitro + th12*nitro^2, dat1,
     start = coef(nls1) ))
}
res1 <- scale(resid(nls1), scale = FALSE) # remove the mean. Why? It is close to 0.
set.seed(5) # Sometime the bootstrap fails, but this seed works
boot1 <- boot(res1, bootfun, R = 500)
boot.ci(boot1, index = 2, type = c("perc"), conf = 0.9)
## Level   Percentile
## 90

## End(Not run)
```

---

**hessling.argentina**  
*Relation between wheat yield and weather in Argentina*

**Description**

Relation between wheat yield and weather in Argentina
**Format**

A data frame with 30 observations on the following 15 variables.

- **yield**: average yield, kg/ha
- **year**: year
- **p05**: precipitation (mm) in May
- **p06**: precip in June
- **p07**: precip in July
- **p08**: precip in August
- **p09**: precip in September
- **p10**: precip in October
- **p11**: precip in November
- **p12**: precip in December
- **t06**: june temperature deviation from normal, deg Celsius
- **t07**: july temp deviation
- **t08**: august temp deviation
- **t09**: september temp deviation
- **t10**: october temp deviation
- **t11**: november temp deviation

**Details**

In Argentina wheat is typically sown May to August. Harvest begins in November or December.

**Source**


**Examples**

```r
## Not run:
library(agridat)
data(hessling.argentina)
dat <- hessling.argentina

# Fig 1 of Hessling. Use avg Aug-Nov temp to predict yield
dat <- transform(dat, avetmp=(t08+t09+t10+t11)/4) # Avg temp
m0 <- lm(yield ~ avetmp, dat)
plot(yield~year, dat, ylim=c(100,1500), type='l', main="hessling.argentina: observed (black) and predicted yield (blue)")
lines(fitted(m0)~year, dat, col="blue")

# A modern, PLS approach
libs(pls)
```
hildebrand.systems

**Description**

Maize yields for four cropping systems at 14 on-farm trials.

**Format**

A data frame with 56 observations on the following 4 variables.

- **village**: village, 2 levels
- **farm**: farm, 14 levels
- **system**: cropping system
- **yield**: yield, t/ha

**Details**

Yields from 14 on-farm trials in Phalombe Project region of south-eastern Malawi. The farms were located near two different villages.

On each farm, four different cropping systems were tested. The systems were: LM = Local Maize, LMF = Local Maize with Fertilizer, CCA = Improved Composite, CCAF = Improved Composite with Fertilizer.

**Source**


**References**

Examples

## Not run:

library(agridat)
data(hildebrand.systems)
dat <- hildebrand.systems

# Piepho 1998 Fig 1
libs(lattice)
dotplot(yield ~ system, dat, groups=village, auto.key=TRUE,
    main="hildebrand.systems", xlab="cropping system by village")

# Plot of risk of 'failure' of System 2 vs System 1
s11 = .30; s22 <- .92; s12 = .34
mu1 = 1.35; mu2 = 2.70
lambda <- seq(from=0, to=5, length=20)
system1 <- pnorm((lambda-mu1)/sqrt(s11))
system2 <- pnorm((lambda-mu2)/sqrt(s22))

# A simpler view
plot(lambda, system1, type="l", xlim=c(0,5), ylim=c(0,1),
    xlab="Yield level", ylab="Prob(yield < level)",
    main="hildebrand.systems - risk of failure for each system")
lines(lambda, system2, col="red")

# Prob of system 1 outperforming system 2. Table 8
pnorm((mu1-mu2)/sqrt(s11+s22-2*s12))
    # .0331

# ----------
libs(asreml,lucid) # asreml4

# Environmental variance model, unstructured correlations

dat <- dat[order(dat$system, dat$farm),]
m1 <- asreml(yield ~ system, data=dat,
    resid = ~us(system):farm)

# Means, table 5
## predict(m1, data=dat, classify="system")$pvals
## system pred.value std.error est.stat
## CCA 1.164 0.2816 Estimable
## CCAF 2.657 0.3747 Estimable
## LM 1.35 0.1463 Estimable
## LMF 2.7 0.2561 Estimable

# Variances, table 5
# vc(m1)[c(2,4,7,11),]
## effect component std.error z.ratio constr
## R!system.CCA:CCA 1.11 0.4354 2.5 pos
## Stability variance model

```r
m2 <- asreml(yield ~ system, data=dat,
             random = ~ farm,
             resid = ~ dsum(~ units|system))
```

m2 <- update(m2)

```r
# predict(m2, data=dat, classify="system")$pvals
```

## # Variances, table 6
## vc(m2)

```
##               effect component std.error z.ratio bound
## farm          0.2998     0.1187 2.5  P  0
## system_CCA!R  0.4133     0.1699 2.4  P  0
## system_CCAF!R 1.265      0.5152 2.5  P  0
## system_LM!R   0.0003805 0.05538 0.0069 P  1.5
## system_LMF!R  0.5294     0.2295 2.3  P  0
```

```
## End(Not run)
```

### Description

Counts of arthropods in a grid-sampled wheat field

### Usage

```r
data("holland.arthropods")
```

### Format

A data frame with 63 observations on the following 8 variables.

- `row` row
- `col` column
- `n.brevicollis` species counts
- `linyphiidae` species counts
- `collembola` species counts
- `carabidae` species counts
- `lycosidae` species counts
- `weedcover` percent weed cover
Details

Arthropods were sampled at 30m x 30m grid in a wheat field near Wimborne, Dorest, UK on 6 dates in Jun/Jul 1996. Arthropod counts were aggregated across the 6 dates.

Holland et al. used SADIE (Spatial Analysis by Distance Indices) to look for spatial patterns. Significant patterns were found for N. brevicollis, Carabidae, Lycosidae. The Lycosidae counts were also significantly associated with weed cover.

Used with permission of John Holland.

Source


Examples

```r
## Not run:

library(agridat)
data(holland.arthropods)
dat <- holland.arthropods

# use log count to make it possible to have same scale for insects
libs(reshape2, lattice)
grays <- colorRampPalette(c("#d9d9d9","#252525"))
dat2 <- melt(dat, id.var=c("row","col"))
contourplot(log(value) ~ col*row|variable, dat2,
           col.regions=grays(7), region=TRUE,
           main="holland.arthropods - log counts in winter wheat")

if(0){
  # individual species
  libs(lattice)
grays <- colorRampPalette(c("#d9d9d9","#252525"))
contourplot(linyphiidae ~ col*row, dat, at=c(0,40,80,120,160,200), region=TRUE,
            col.regions=grays(5),
            main="holland.arthropods - linyphiidae counts in winter wheat")
contourplot(n.brevicollis ~ col*row, dat, region=TRUE)
contourplot(linyphiidae ~ col*row, dat, region=TRUE)
contourplot(collembola ~ col*row, dat, region=TRUE)
contourplot(carabidae ~ col*row, dat, region=TRUE)
contourplot(lycosidae ~ col*row, dat, region=TRUE)
contourplot(weedcover ~ col*row, dat, region=TRUE)
}
```

## End(Not run)
holshouser.splitstrip  Split-strip-plot of soybeans

Description

Split-strip-plot of soybeans

Format

A data frame with 160 observations on the following 8 variables.

- block: block factor, 4 levels
- plot: plot number
- cultivar: cultivar factor, 4 levels
- spacing: row spacing
- pop: population (thousand per acre)
- yield: yield
- row: row
- col: column

Details

Within each block, cultivars were whole plots. Within whole plots, spacing was applied in strips vertically, and population was applied in strips horizontally.

Used with permission of David Holshouser at Virginia Polytechnic.

Source


Examples

```r
## Not run:

library(agridat)
data(holshouser.splitstrip)
dat <- holshouser.splitstrip
dat$spacing <- factor(dat$spacing)
dat$pop <- factor(dat$pop)

# Experiment layout and field trends
libs(desplot)
desplot(dat, yield ~ col*row, 
   out1=block, # unknown aspect
   main="holshouser.splitstrip")
```
holtsmark.timothy.uniformity

holtsmark.timothy.uniformity

Uniformity trial of timothy

Description

Uniformity trial of timothy hay circa 1905

Usage

data("holtsmark.timothy.uniformity")
**Format**

A data frame with 240 observations on the following 3 variables.

- **row** row
- **col** column
- **yield** yield per plot, kg

**Details**

Field width: 40 plots * 5 m = 200 m.
Field length: 6 plots * 5 m = 30 m

Holtsmark & Larsen used this trial to compare standard deviations of different sized plots (combined from smaller plots).

**Source**

Holtsmark, G and Larsen, BR (1905). Om Muligheder for at indskraenke de Fejl, som ved Markfor-

**References**

Theodor Roemer (1920). *Der Feldversuch*.

**Examples**

```r
## Not run:
library(agridat)
data(holtsmark.timothy.uniformity)
dat <- holtsmark.timothy.uniformity

# Define diagonal 'check' plots like Holtsmark does
dat <- transform(dat,
  check = ifelse(floor((row+col)/3)==(row+col)/3, "C", "")
)

libs(desplot)
desplot(dat, yield ~ col*row,
  flip=TRUE, text=check, show.key=FALSE,
  aspect=30/200, # true aspect
  main="holtsmark.timothy.uniformity")
# sd(dat$yield) # 2.92 matches Holtsmark p. 348

## End(Not run)
```
Multi-environment trial of wheat to illustrate stability statistics

Description

Multi-environment trial to illustrate stability statistics

Usage

data("huehn.wheat")

Format

A data frame with 200 observations on the following 3 variables.

gen genotype
env environment
yield yield dt/ha

Details

Yields for a winter-wheat trial of 20 genotypes at 10 environments.
Note: Huehn 1979 does not use genotype-centered data when calculating stability statistics.

Source


References


Examples

```r
## Not run:
library(agridat)
data(huehn.wheat)
dat <- huehn.wheat

# Nassar & Huehn, p. 51 "there is no evidence for differences in stability
# among the 20 varieties".
labs(gge)
m1 <- gge(dat, yield ~ gen*env)
biplot(m1, main="huehn.wheat")
```
hughes.grapes

**RCB experiment of grape, disease incidence**

**Description**

Disease incidence on grape leaves in RCB experiment with 6 different treatments.

**Format**

A data frame with 270 observations on the following 6 variables.

- **block**: Block factor, 1-3
- **trt**: Treatment factor, 1-6
- **vine**: Vine factor, 1-3
- **shoot**: Shoot factor, 1-5
- **diseased**: Number of diseased leaves per shoot
- **total**: Number of total leaves per shoot
Details

These data come from a study of downy mildew on grapes. The experiment was conducted at Wooster, Ohio, on the experimental farm of the Ohio Agricultural Research and Development Center, Ohio State University.

There were 3 blocks with 6 treatments. Treatment 1 is the unsprayed control. On 30 Sep 1990, disease incidence was measured. For each plot, 5 randomly chosen shoots on each of the 3 vines were observed. The canopy was closed and shoots could be intertwined. On each shoot, the total number of leaves and the number of infected leaves were recorded.

Used with permission of Larry Madden.

Source


References


Examples

```r
# Not run:

library(agridat)
data(hughes.grapes)
dat <- hughes.grapes

dat <- transform(dat, rate = diseased/total, plot=trt:block)

# Trt 1 has higher rate, more variable, Trt 3 lower rate, less variable
libs(lattice)
foo <- bwplot(rate ~ vine|block*trt, dat, main="hughes.grapes",
             xlab="vine")
libs(latticeExtra)
useOuterStrips(foo)

# Table 1 of Piepho 1999
tapply(dat$rate, dat$trt, mean) # trt 1 does not match Piepho
tapply(dat$rate, dat$trt, max)

# Piepho model 3. Binomial data. May not be exactly the same model

# Use the binomial count data with lme4
libs(lme4)
m1 <- glmer(cbind(diseased, total-diseased) ~ trt + block + (1|plot/vine),
             data=dat, family=binomial)
```
m1

# Switch from binomial counts to bernoulli data
libs(aod)
bdat <- splitbin(cbind(diseased, total-diseased) ~ block+trt+plot+vine+shoot,
    data=dat)$tab
names(bdat)[2] <- 'y'

# Using lme4
m2 <- glmer(y ~ trt + block + (1|plot/vine), data=bdat, family=binomial)
m2

# Now using MASS::glmmPQL
libs(MASS)
m3 <- glmmPQL(y ~ trt + block, data=bdat,
    random=1|plot/vine, family=binomial)
m3

## End(Not run)

---

 hunter.corn  Multi-environment trial of corn with nitrogen fertilizer

Description

Corn yield response to nitrogen

Format

A data frame with 54 observations on the following 4 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>nitro</td>
<td>nitrogen fertilizer, pound/acre</td>
</tr>
<tr>
<td>year</td>
<td>year</td>
</tr>
<tr>
<td>loc</td>
<td>location</td>
</tr>
<tr>
<td>yield</td>
<td>yield, bu/ac</td>
</tr>
</tbody>
</table>

Details

Experiments were conducted in eastern Oregon during the years 1950-1952.
Planting rates varied from 15,000 to 21,000 planter per acre.

Source

igue.sugarcane.uniformity

References

Examples

```r
library(agridat)
data(hunter.corn)
dat <- hunter.corn

dat <- transform(dat, env=factor(paste(loc,year)))

libs(lattice)
xyplot(yield~nitro|env, dat, type='b',
main="hunter.corn - nitrogen response curves")
```

ique.sugarcane.uniformity

Uniformity trial with sugarcane

Description
Uniformity trial with sugarcane in Brazil, 1982.

Usage
data("igue.sugarcane.uniformity")

Format
A data frame with 1512 observations on the following 3 variables.

- row row
- col column
- yield yield, kg/plot

Details
A uniformity trial with sugarcane in the state of Sao Paulo, Brazil, in 1982. The field was 40 rows, each 90 m long, with 1.5 m between rows.
Field width: 36 plots * 1.5 m = 54 m
Field length: 42 plots * 2 m = 84 m
ilri.sheep

Birth weight and weaning weight of Dorper x Red Maasi lambs

Description

Birth weight and weaning weight of 882 lambs from a partial diallel cross of Dorper and Red Maasi breeds.

Format

A data frame with 882 observations on the following 12 variables.

- lamb: lamb id
- sex: sex of lamb, M=Male/F=Female
- gen: genotype, DD, DR, RD, RR
- birthwt: weight of lamb at birth, kg
- weanwt: weight of lamb at weaning, kg

Examples

```r
library(agridat)
data(igue.sugarcane.uniformity)
dat <- igue.sugarcane.uniformity

# match Igue CV top row of page 171
sd(dat$yield)/mean(dat$yield) # 16.4

libs(desplot)
desplot(dat, yield ~ col*row, 
    flip=TRUE, tick=TRUE, aspect=(42*2)/(36*1.5), 
    main="igue.sugarcane.uniformity")
```

## End(Not run)
Red Maasai sheep in East Africa are perceived to be resistant to certain parasites. ILRI decided in 1990 to investigate the degree of resistance exhibited by this Red Maasai breed and initiated a study in Kenya. A susceptible breed, the Dorper, was chosen to provide a direct comparison with the Red Maasai. The Dorper is well-adapted to this area and is also larger than the Red Maasai, and this makes these sheep attractive to farmers.

Throughout six years from 1991 to 1996 Dorper (D), Red Maasai (R) and Red Maasai x Dorper crossed ewes were mated to Red Maasai and Dorper rams to produce a number of different lamb genotypes. For the purposes of this example, only the following four offspring genotypes are considered (Sire x Dam): D x D, D x R, R x D and R x R.

Records are missing in 182 of the lambs, mostly because of earlier death.
dat <- within(dat, {
  dl <- damage
  dl <- ifelse(dl < 3, 2, dl)
  dl <- ifelse(dl > 7, 8, dl)
  dq <- dl^2
})

dat <- subset(dat, !is.na(weanage))

# EDA
lattice
## bwplot(weanwt ~ year, dat, main="ilri.sheep", xlab="year", ylab="Wean weight",
##       panel=panel.violin) # Year effect
bwplot(weanwt ~ factor(dl), dat,
      main="ilri.sheep", xlab="Dam age", ylab="Wean weight") # Dam age effect
# bwplot(weanwt ~ gen, dat,
#       main="ilri.sheep", xlab="Genotype", ylab="Wean weight") # Genotype differences
xyplot(weanwt ~ weanage, dat, type=c("p", "smooth"),
       main="ilri.sheep", xlab="Wean age", ylab="Wean weight") # Age covariate

# case study page 4.18
lm1 <- lm(weanwt ~ year + sex + weanage + dl + dq + ewegen + ramgen, data=dat)
summary(lm1)
anova(lm1)

# ----------

lme4
lme1 <- lmer(weanwt ~ year + sex + weanage + dl + dq + ewegen + ramgen +
              (1|ewe) + (1|ram), data=dat)
print(lme1, corr=FALSE)
lme2 <- lmer(weanwt ~ year + sex + weanage + dl + dq + ewegen + ramgen +
              (1|ewe), data=dat)
lme3 <- lmer(weanwt ~ year + sex + weanage + dl + dq + ewegen + ramgen +
              (1|ram), data=dat)
anova(lme1, lme2, lme3)

# ----------

asreml,lucid
# case study page 4.20
m1 <- asreml(weanwt ~ year + sex + weanage + dl + dq + ramgen + ewegen,
             data=dat)
# wald(m1)

# case study page 4.26
m2 <- asreml(weanwt ~ year + sex + weanage + dl + dq + ramgen + ewegen,
             random = ~ ram + ewe, data=dat)
# wald(m2)

# case study page 4.37, year means
# predict(m2, data=dat, classify="year")
## year predicted.value standard.error est.status
Immer’s sugarbeet uniformity

**Uniformity trial of sugarbeets, measurements of yield, sugar, purity**

**Description**

Uniformity trial of sugarbeets, at Minnesota, 1930, with measurements of yield, sugar, purity.

**Format**

A data frame with 600 observations on the following 5 variables.

- `row` row
- `col` column
- `yield` yield, pounds
- `sugar` sugar percentage
- `purity` apparent purity

**Details**

Beets were planted in rows 22 inches apart, thinned to 1 plant per row. At harvest, the rows were marked into segments 33 feet long with 2 foot alleys between ends of plots. The harvested area was 60 rows 350 feet long.

Field width: 10 plots * 33 feet + 9 alleys * 2 feet = 348 feet
Field length: 60 plots/rows * 22 in/row / 12 in/feet = 110 feet

Planted in 1930. Field conditions were uniform. Beets were planted in rows 22 inches apart. After thinning, one beet was left in each 12-inch unit. At harvest, the field was marked out in plot 33 feet long, with a 2-foot alley between plots to minimize carryover from the harvester. A sample of 10 beets was taken uniformly (approximately every third beet) and measured for sugar percentage and apparent purity. The beets were counted at weighing time and the yields were calculated on the basis of 33 beets per plot.

Immer found that aggregating the data from one row to two resulted in a dramatic reduction in the standard error (for yield).
**Source**


**Examples**

```r
### Not run:

library(agridat)

data(immer.sugarbeet.uniformity)
dat <- immer.sugarbeet.uniformity

# Immer numbers rows from the top
libs(desplot)
desplot(dat, yield~col*row,
       aspect=110/348, tick=TRUE, flip=TRUE, # true aspect
       main="immer.sugarbeet.uniformity - yield")
desplot(dat, sugar~col*row,
       aspect=110/348, tick=TRUE, flip=TRUE,
       main="immer.sugarbeet.uniformity - sugar")
desplot(dat, purity~col*row,
       aspect=110/348, tick=TRUE, flip=TRUE,
       main="immer.sugarbeet.uniformity - purity")

pairs(dat[,c('yield','sugar','purity')],
       main="immer.sugarbeet.uniformity")

### End(Not run)
```

---

**ivins.herbs**

Percent ground cover of herbage species and nettles.

**Description**

Percent ground cover of herbage species and nettles.

**Format**

A data frame with 78 observations on the following 4 variables.

- **block** block, 6 levels
- **gen** genotype, 13 levels
- **nettle** percent ground cover of nettles
- **herb** percent ground cover of herbage species
Details

On the University of Nottingham farm, 13 different strains and species of herbage plants were sown on about 4 acres in an RCB design. Each grass species was sown together with white clover seed.

During establishment of the herbage plants, it became apparent that *Urtica dioica* (nettle) became established according to the particular herbage plant in each plot. In particular, nettle became established in plots sown with leguminous species and the two grass species. The graminaceous plots had less nettles.

The data here are the percentage ground cover of nettle and herbage plants in September 1951. Note, some of the percent ground cover amounts were originally reported as 'trace'. These have been arbitrarily set to 0.1 in this data.

<table>
<thead>
<tr>
<th>gen</th>
<th>species</th>
<th>strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>G01</td>
<td>Lolium perenne</td>
<td>Irish perennial ryegrass</td>
</tr>
<tr>
<td>G02</td>
<td>Lolium perenne</td>
<td>S. 23 perennial ryegrass</td>
</tr>
<tr>
<td>G03</td>
<td>Dactylis glomerata</td>
<td>Danish cocksfoot</td>
</tr>
<tr>
<td>G04</td>
<td>Dactylis glomerata</td>
<td>S. 143 cocksfoot</td>
</tr>
<tr>
<td>G05</td>
<td>Phleum pratense</td>
<td>American timothy</td>
</tr>
<tr>
<td>G06</td>
<td>Phleum pratense</td>
<td>S. 48 timothy</td>
</tr>
<tr>
<td>G07</td>
<td>Festuca pratensis</td>
<td>S. 215 meadow fescue</td>
</tr>
<tr>
<td>G08</td>
<td>Poa trivialis</td>
<td>Danish rough stalked meadow grass</td>
</tr>
<tr>
<td>G09</td>
<td>Cynosurus cristatus</td>
<td>New Zealand crested dogstail</td>
</tr>
<tr>
<td>G10</td>
<td>Trifolium pratense</td>
<td>Montgomery late red clover</td>
</tr>
<tr>
<td>G11</td>
<td>Medicago lupulina</td>
<td>Commercial black medick</td>
</tr>
<tr>
<td>G12</td>
<td>Trifolium repens</td>
<td>S. 100 white clover</td>
</tr>
<tr>
<td>G13</td>
<td>Plantago lanceolata</td>
<td>Commercial ribwort plantain</td>
</tr>
</tbody>
</table>

Source


References


Examples

```r
library(agridat)
data(ivins.herbs)
dat <- ivins.herbs

# Nettle is primarily established in legumes.
```
libs(lattice)
xypplot(herb~nettle|gen, dat, main="ivins.herbs - herb yield vs weeds",
       xlab="Percent groundcover in nettles",
       ylab="Percent groundcover in herbs")

# O'Brien used first 7 species to test gen differences
dat7 <- droplevels(subset(dat, is.element(gen, c('G01','G02','G03','G04','G05','G06','G07'))))
m1 <- lm(herb ~ gen + block, data=dat7)
anova(m1) # gen p-value is .041
## Response: herb
## Df  Sum Sq Mean Sq  F value Pr(>F)
## gen  6 1083.24 180.540 2.5518 0.04072 *
## block 5  590.69 118.138 1.6698 0.17236
## Residuals 30 2122.48  70.749

friedman.test(herb ~ gen|block, dat7) # gen p-value .056

—

iyer.wheat.uniformity  Uniformity trials of wheat in India

Description

Uniformity trials of wheat in India.

Usage

data("iyer.wheat.uniformity")

Format

A data frame with 2000 observations on the following 3 variables.

row row
col column
yield yield, ounces per plot

Details

Data collected at the Agricultural Sub-station in Karnal, India, in April 1978. A net area of 400 ft x 125 ft was harvested by dividing it into 80x25 units 5 ft x 5 ft after eliminating a minimum border of 3.5 ft all around the net area.

Field width: 80 plots * 5 feet = 400 feet
Field length: 25 rows * 5 feet = 125 feet

In a second paper, Iyer used this data to compare random vs. balanced arrangements of treatments to plots, with the conclusion that "it is very difficult to say which [method] is better. However, there is some tendency for the randomized arrangements to give more accurate results."
jansen.apple

Source


References

None.

Examples

```r
# Not run:
library(agridat)
data(iyer.wheat.uniformity)
dat <- iyer.wheat.uniformity
libs(desplot)
desplot(dat, yield ~ col*row,
  main="iyer.wheat.uniformity", tick=TRUE,
  aspect=(25*5)/(80*5)) # true aspect

# not exactly the same as Iyer table 1, p. 241
var(subset(dat, col <= 20)$yield)
var(subset(dat, col > 20 & col <= 40)$yield)
var(subset(dat, col > 40 & col <= 60)$yield)
var(subset(dat, col > 60)$yield)

# cv for 1x1 whole-field
# sd(dat$yield)/mean(dat$yield)
# 18.3

# End(Not run)
```

---

jansen.apple

Infestation of apple shoots by apple canker.

Description

Infestation of apple shoots by apple canker.

Usage

data("jansen.apple")
Format

A data frame with 36 observations on the following 5 variables.

- inoculum: inoculum level
- gen: genotype/variety
- block: block
- y: number of inoculations developing canker
- n: number of inoculations

Details

Shoots of apple trees were infected with fungus *Nectria galligena*, which may cause apple canker. The inoculum density treatment had 3 levels, measured in macroconidia per ml. There were 4 blocks.

Used with permission of J. Jansen. Electronic version supplied by Miroslav Zoric.

Source


References

None.

Examples

```r
## Not run:
library(agridat)
data(jansen.apple)
dat <- jansen.apple

libs(lattice)
xyplot(inoculum ~ y/n|gen, data=dat, group=block,
       layout=c(3,1),
       main="jansen.apple",
       xlab="Proportion infected per block/inoculum",
       ylab="Inoculum level")

## libs(lme4)
## # Tentative model. Needs improvement.
## m1 <- glmer(cbind(y,n-y) ~ gen + factor(inoculum) + (1|block),
##           data=dat, family=binomial)
## summary(m1)

## End(Not run)
```
Infestation of 16 carrot genotypes by fly larvae, comparing 2 treatments in 16 blocks.

Usage

data("jansen.carrot")

Format

A data frame with 96 observations on the following 5 variables.

trt  treatment
gen  genotype
block  block
n  number of carrots sampled per plot
y  number of carrots infested per plot

Details

This experiment was designed to compare different genotypes of carrots with respect to their resistance to infestation by larvae of the carrotfly.

There were 16 genotypes, 2 levels of pest-control treatments, conducted in 3 randomized complete blocks. About 50 carrots were sampled from each plot and evaluated. The data show the number of carrots and the number infested by fly larvae.

Used with permission of J. Jansen. Electronic version supplied by Miroslav Zoric.

Source


References

None.
Examples

```r
library(agridat)
data(jansen.carrot)
dat <- jansen.carrot

libs(lattice)
dotplot(gen ~ y/n, data=dat, group=trt, auto.key=TRUE,
       main="jansen.carrot",
       xlab="Proportion of carrots infected per block", ylab="Genotype")
```

```r
## Not run:
# Not run because CRAN wants < 5 seconds per example. This is close.
libs(lme4)
# Tentative model. Needs improvement.
m1 <- glmer(cbind(y,n-y) ~ gen*trt + (1|block),
             data=dat, family=binomial)
summary(m1)
# Todo: Why are these results different from Jansen?
# Maybe he used ungrouped bernoulli data? Too slow with 4700 obs

## End(Not run)
```

---

**jansen.strawberry**  
*Ordered disease ratings of strawberry crosses.*

Description

Ordered disease ratings of strawberry crosses.

Usage

```r
data("jansen.strawberry")
```

Format

A data frame with 144 observations on the following 5 variables.

- **male**: male parent
- **female**: female parent
- **block**: block
- **category**: disease damage, C1 < C2 < C3
- **count**: number of plants in each category
Details

In strawberries, red core disease is caused by a fungus, *Phytophthora fragariae*. This experiment evaluated different populations for damage caused by red core disease.

There were 3 male strawberry plants and 4 DIFFERENT female strawberry plants that were crossed to create 12 populations. Note: Jansen labeled the male parents 1,2,3 and the female parents 1,2,3,4. To reduce confusion, this data labels the female parents 5,6,7,8.

The experiment had four blocks with 12 plots each (one for each population). Plots usually had 10 plants, but some plots only had 9 plants. Each plant was assessed for damage from fungus and rated as belonging to category C1, C2, or C3 (increasing damage).

Used with permission of Hans Jansen.

Source


Examples

```r
## Not run:

library(agridat)

data(jansen.strawberry)
dat <- jansen.strawberry
dat <- transform(dat, category=ordered(category, levels=c('C1','C2','C3')))

dtab <- xtabs(count ~ male + female + category, data=dat)
ftable(dtab)
mosaicplot(dtab,
  color=c("lemonchiffon1","lightsalmon1","indianred"),
  main="jansen.strawberry disease ratings",
  xlab="Male parent", ylab="Female parent")

libs(MASS,vcd)
# Friendly suggests a minimal model is [MF][C]
# m1 <- loglm(~ 1*2 + 3, dtab) # Fails, only with devtools
# mosaic(m1)

## End(Not run)
```

---

**Bamboo progeny trial**

Description

Bamboo progeny trial in 2 locations, 3 blocks
Usage

data("jayaraman.bamboo")

Format

A data frame with 216 observations on the following 5 variables.

- loc  location factor
- block block factor
- tree tree factor
- family family factor
- height height, cm

Details

Data from a replicated trial of bamboo at two locations in Kerala, India. Each location had 3 blocks. In each block were 6 families, with 6 trees in each family.

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(jayaraman.bamboo)
dat <- jayaraman.bamboo

# very surprising differences between locations
library(lattice)
bwplot(height ~ family|loc, d2, main="jayaraman.bamboo")
# match Jayarman's anova table 6.3, page 173
# m1 <- aov(height ~ loc+loc:block + family + family:loc + family:loc:block, data=d2)
# anova(m1)

# more modern approach with mixed model, match variance components needed
# for equation 6.9, heritability of the half-sib averages as
m2 <- lme4::lmer(height ~ 1 + (1|loc/block) + (1|family[loc/block], data=d2)
lucid::vc(m2)

## End(Not run)
```
Yields from treatment for mildew control

Description

Yields from treatment for mildew control

Format

A data frame with 38 observations on the following 4 variables.

- **plot**: plot number
- **trt**: treatment factor, 4 levels
- **block**: block factor, 9 levels
- **yield**: grain yield, tons/ha

Details

There were four spray treatments: 0 (none), 1 (early), 2 (late), R (repeated).
Each treatment occurs once between each of the 9 ordered pairs of the other treatments.
The first and last plot are not assigned to a block.

Source


References

Maria Durban, Christine Hackett, Iain Currie. Blocks, Trend and Interference in Field Trials.

Examples

```r
library(agridat)
data(jenkyn.mildew)
dat <- jenkyn.mildew

bwplot(yield ~ trt, dat, main="jenkyn.mildew", xlab="Treatment")

# Residuals from treatment model show obvious spatial trends
m0 <- lm(yield ~ trt, dat)
xyplot(resid(m0)~plot, dat, ylab="Residual",
       main="jenkyn.mildew - treatment model")

# The blocks explain most of the variation
```
m1 <- lm(yield ~ trt + block, dat)
xyplot(resid(m1)~plot, dat, ylab="Residual",
       main="jenkyn.mildew - block model")

john.alpha 

Alpha lattice design of spring oats

Description

Alpha lattice design of spring oats

Format

A data frame with 72 observations on the following 5 variables.

  plot  plot number
  rep   replicate
  block incomplete block
  gen   genotype (variety)
  yield dry matter yield (tonnes/ha)
  row   Row ordinate
  col   Column ordinate

Details

A spring oats trial grown in Craibstone, near Aberdeen. There were 24 varieties in 3 replicates, each consisting of 6 incomplete blocks of 4 plots. Planted in a resolvable alpha design.

Caution: Note that the table on page 146 of John & Williams (1995) is NOT the physical layout. The plots were laid out in a single line.

Source


References


Examples

```r
## Not run:
library(agridat)
data(john.alpha)
dat <- john.alpha

# RCB (no incomplete block)
m0 <- lm(yield ~ 0 + gen + rep, data=dat)

# Block fixed (intra-block analysis) (bottom of table 7.4 in John)
m1 <- lm(yield ~ 0 + gen + rep + rep:block, dat)
anova(m1)

# Block random (combined inter-intra block analysis)
libs(lme4, lucid)
m2 <- lmer(yield ~ 0 + gen + rep + (1|rep:block), dat)
anova(m2)

## Analysis of Variance Table
## Df Sum Sq Mean Sq F value
## gen 24 380.43 15.8513 185.9942
## rep 2 1.57 0.7851 9.2123
vc(m2)
## grp var1 var2 vcov sdcor
## rep:block (Intercept) <NA> 0.06194 0.2489
## Residual <NA> <NA> 0.08523 0.2919

# Variety means. John and Williams table 7.5. Slight, constant
# difference for each method as compared to John and Williams.
means <- data.frame(rcb=coef(m0)[1:24],
                    ib=coef(m1)[1:24],
                    intra=fixef(m2)[1:24])
head(means)
## rcb  ib  intra
## genG01 5.201233 5.268742 5.146433
## genG02 4.552933 4.665389 4.517265
## genG03 3.381800 3.803790 3.537934
## genG04 4.439400 4.728175 4.528828
## genG05 5.103100 5.225708 5.075944
## genG06 4.749067 4.618234 4.575394

libs(lattice)
splom(means, main="john.alpha - means for RCB, IB, Intra-block")
```

# ----------
# asreml4
libs(asreml,lucid)
# Heritability calculation of Piepho & Mohring, Example 1

```r
m3 <- asreml(yield ~ 1 + rep, data=dat, random=~ gen + rep:block)
sg2 <- summary(m3)$varcomp['gen','component'] # .142902

# Average variance of a difference of two adjusted means (BLUP)

p3 <- predict(m3, data=dat, classify="gen", sed=TRUE)
# Matrix of pair-wise SED values, squared
vdiff <- p3$sed^2
# Average variance of two DIFFERENT means (using lower triangular of vdiff)
vblup <- mean(vdiff[lower.tri(vdiff)]) # .05455038

# Note that without sed=TRUE, asreml reports square root of the average variance
# of a difference between the variety means, so the following gives the same value
# predict(m3, data=dat, classify="gen")$avsed ^ 2 # .05455038

# Average variance of a difference of two adjusted means (BLUE)
m4 <- asreml(yield ~ 1 + gen + rep, data=dat, random = ~ rep:block)
p4 <- predict(m4, data=dat, classify="gen", sed=TRUE)
vdiff <- p4$sed^2
vblue <- mean(vdiff[lower.tri(vdiff)]) # .07010875
# Again, could use predict(m4, data=dat, classify="gen")$avsed ^ 2

# H^2 Ad-hoc measure of heritability
sg2 / (sg2 + vblue/2) # .803

# H^2c Similar measure proposed by Cullis.
1-(vblup / (2*sg2)) # .809
```

# ---------
# lme4 to calculate Cullis H2
# https://stackoverflow.com/questions/38697477

```r
libs(lme4)
cov2sed <- function(x){
  # Convert var-cov matrix to SED matrix
  # sed[i,j] = sqrt( x[i,i] + x[j,j]- 2*x[i,j] )
  n <- nrow(x)
  vars <- diag(x)
  sed <- sqrt( matrix(vars, n, n, byrow=TRUE) +
               matrix(vars, n, n, byrow=FALSE) - 2*x )
  diag(sed) <- 0
  return(sed)
}

# Same as asreml model m4. Note 'gen' must be first term
m5blue <- lmer(yield ~ 0 + gen + rep + (1|rep:block), dat)
libs(emmeans)
ls5blue <- emmeans(m5blue, "gen")
```
johnson.blight

johnson.blight

Potato blight due to weather in Prosser, Washington

con <- ls5blue@linfct[,1:24] # contrast matrix for genotypes
# The 'con' matrix is identity diagonal, so we don't need to multiply,
# but do so for a generic approach
# sed5blue <- cov2sed(con)
tmp <- tcrossprod(crossprod(t(con), vcov(m5blue)[1:24,1:24]), con)
sed5blue <- cov2sed(tmp)

# vblue Average variance of difference between genotypes
vblue <- mean(sed5blue[upper.tri(sed5blue)]^2)
vblue # .07010875 matches 'vblue' from asreml

# Now blups
m5blup <- lmer(yield ~ 0 + (1|gen) + rep + (1|rep:block), dat)
# Need lme4::ranef in case ordinal is loaded
re5 <- lme4::ranef(m5blup,condVar=TRUE)
vv1 <- attr(re5$gen,"postVar")
vblup <- 2*mean(vv1) # .0577 not exactly same as 'vblup' above

# H^2 Ad-hoc measure of heritability
sg2 <- c(lme4::VarCorr(m5blup)[["gen"]]) # 0.142902
sg2 / (sg2 + vblue/2) # .803 matches asreml

# H^2c Similar measure proposed by Cullis.
1-(vblup / 2 / sg2) # .809 from asreml, .800 from lme4

# Sommer to calculate Cullis H2
libs(sommer)
m2.ran <- mmer(fixed = yield ~ rep,
random = ~ gen + rep:block,
data = dat)

vc_g <- m2.ran$sigma$gen # genetic variance component
n_g <- n_distinct(dat$gen) # number of genotypes
C22_g <- m2.ran$PevU$gen$yield # Prediction error variance matrix for genotypic BLUPs
trC22_g <- sum(diag(C22_g)) # trace
# Mean variance of a difference between genotypic BLUPs. Smith eqn 26
# I do not see the algebraic reason for this...2
av2 <- 2/n_g * (trC22_g - (sum(C22_g)-trC22_g) / (n_g-1))

### H2 Cullis
1-(av2 / (2 * vc_g)) #0.8091

### End(Not run)
Description

Potato blight due to weather in Prosser, Washington

Format

A data frame with 25 observations on the following 6 variables.

- **year**
- **area** area affected, hectares
- **blight** blight detected, 0/1 numeric
- **rain.am** number of rainy days in April and May
- **rain.ja** number of rainy days in July and August
- **precip.m** precipitation in May when temp > 5C, milimeters

Details

The variable `blight detected` is 1 if `area > 0`.

Source


References

Vinayanand Kandala, *Logistic Regression*

Examples

```r
library(agridat)

data(johnson.blight)
dat <- johnson.blight

# Define indicator for blight in previous year
dat$blight.prev[1] <- 0 # Need this to match the results of Johnson
dat$blight.prev <- factor(dat$blight.prev)
dat$blight <- factor(dat$blight)

# Johnson et al developed two logistic models to predict outbreak of blight

m1 <- glm(blight ~ blight.prev + rain.am + rain.ja, data=dat, family=binomial)
summary(m1)

## Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -11.4699   5.5976  -2.049  0.0405 *
## blight.prev1   3.8796   1.8066   2.148  0.0318 *
```
## rain.am 0.7162 0.3665 1.954 0.0507 .
## rain.ja 0.2587 0.2468 1.048 0.2945
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## (Dispersion parameter for binomial family taken to be 1)

## Null deviance: 34.617 on 24 degrees of freedom
## Residual deviance: 13.703 on 21 degrees of freedom
## AIC: 21.703

m2 <- glm(blight ~ blight.prev + rain.am + precip.m, data=dat, family=binomial)
summary(m2)

## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.5483 3.8070 -1.983 0.0474 *
## blight.prev1 3.5526 1.6061 2.212 0.0270 *
## rain.am 0.6290 0.2763 2.276 0.0228 *
## precip.m -0.0904 0.1144 -0.790 0.4295
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## (Dispersion parameter for binomial family taken to be 1)

## Null deviance: 34.617 on 24 degrees of freedom
## Residual deviance: 14.078 on 21 degrees of freedom
## AIC: 22.078

libs(lattice)
splom(dat[,c("blight","rain.am","rain.ja","precip.m")],
     main="johnson.blight - indicator of blight")

---

johnson.douglasfir  A study of small-plots of old-growth Douglas Fir in Oregon.

---

Description

A study of small-plots of old-growth Douglas Fir in Oregon.

Usage

data("johnson.douglasfir")

Format

A data frame with 1600 observations on the following 3 variables.

row row
jones.corn.uniformity

Uniformity trial of corn.

Description

Uniformity trial of corn in Iowa in 2016.
Usage

data("jones.corn.uniformity")

Format

A data frame with 144 observations on the following 3 variables.

col  column ordinate  
row  row ordinate  
yield yield, bu/ac

Details

This data corresponds to field "ISU.SE" in the paper by Jones.

Field width: 12 columns, 4.6 meters each.

Field length: 12 rows, 3 meters each.

Electronic version provided as an online supplement. The "row" and "col" variables in the supplement have been swapped for the presentation of the data here in order to be more consistent with the figures in the paper.

Used with permission of Marcus Jones.

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(jones.corn.uniformity)
dat <- jones.corn.uniformity
library(desplot)
# Compare to figure 5 of Jones et al.
desplot(dat, yield ~ col*row,
       aspect=(12*4.6)/(12*3),
       main="jones.corn.uniformity")
## End(Not run)
```
kadam.millet.uniformity

Uniformity trial of millet

Description
Uniformity trial of millet in India during 2 years

Usage
data("kadam.millet.uniformity")

Format
A data frame with 240 observations on the following 4 variables.

year year
row row
col column
yield yield, ounces

Details
Uniformity trials conducted during the kharip (monsoon) seasons of 1933 and 1934 at Kundewadi, Niphad, in the district of Nasik, India. Bajari (pearl millet) strain 54 was used.

In 1933:
Field width: 8 plots * 16.5 feet
Field length: 10 plots * 33 feet

In 1934:
Field width: 8 plots * 16.5 feet
Field length: 20 plots * 16.5 feet

Source

References
None.
Examples

```r
## Not run:

library(agridat)

data(kadam.millet.uniformity)
dat <- kadam.millet.uniformity

# similar to Kadam fig 1
libs(desplot)
desplot(dat, yield ~ col*row,
       subset=year==1933,
       flip=TRUE, aspect=(10*33)/(8*16.5), # true aspect
       main="kadam.millet.uniformity 1933")

desplot(dat, yield ~ col*row,
       subset=year==1934,
       flip=TRUE, aspect=(20*16.5)/(8*16.5), # true aspect
       main="kadam.millet.uniformity 1934")

## End(Not run)
```

---

**kalamkar.potato.uniformity**

*Uniformity trial of potatoes*

**Description**

Uniformity trial of potatoes at Saskatchewan, Canada, 1929.

**Usage**

```r
data("kalamkar.potato.uniformity")
```

**Format**

A data frame with 576 observations on the following 3 variables.

- `row` row
- `col` column
- `yield` yield of potatoes per plot
Details

The data is for potato yields in 96 rows, each 132 feet long, with 3 feet between rows. Each row was harvested as six plots, each 22 feet long. Each hill had one seed piece. Hills were spaced 2 feet apart in each row.

Field width: 6 plots * 22 feet = 132 feet
Field length: 96 rows * 3 feet = 288 feet

Source


Examples

```r
## Not run:
library(agridat)
data(kalamkar.potato.uniformity)
dat <- kalamkar.potato.uniformity

# Similar to figure 1 of Kalamkar
libs(desplot)
desplot(dat, yield~col*row,
  flip=TRUE, tick=TRUE, aspect=288/132, # true aspect
  main="kalamkar.potato.uniformity")

## End(Not run)
```

---

**kalamkar.wheat.uniformity**

*Uniformity trial of wheat*

Description

Uniformity trial of wheat at Rothamsted, UK in 1931.

Usage

data("kalamkar.wheat.uniformity")

Format

A data frame with 1280 observations on the following 4 variables.

- row row
- col column
- yield yield, grams/half-meter
- ears ears per half-meter
Details

Plot 18 of the Four Course Rotation Experiment, Great Hoos, at Rothamsted, UK was used. Sown with Yeoman II wheat.
Field width = 16 segments * 0.5 meters = 8 meters.
Field length: 80 rows * 6 inches apart = 40 feet.
The grain yield and number of ears for each half-meter length were recorded. This is quite a small field, only 1/40 acre in size.
Edge rows have higher yields. Row-end units have higher yields than interior units. These border effects are significant. Variation between rows is greater than variation within rows. Negative correlation between rows may indicate competition effects.
For ears, Kalamkar discarded 4 rows from each side and 3 half-meter lengths at each end.
Kalamkar suggested using four parallel half-meter rows as a sampling unit.

Note, the Rothamsted report for 1931, page 57, says: During the year three workers (F. R. Immer, S. H. Justensen and R. J. Kalamkar) have taken up the question of the most efficient use of land in experiments in which an edge row must be discarded...

Source


References

None.

Examples

```r
## Not run:
library(agridat)
data(kalamkar.wheat.uniformity)
dat <- kalamkar.wheat.uniformity
plot(yield ~ ears, dat, main="kalamkar.wheat.uniformity")
# totals match Kalamkar
# sum(dat$yield) # 24112.5
# sum(dat$ears) # 25850
libs(desplot)
desplot(dat, ears ~ col*row,
flp=TRUE, aspect=(80*0.5)/(16*1.64042), # true aspect
main="kalamkar.wheat.uniformity - ears")
desplot(dat, yield ~ col*row,
flp=TRUE, aspect=(80*0.5)/(16*1.64042), # true aspect
main="kalamkar.wheat.uniformity - yield")
```
# ------

libs(asreml,lucid) # asreml4

# Show the negative correlation between rows

dat <- transform(dat,
    rowf=factor(row), colf=factor(col))

dat <- dat[order(dat$rowf, dat$colf),]

m1 = asreml(yield ~ 1, data=dat, resid= ~ ar1(rowf):ar1(colf))

vc(m1)

## effect component std.error z.ratio bound pctch
## rowf:colf!R 81.53 3.525 23 P 0
## rowf:colf!rowf!cor -0.09464 0.0277 -3.4 U 0.1
## rowf:colf!colf!cor 0.2976 0.02629 11 U 0.1

## End(Not run)

---

**kang.maize**  
*Multi-environment trial of maize in Louisianna at 4 locs in 3 years*

**Description**

Maize yields at 4 locs in 3 years in Louisianna.

**Usage**

```r
data("kang.maize")
```

**Format**

- **gen**: genotype, 17 levels
- **env**: environment, 12 levels
- **yield**: yield, tonnes/ha
- **environment**: 13 levels
- **year**: year, 85-87
- **loc**: location, 4 levels

**Details**

Yield trials were conducted at four locations (Alexandria, Baton Rouge, Bossier City, and St. Joseph) in Louisiana during 1985 to 1987. Each loc was planted as RCB design with 4 reps. Mean yields are given in this data.

Used with permission of Dan Gorman.
Source

Examples
```r
## Not run:
library(agridat)
data(kang.maize)
dat <- kang.maize

# Sweep out loc means, then show interaction plot.
libs(reshape2)
mat <- acast(dat, gen~env, value.var='yield')
mat <- sweep(mat, 2, colMeans(mat))
datat <- melt(mat)
names(datat) <- c('gen','env','yield')
libs(lattice)
xyplot(yield~env|gen, data=datat, type='l', group=gen,
   panel=function(x,y,...){
     panel.abline(h=0,col="gray70")
     panel.xyplot(x,y,...)
   },
   ylab="Environment-centered yield",
   main="kang.maize - maize hybrid yields", scales=list(x=list(rot=90)))

# Weather covariates for each environment.
cov <- data.frame(env=c("AL85","AL86","AL87", "BR85","BR86","BR87",
                         "BC85","BC86","BC87", "SJ85","SJ86","SJ87"),
   max=c(30.7,30.2,29.7,31.5,29.4,28.5, 31.9, 30.4,31.7, 32,29.6,28.9),
   rain=c(.2,.34,.22, .2,.36,.61, .2,.43,.2, .36,.41,.22),
   humid=c(82.8,91.1,85.4, 88.1,90.9,88.6, 95.4,90.4,86.7, 95.6,89.5,85))

## End(Not run)
```

---

**kang.peanut**  
*Multi-environment trial of peanuts for 10 genotypes in 15 environments*

**Description**  
Peanut yields for 10 genotypes in 15 environments

**Usage**  
data("kang.peanut")
Format

A data frame with 590 observations on the following 4 variables.

- gen  genotype factor, 10 levels
- rep  replicate factor, 4 levels
- yield yield
- env  environment factor, 15 levels

Details

Florman, Tegua, mf484, mf485, mf487, mf489 have a long crop cycle. The others have a short crop cycle.

This data is also likely used in Casanoves et al 2005, "Evaluation of Multi-environment Trials of Peanut Cultivars", but this appears to be a slightly smaller subset (only 10 genotypes, and perhaps only the years 96, 97, 98, 99). Based on the d.f. in their table 5, it appears that environment E13 was grown in 1998. (5 loc * (4-1) = 15, but the table has 14, and 98-99 had only 3 reps instead of 4 reps.)

Data from National Institute of Agricultural Technology, Argentina.

Source


References


Examples

```r
## Not run:

library(agridat)
data(kang.peanut)
dat <- kang.peanut

# Table 5.1 of Kang et al. (Chapter 5 of Saxton)
libs(reshape2)
Y0 <- acast(dat, env~gen, value.var='yield', fun=mean)
round(Y0,2)

# GGE biplot of Kang, p. 82.
libs(gge)
m1 <- gge(dat, yield~gen*env, scale=FALSE)
biplot(m1, flip=c(1,1), main="kang.peanut - GGE biplot")

# Forkman 2019, fig 2
```
karcher.turfgrass

Turfgrass ratings for different treatments

Description

Turfgrass ratings for different treatments

Format

A data frame with 128 observations on the following 6 variables.

- **week**: week number
- **rep**: blocking factor
- **manage**: management factor, 4 levels
- **nitro**: nitrogen factor, 2 levels
- **rating**: turfgrass rating, 4 ordered levels
- **count**: number of samples for a given rating

Details

Turf color was assessed on a scale of Poor, Average, Good, Excellent.

The data are the number of times that a combination of management style and nitrogen level received a particular rating across four replicates and four sampling weeks. The eight treatments were in a completely randomized design.

Nitrogen level 1 is 2.5 g/m^2, level 2 is 5 g/m^2.

Management 1 = N applied with no supplemental water injection.
M2 = surface applied with supplemental water injection.
M3 = nitrogen injected 7.6 cm deep
M4 = nitrogen injected 12.7 cm deep.

Source

Examples

```r
library(agridat)

data(karcher.turfgrass)
adat <- karcher.turfgrass

dat$rating <- ordered(dat$rating, levels=c('Poor', 'Average', 'Good', 'Excellent'))

ftable(xtabs(~manage+nitro+rating, dat)) # Table 6.19 of Schabenberger

# Probably would choose management M3, nitro N2
mosaicplot(xtabs(count ~ manage + rating + nitro, dat),
            shade=TRUE, dir=c('h','v','h'),
            main="karcher.turfgrass - turfgrass ratings")

# Multinomial logistic model. Probit Ordered Logistic Regression.
libs(MASS)
m1 <- polr(rating ~ nitro*manage + week, dat, weights=count, Hess=TRUE, method='logistic')
summary(m1)

# Try to match the "predicted marginal probability distribution" of
# Schabenberger table 6.20. He does't define "marginal".
# Are the interaction terms included before aggregation?
# Are 'margins' calculated before/after back-transforming?
# At what level is the covariate 'week' included?

# Here is what Schabenberger presents:
## M1 M2 M3 M4 | N1 N2
## Poor .668 .827 .001 .004 | .279 .020
## Avg .330 .172 .297 .525 | .712 .826
## Good .002 .001 .695 .008 | .008 .153
## Exc .000 .000 .007 .003 | .001 .001

## We use week=3.5, include interactions, then average
newd <- expand.grid(manage=levels(dat$manage), nitro=levels(dat$nitro), week=3.5)
newd <- cbind(newd, predict(m1, newdata=newd, type='probs')) # probs
print(aggregate( ~ manage, data=newd, mean), digits=2)
```

### kayad.alfalfa

Yield monitor data for 4 cuttings of alfalfa in Saudi Arabia.

**Description**

Yield monitor data for 4 cuttings of alfalfa in Saudi Arabia.
Usage

data("kayad.alfalfa")

Format

A data frame with 8628 observations on the following 4 variables.

- `harvest`  harvest number
- `lat`     latitude
- `long`    longitude
- `yield`   yield, tons/ha

Details

Data was collected from a 23.5 ha field of alfalfa in Saudia Arabia. The field was harvested four consecutive times (H8 = 5 Dec 2013, H9 = 16 Feb 2014, H10 = 2 Apr 2014, H11 = 6 May 2014). Data were collected using a geo-referenced yield monitor. Supporting information contains yield monitor data for 4 hay harvests on a center-pivot field.

# TODO: Normalize the yields for each harvest, then average together # to create a productivity map. Two ways to normalize: # Normalize to 0-100: 
((mapValue - min) * 100) / (max - min) #
Standardize: 
((mapValue - mean) / stdev) * 100

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(kayad.alfalfa)
dat <- kayad.alfalfa

# match Kayad table 1 stats
libs(dplyr)
dat <- group_by(dat, harvest)
summarize(dat, min=min(yield), max=max(yield),
          mean=mean(yield), stdev=sd(yield), var=var(yield))

# Figure 4 of Kayad
libs(latticeExtra)
catcols <- c("cccccc","ff0000","ffffff","55ff00","0070ff","c500ff","73004c")
levelplot(yield ~ long*lat | harvest, dat,
          aspect=1, at = c(0,2,3,4,5,6,7,10), col.regions=catcols,
```
keen.potatodamage

Damage to potato tubers from lifting rods.

Description

Damage to potato tubers from lifting rods.

Usage

data("keen.potatodamage")

Format

A data frame with 1152 observations on the following 6 variables.

energy  energy factor
weight  weight class
gen   genotype/variety factor
rod    rod factor
damage damage category
count  count of tubers in each combination of categories

Details

Experiments performed at Wageningen, Netherlands.

Potatoes can be damaged by the lifter. In this experiment, eight types of lifting rod were compared. Two energy levels, six genotypes/varieties and three weight classes were used. Most combinations of treatments involved about 20 potato tubers. Tubers were rated as undamaged (D1) to severely damaged (D4).

The main interest is in differences between rods, and not in interactions. The other factors (besides rod) were introduced to create variety in experimental conditions and are not of interest.

Keen and Engle estimated the following rod effects.

# Rod: 1 2 3 4 5 6 7 8
# Effect: 0 -1.26 -0.42 0.55 -1.50 -1.85 -1.76 -2.09

Used with permission of Bas Engel.
### Source


### References


### Examples

```r
## Not run:
library(agridat)

data(keen.potatodamage)
dat <- keen.potatodamage

# Energy E1, Rod R4, Weight W1 have higher proportions of severe damage
# Rod 8 has the least damage
d2 <- xtabs(count~energy+rod+gen+weight+damage, data=dat)
mosaicplot(d2, color=c("lemonchiffon1","moccasin","lightsalmon1","indianred"),
          xlab="Energy / Genotype", ylab="Rod / Weight", main="keen.potatodamage")

# Not run because CRAN prefers examples less than 5 seconds.
libs(ordinal)
# Note, the clmm2 function can have only 1 random term. Results are
# similar to Keen & Engle, but necessarily different (they had multiple
# random terms).
m1 <- clmm2(damage ~ rod + energy + gen + weight, data=dat,
            weights=count, random=rod:energy, link="probit")
round(coef(m1)[4:10],2)
## rodR2  rodR3  rodR4  rodR5  rodR6  rodR7  rodR8
## -1.19 -0.41  0.50 -1.46 -1.73 -1.67 -1.99

# Alternative
# m2 <- clmm(damage ~ rod + energy + gen + weight +
#   (1|rod:energy), data=dat, weights=count, link='probit')

## End(Not run)
```

### Description

Format

A data frame with 196 observations on the following 3 variables.

row  row
col  column
yield  grain yield, kg

Details

Each plot is 5 feet wide, 14 feet long.
Field width: 7 plots * 14 feet = 98 feet
Field length: 28 plots * 5 feet = 140 feet

Source


References


Examples

## Not run:

library(agridat)
data(kempton.barley.uniformity)
dat <- kempton.barley.uniformity

libs(desplot)
desplot(dat, yield~col*row,
       aspect=140/98, tick=TRUE, # true aspect
       main="kempton.barley.uniformity")

# Kempton estimated auto-regression coefficients b1=0.10, b2=0.91
dat <- transform(dat, xf = factor(col), yf=factor(row))

# --------

libs(asreml,lucid) # asreml4
dat <- transform(dat, xf = factor(col), yf=factor(row))
m1 <- asreml(yield ~ 1, data=dat, resid = ~ar1(xf):ar1(yf))

# vc(m1)
## effect component std.error z.ratio bound
## xf:yf!R 0.1044 0.02197 4.7 P 0
## xf:yf!xf!cor 0.2458 0.07484 3.3 U 0
## xf:yf!yf!cor 0.8186 0.03821 21 U 0

# asreml estimates auto-regression correlations of 0.25, 0.82
# Kempton estimated auto-regression coefficients b1=0.10, b2=0.91

if(0){
  # Kempton defines 4 blocks, randomly assigns variety codes 1-49 in each block, fits
  # RCB model, computes mean squares for variety and residual. Repeat 40 times.
  # Kempton's estimate: variety = 1032, residual = 1013
  # Our estimate: variety = 825, residual = 1080
  fitfun <- function(dat){
    dat <- transform(dat, block=factor(ceiling(row/7)),
                      gen=factor(c(sample(1:49),sample(1:49),sample(1:49),sample(1:49))))
    m2 <- lm(yield*100 ~ block + gen, dat)
    anova(m2)[2:3,
             'Mean Sq']
  }
  set.seed(251)
  out <- replicate(50, fitfun(dat))
  rowMeans(out) # 826 1079
}

## End(Not run)

---

kempton.competition  Sugar beet trial with competition effects

Description

Yield of sugar beets for 36 varieties in a 3-rep RCB experiment. Competition effects are present.

Format

A data frame with 108 observations on the following 5 variables.

- gen  genotype, 36 levels
- rep  rep, 3 levels
- row  row
- col  column
- yield yield, kg/plot
Details

Entries are grown in 12m rows, 0.5m apart. Guard rows were grown alongside replicate boundaries, but yields of these plots are not included.

Source


Examples

```r
## Not run:
library(agridat)
data(kempton.competition)
dat <- kempton.competition

# Raw means in Kempton table 2
dat$yield, dat$gen, mean),2)

# Fixed genotype effects, random rep effects,
# Autocorrelation of neighboring plots within the same rep, phi = -0.22
ml1 <- lme(yield ~ -1+gen, random=-1|rep, data=dat,
corr=corAR1(form=-col|rep))

# Lag 1 autocorrelation is negative--evidence of competition
plot(ACF(ml1), alpha=.05, grid=TRUE, main="kempton.competition",
     ylab="Autocorrelation between neighborng plots")

# Genotype effects
round(fixef(ml1),2)

# Variance of yield increases with yield
plot(ml1, main="kempton.competition")

## End(Not run)
```

Description

Row-column experiment of wheat, 35 genotypes, 2 reps.
**Format**

A data frame with 68 observations on the following 5 variables.

- `rep`: replicate factor, 2 levels
- `row`: row
- `col`: column
- `gen`: genotype factor, 35 levels
- `yield`: yield

**Details**

Included to illustrate REML analysis of a row-column design.

**Source**


**Examples**

```r
## Not run:

library(agridat)
data(kempton.rowcol)
dat <- kempton.rowcol
dat <- transform(dat, rowf=factor(row), colf=factor(col))

libs(desplot)
desplot(dat, yield~col*row|rep, 
   num=gen, out1=rep, # unknown aspect
   main="kempton.rowcol")

# Use "-1" so that the vcov matrix doesn't include intercept
libs(lme4)
m1 <- lmer(yield ~ -1 + gen + rep + (1|rep:rowf) + (1|rep:colf), data=dat)

# Variance components match Kempton.
print(m1, corr=FALSE)

# Standard error of difference for genotypes. Kempton page 62, bottom.
covs <- as.matrix(vcov(m1)[1:35, 1:35])
vars <- diag(covs)
vdiff <- outer(vars, vars, "+") - 2 * covs
sed <- sqrt(vdiff[upper.tri(vdiff)])
min(sed) # Minimum SED
mean(sed) # Average SED
max(sed) # Maximum SED
```
Description

Yields for a Slate Hall Farm 1976 spring wheat trial.

Format

A data frame with 150 observations on the following 5 variables.

rep  rep, 6 levels
row  row
col  column
gen  genotype, 25 levels
yield  yield (grams/plot)

Details

The trial was a balanced lattice with 25 varieties in 6 replicates, 10 ranges of 15 columns. The plot size was 1.5 meters by 4 meters. Each row within a rep is an (incomplete) block.

Field width: 15 columns * 1.5m = 22.5m
Field length: 10 ranges * 4m = 40m

Source


References

Examples

```r
## Not run:
library(agridat)
data(kempton.slatehall)
dat <- kempton.slatehall

# Besag 1993 figure 4.1 (left panel)
libs(desplot)
grays <- colorRampPalette(c("#d9d9d9","#252525"))
desplot(dat, yield ~ col * row,
    aspect=40/22.5, # true aspect
    num=gen, out1=rep, col.regions=grays, # unknown aspect
    main="kempton.slatehall - spring wheat yields")

# Incomplete block model of Gilmour et al 1995
libs(lme4, lucid)
dat <- transform(dat, xf=factor(col), yf=factor(row))
m1 <- lmer(yield ~ gen + (1|rep) + (1|rep:yf) + (1|rep:xf), data=dat)
vc(m1)

## groups name variance stddev
## rep:xf (Intercept) 14810 121.7
## rep:yf (Intercept) 15600 124.9
## rep (Intercept) 4262 65.29
## Residual 8062 89.79

# asreml3 & asreml4
libs(asreml,lucid)

# Incomplete block model of Gilmour et al 1995
dat <- transform(dat, xf=factor(col), yf=factor(row))
m2 <- asreml(yield ~ gen, random = ~ rep/(xf+yf), data=dat)
vc(m2)

## effect component std.error z.ratio constr
## rep!rep.var 4262 6890 0.62 pos
## rep:xf!rep.var 14810 4865 3 pos
## rep:yf!rep.var 15600 5091 3.1 pos
## R!variance 8062 1340 6 pos

# Table 4
# asreml3
# predict(m2, data=dat, classify="gen")$predictions$pvals
# asreml4
# predict(m2, data=dat, classify="gen")$pvals
```
## Description
Repeated measurements of the weights of calves from a trial on the control of intestinal parasites.

## Usage
```r
data("kenward.cattle")
```

## Format
A data frame with 660 observations on the following 4 variables.

- `animal` animal factor
- `trt` treatment factor, A or B
- `day` day, numeric, 0-133
- `weight` bodyweight, kg

## Details
Grazing cattle can ingest larvae, which deprives the host animal of nutrients and weakens the immune system, affecting the growth of the animal.

Two treatments A and B were applied randomly to 60 animals (30 each in two groups) to control the disease.

Each animal was weighed 11 times at two-week intervals (one week between the final two measurements).

Is there a difference in treatments, and when does that difference first become manifest?

## Source

## References
Examples

## Not run:

```r
library(agridat)
data(kenward.cattle)
dat <- kenward.cattle

# Profile plots
libs(lattice)
foo1 <- xyplot(weight~day|trt, data=dat, type='l', group=animal,
               xlab="Day", ylab="Animal weight", main="kenward.cattle")
print(foo1)

# lme4. Fixed treatment intercepts, treatment polynomial trend.
# Random deviation for each animal
libs(lme4)
m1a <- lmer(weight ~ trt*poly(day, 4) + (1|animal), data=dat,
             REML = FALSE)
# Change separate polynomials into common polynomial
m1b <- lmer(weight ~ trt + poly(day, 4) + (1|animal), data=dat,
             REML = FALSE)
# Drop treatment differences
m1c <- lmer(weight ~ poly(day, 4) + (1|animal), data=dat,
             REML = FALSE)
anova(m1a, m1b, m1c) # Significant differences between trt polynomials

# Overlay polynomial predictions on plot
libs(latticeExtra)
dat$pred <- predict(m1a, re.form=NA)
foo1 + xyplot(pred ~ day|trt, data=dat,
              lwd=2, col="black", type='l')

# A Kenward-Roger Approximation and Parametric Bootstrap
# libs(pbkrtest)
# KRmodcomp(m1b, m1c) # Non-signif
# Model comparison of nested models using parametric bootstrap methods
# PBmodcomp(m1b, m1c, nsim=500)
## Parametric bootstrap test; time: 13.20 sec; samples: 500 extremes: 326;
## large : weight ~ trt + poly(day, 4) + (1 | animal)
## small : weight ~ poly(day, 4) + (1 | animal)
## stat df p.value
## LRT 0.2047 1 0.6509
## PBtest 0.2047 0.6527
```

# asreml3 & asreml4

# ASREML approach to model. Not final by any means.
# Maybe a spline curve for each treatment, plus random deviations for each time


```r
libs(asreml)

m1 <- asreml(weight ~ 1 + lin(day) + # overall line
trt + trt:lin(day), # different line for each treatment
data=dat,
random = ~ spl(day) + # overall spline
trt:spl(day) + # different spline for each treatment
develop(day) + trt:dev(day) ) # non-spline deviation at each time*trt

p1 <- predict(m1, data=dat, classify="trt:day")
p1 <- p1$pvals

foo2 <- xyplot(predicted.value ~ day|trt, p1, type='l', lwd=2, lty=1, col="black")

libs(latticeExtra)
print(foo1 + foo2)

# Not much evidence for treatment differences

# wald(m1)
## Df Sum of Sq Wald statistic Pr(Chisq)
## (Intercept) 1 37128459 139060 <2e-16 ***
## trt 1 455 2 0.1917
## lin(day) 1 570798 2138 <2e-16 ***
## trt:lin(day) 1 283 1 0.3031
## residual (MS) 267

# vc(m1)
## effect component std.error z.ratio constr
## spl(day) 25.29 24.09 1 pos
## dev(day) 1.902 4.923 0.39 pos
## trt:spl(day)|trt.var 0.0000003 0.0000002 18 bnd
## trt:dev(day)|trt.var 0.0000003 0.0000002 18 bnd
## R!variance 267 14.84 18 pos

## End(Not run)
```

### kerr.sugarcane.uniformity

**Uniformity trials of sugarcane, 4 fields**

---

**Description**

Uniformity trials of sugarcane, 4 fields

**Usage**

```r
data("kerr.sugarcane.uniformity")
```
**Format**

A data frame with 564 observations on the following 4 variables.

- row  row
- col  column
- yield  yield, pounds per plot
- trial  trial number

**Details**

Experiment conducted at the Sugar Experiment Station, Brisbane, Queensland, Australia in 1937. Four trials were harvested, each 12 plots by 12 plots, each plot 19 feet by 19 feet (one field used 18-foot plots).

- Trial 1 is plant cane.
- Trial 2 is ratoon cane.
- Trial 3 plant cane, irrigated.
- Trial 4 is ratoon cane, irrigated.

Field length: 12 plots * 19 feet = 228 feet.
Field width: 12 plots * 19 feet = 228 feet.

**Source**


**References**

None

**Examples**

```r
## Not run:
library(agridat)
data(kerr.sugarcane.uniformity)
dat <- kerr.sugarcane.uniformity

# match Kerr figure 4
libs(desplot)
desplot(dat, yield ~ col*row|trial,
       flip=TRUE, aspect=1, # true aspect
       main="kerr.sugarcane.uniformity")

# CV matches Kerr table 2, page 768
# aggregate(yield ~ trial, dat, FUN= function(x) round(100*sd(x)/mean(x),2))
## trial yield
## 1  T1  7.95
## 2  T2  9.30
```
### Uniformity trial of rice

#### Description

#### Usage
```r
data("khin.rice.uniformity")
```

#### Format
A data frame with 1080 observations on the following 3 variables.

- `row`
- `col`
- `yield`

#### Details
A uniformity trial of rice. Conducted at the Mudon Agricultural Station, Burma, in 1947-48. Basic plots were 3 feet square.

- Field width: 30 plots * 3 feet.
- Field length: 36 plots * 3 feet.

#### Source
Khin, San. 1950. Investigation into the relative costs of rice experiments based on the efficiency of designs. Dissertation: Imperial College of Tropical Agriculture (ICTA). Appendix XV. https://hdl.handle.net/2139/42422

#### References
None.
## Description

Uniformity trial of oats at Nebraska in 1916.

## Usage

```r
library(agridat)

data(khinn.rice.uniformity)

dat <- khinn.rice.uniformity

libs(desplot)

desplot(dat, yield ~ col*row, 
   flip=TRUE, 
   main="khinn.rice.uniformity", 
   aspect=(36*3)/(30*3)) # true aspect
```

## Format

A data frame with 207 observations on the following 3 variables.

- `row`  row
- `col`  column
- `yield`  yield bu/ac

## Details

Experiment conducted in 1916. Crop was Kerson oats. Each plot covered 1/30th acre. Oats were drilled in plats 66 inches wide by 16 rods long. The drill was 66 inches wide. Plats were separated by a space of 16 inches between outside drill rows. The source document includes three photographs of the field.

1 acre = 43560 sq feet

1/30 acre = 1452 sq feet = 16 rods * 16.5 ft/rod * 5.5 ft

Field width: 3 plats * 16 rods/plat * 16.5 ft/rod = 792 feet

Field length: 69 plats * 5.5 ft + 68 gaps * 1.33 feet = 469 feet
Source
Kiesselbach, Theodore A. (1917). Studies Concerning the Elimination of Experimental Error in
Comparative Crop Tests. University of Nebraska Agricultural Experiment Station Research Bulletin
https://digitalcommons.unl.edu/extensionhist/430/

References
None.

Examples
```r
## Not run:
library(agridat)

data(kiesselbach.oats.uniformity)
dat <- kiesselbach.oats.uniformity

range(dat$yield) # 56.7 92.8 match Kiesselbach p 64.

libs(desplot)
desplot(dat, yield ~ col*row,  
tick=TRUE, flip=TRUE, aspect=792/469, # true aspect
main="kiesselbach.oats.uniformity")

## End(Not run)
```

---

**kreusler.maize**

**Growth of maize plants in Germany during 1875-1878**

Description
Growth of maize plants in Germany during 1875-1878.

Usage
data("kreusler.maize")

Format
A data frame with 165 observations on the following 17 variables.

- gen genotype
- year year
- date calendar date
- raindays number of days of rain per week (zahl der regenstage)
Details

Experiments were performed at Poppelsdorf, Germany (near Bonn) during the years 1875 to 1878. Observations were collected weekly throughout the growing season.

Five varieties were grown in 1875. Two in 1876, and one in 1877 and 1878.

The plants were selected by eye as representative, with the number of plants chosen decreasing during the growing season. For example, the dry-weight data was based on the following number of plants:

In 1875 the number sampled began at 20 and dropped to 10.
In 1876 the number sampled began at 45 and dropped to 24.
In 1877 the number sampled began at 90 and dropped to 36.
In 1878 the number sampled began at 120 and dropped to 40.

Most of the observations included fresh weight and dry weight of entire plants, along with leaf area, date of inflorescence, fertilization, and kernel development.

The data of Hornberger 71 are the same as Kreusler/Hornberger, but more complete.

The temperature data was originally given in degrees Reaumur in 1875 and 1876, and degrees Celsius in 1877 and 1878. All temperatures in this data are degrees Celsius. Note: deg C = 1.25 deg R. Briggs, Kidd & West (1920) give all temperature in Celsius.

Source

The 1875-1876 data are from:


The 1877 data are from:


The 1878 data are from:

Dry plant weight and leaf area for all genotypes and years are repeated by:


**References**


**Examples**

data(kreusler.maize)
dat <- kreusler.maize
dat$date2 <- as.Date(dat$date, "%d %b %Y")
dat$doy <- as.numeric(strftime(dat$date2, format="%j"))

# Hunt & Evans Fig 2a
libs(lattice)
xyplot(log10(plantweight)~doy|factor(year), data=dat, group=gen,
type=c("p","smooth"), span=.4, as.table=TRUE,
xlab="Day of year", main="kreusler.maize - growth of maize",
auto.key=list(columns=5))

## Not run:

# Hunt & Evans Fig 2b
xyplot(log10(plantweight)~doy|gen, data=dat, group=factor(year),
type=c("p","smooth"), span=.5, as.table=TRUE,
xlab="Day of year",
auto.key=list(columns=4))

# Hunt & Evans Fig 3a
xyplot(log10(leafarea)~doy|factor(year), data=dat, group=gen,
type=c("p","smooth"), span=.5, as.table=TRUE,
xlab="Day of year",
auto.key=list(columns=5))

# Hunt & Evans Fig 3a
xyplot(log10(leafarea)~doy|gen, data=dat, group=factor(year),
type=c("p","smooth"), span=.5, as.table=TRUE,
xlab="Day of year",
auto.key=list(columns=4))

# All traits

xyplot(raindays~doy|factor(year), data=dat, group=gen,
type="l", auto.key=list(columns=5), as.table=TRUE, layout=c(1,4))
xyplot(rain~doy|factor(year), data=dat, group=gen,
Uniformity trial of barley

**Description**

Uniformity trial of barley conducted in Denmark, 1905.

**Usage**

```r
data("kristensen.barley.uniformity")
```

**Format**

A data frame with 718 observations on the following 3 variables.

- **row** row
- **col** column
- **yield** yield, hectograms/plot
Details

Experiment conducted in 1905 at Askov, Denmark. Harvested plot size was 10 x 14 ‘alen’, 6.24 x 8.79 meters. The soil was uniform, but an attack of mildew spread from an adjacent field. Yield is measured in hectograms/plot for straw and grain together. (Page 468).

Orientation of the plots dimensions is not clear from the text, but the aspect used in the example below aligns well with Kristensen figure 1.

Field width: 22 plots * 8.79 m
Field length: 11 plots * 6.24 m

Notes from Kristensen: Fig 5 is a 3x3 moving average, Fig 6 is deviation from the trend, Fig 7 is the field average added to the deviation. Fig 13 is another uniformity trial of barley in 1924, Fig 14 is a uniformity trial of oats in 1924.

Source


References


Examples

```r
## Not run:

library(agridat)
data(kristensen.barley.uniformity)
dat <- kristensen.barley.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE, aspect=(11*6.24)/(22*8.79),
       main="kristensen.barley.uniformity")

## End(Not run)
```

---

**Description**

Uniformity trial of sorghum in India, 3 years on the same plots 1930-1932.
Usage

```r
data("kulkarni.sorghum.uniformity")
```

Format

A data frame with 480 observations on the following 4 variables.

```r
row row
col column
yield grain yield, tolas per plot
year year
```

Details

The experiment was conducted in the Sholapur district in India for three consecutive years in 1930-1932.

One acre of land (290 ft x 150 ft) was chosen in the midst of a bigger area (plot 13 on the Mohol Plot) for sowing to sorghum. It was harvested in plots of 1/160 acre (72 ft 6 in x 3 ft 9 in) each containing three rows of plants 15 in. apart. The 160 plots were arranged in forty rows of four columns, and the yields were measured in tolas. The plot division was kept intact for three years, and the yields of the 160 plots are available for three consecutive harvests. The original data are given in Appendix I.

Field width: 4 plots * 72.5 feet = 290 feet
Field length: 40 plots * 3.75 feet = 150 feet

Conclusions: "Thus, highly narrow strips of plots (length much greater than breadth) lead to greater precision than plots of same area but much wider and not so narrow."

Correlation of plots from year to years was low.

Source


References

None.

Examples

```r
## Not run:
library(agridat)
data(kulkarni.sorghum.uniformity)
dat <- kulkarni.sorghum.uniformity
```
# lambert.soiltemp

Average monthly soil temperature near Zurich

Description

Average monthly soil temperature near Zurich, at seven depths, averaged over four years.

Format

A data frame with 84 observations on the following 3 variables.

- month
- depth
- temp

Details

This is one of the earliest time series in scientific literature.

These data show the monthly soil temperature near Zurich, averaged over four years (beginning in 1762), at 7 different depths.

The temperature measurements are related to the 'du Crest' scale. (The measurements do not seem to be exactly according to the du Crest scale. If you can read German, use the Google books link to see if you can figure out why.) Even the scale on Lambert’s own graph doesn’t match the data.

Greater depths show less variation and a greater lag in temperature responsiveness to the air temperature.

This data also appears in Pedometrics, issue 23, December 2007. But, the formula for converting the temperature does not make sense and the data in Table 1 do not directly match the corresponding figure.

---

```r
# match means on page 462
# tapply(dat$yield, dat$year, mean)
# 1930 1931 1932
# 116.2875 67.2250 126.3688

libs(reshape2)
libs(lattice)
dmat <- acast(dat, row+col ~ year, value.var="yield")
splom(dmat, main="kulkarni.sorghum.uniformity")
cor(dmat)

libs(desplot)
desplot(dat, yield ~ col*row|year,
        flip=TRUE, aspect=150/290,
        main="kulkarni.sorghum.uniformity")

## End(Not run)
```
**lander.multi.uniformity**

Uniformity trials of wheat and chari, 4 years on the same land.

**Description**
Uniformity trials of wheat and chari, 4 years on the same land, in India.

**Usage**
```
data("lander.multi.uniformity")
```

**Format**
A data frame with 780 observations on the following 5 variables.
```
  row row
  col column
  yield yield, maunds per plot
  year year
  crop crop
```
Details

Note, "chari" in this paper is *Andropogon Sorghum*, and "wheat" is *Triticum vulgare*.

Uniformity trials carried out at Rawalpindi, India.

The area consisted of 5 fields (D4,D5,D6,D7,D8), each 5 acres in size. Each of these 5 fields was divided into three sub-divisions A, B, C, by means of two strong bunds each 5 feet wide. These 3 sub-divisions were divided into 5 blocks, each consisting of 13 experimental plots with 14 non-experiment strips 5 feet wide separating the plots from the other. The dimensions of the plot were 207 ft 5 in by 19 ft 1 in.

The same land was used for 4 consecutive crops. The first crop was wheat, followed by chari/sorghum, followed by wheat 2 times.

Field width: 207.42 * 5 plots = 1037.1 feet
Field length: (19.08+5)*39 rows = 939.12 feet

Conclusions: It is evident, therefore, that soil heterogeneity as revealed by any one crop cannot be a true index of the subsequent behavior of that area with respect to other crops. Even the same crop raised in different seasons has not shown any constancy as regards soil heterogeneity.

Source


References

None

Examples

```r
## Not run:
library(agritdat)
data(lander.multi.uniformity)
dat <- lander.multi.uniformity

# Yearly means, similar to Lander table 7
## filter(dat)
## 1 1929 18.1
## 2 1930 58.3
## 3 1931 22.8
## 4 1932 14.1

# heatmaps for all years
libs(desplot)
dat$year <- factor(dat$year)
desplot(dat, yield ~ col*row|year,
       flip=TRUE, aspect=(1037.1/939.12),
       main="lander.multi.uniformity")

## End(Not run)
```
Description

Yield monitor data for a corn field in Argentina with variable nitrogen.

Usage

data("lasrosas.corn")

Format

A data frame with 3443 observations on the following 8 variables.

year  year, 1999 or 2001
lat   latitude
long  longitude
yield yield, quintals/ha
nitro nitrogen fertilizer, kg/ha
topo  topographic factor
bv    brightness value (proxy for low organic matter content)
rep   rep factor
nf    nitrogen as a factor, N0-N4

Details

Corn yield and nitrogen fertilizer treatment with field characteristics for the Las Rosas farm, Rio Cuarto, Cordoba, Argentina.

Data has 6 nitro treatments, 3 reps, in strips.

Data collected using yield monitor, for harvests in 1999 and 2001.

The points within each long strip have been averaged so that the distance between points _within_ a strip is the same as the distance _between_ strips (9.8 meters).

The topographic factor a factor with levels W = West slope, HT = Hilltop, E = East slope, LO = Low East.

The 'rep' factor in this data was added by hand and did not appear in the original data.

Slightly different levels of nitrogen were used in the two years, so the nitrogen factor 'nf' was created to have common levels across years.

Published descriptions of the data describe the experiment design as having randomized nitrogen treatments. The nitrogen treatments were randomized within one rep, but the same randomization was used in the other two reps.

Anselin et al. used corn grain price of $6.85/quintal and nitrogen cost of $0.4348/kg.
The corners of the field in 1999 are: https://www.google.com/maps/place/-33.0501258,-63.8488636
https://www.google.com/maps/place/-33.0522963,-63.84181819

Anselin et al. found a significant response to nitrogen for slope. However, Bongiovanni and Lowenberg-DeBoer (2002) found that slope position was NOT significant in 2001.

Used with permission of the ASU GeoDa Center.

Source

The Las Rosas data files were obtained from https://geodacenter.asu.edu/sdata and converted from ESRI shape files to a flat data.frame.

References


Examples

```r
## Not run:

library(agridat)
data(lasrosas.corn)
dat <- lasrosas.corn

data(lasrosas.corn)
dat <- lasrosas.corn

# yield map
libs(lattice,latticeExtra) # for panel.levelplot.points
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(yield ~ long*lat|factor(year), data=dat,
main="lasrosas.corn grain yield", xlab="Longitude", ylab="Latitude",
scales=list(alternating=FALSE),
prepanel = prepanel.default.xyplot,
panel = panel.levelplot.points,
type = c("p", "g"), aspect = "iso", col.regions=redblue)

d1 <- subset(dat, year==1999)

# Experiment design
xyplot(lat~long, data=d1, col=as.numeric(as.factor(d1$nitro)), pch=d1$topo,
```
A quadratic response to nitrogen is suggested

```r
xyplot(yield~nitro|topo, data=d1, type=c('p','smooth'), layout=c(4,1),
main="lasrosas.corn yield by topographic zone 1999")
```

# Full-field quadratic response to nitrogen. Similar to Bongiovanni 2000, # table 1.

```r
m1 <- lm(yield ~ 1 + nitro + I(nitro^2), data=d1, subset=year==1999)
coef(m1)
```

## End(Not run)

---

**lavoranti.eucalyptus**  
*Height of Eucalyptus trees in southern Brazil*

### Description

Height of Eucalyptus trees in southern Brazil

### Format

A data frame with 490 observations on the following 4 variables.

- **gen**: genotype (progeny) factor
- **origin**: origin of progeny
- **loc**: location
- **height**: height, meters

### Details

The genotypes originated from three different locations in Queensland, Australia, and were tested in southern Brazil. The experiment was conducted as a randomized complete block design with 6 plants per plot and 10 blocks. Mean tree height is reported.

The testing locations are described in the following table:

<table>
<thead>
<tr>
<th>Loc</th>
<th>City</th>
<th>Lat (S)</th>
<th>Long (W)</th>
<th>Altitude</th>
<th>Avg min temp</th>
<th>Avg max temp</th>
<th>Avg temp (C)</th>
<th>Precip</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Barra Ribeiro, RS</td>
<td>30.33</td>
<td>51.23</td>
<td>30</td>
<td>9</td>
<td>25</td>
<td>19</td>
<td>1400</td>
</tr>
<tr>
<td>L2</td>
<td>Telemaco Borba, PR</td>
<td>24.25</td>
<td>20.48</td>
<td>850</td>
<td>11</td>
<td>26</td>
<td>19</td>
<td>1480</td>
</tr>
<tr>
<td>L3</td>
<td>Boa Experanca de Sul, SP</td>
<td>21.95</td>
<td>48.53</td>
<td>540</td>
<td>15</td>
<td>23</td>
<td>21</td>
<td>1300</td>
</tr>
<tr>
<td>L4</td>
<td>Guanhaes, MG</td>
<td>18.66</td>
<td>43</td>
<td>900</td>
<td>14</td>
<td>24</td>
<td>19</td>
<td>1600</td>
</tr>
<tr>
<td>L5</td>
<td>Ipatinga, MG</td>
<td>19.25</td>
<td>42.33</td>
<td>250</td>
<td>15</td>
<td>24</td>
<td>22</td>
<td>1250</td>
</tr>
<tr>
<td>L6</td>
<td>Aracruz, ES</td>
<td>19.8</td>
<td>40.28</td>
<td>50</td>
<td>15</td>
<td>26</td>
<td>24</td>
<td>1360</td>
</tr>
<tr>
<td>L7</td>
<td>Cacapva, SP</td>
<td>23.05</td>
<td>45.76</td>
<td>650</td>
<td>14</td>
<td>24</td>
<td>20</td>
<td>1260</td>
</tr>
</tbody>
</table>

Arciniegas-Alarcon (2010) used the 'Ravenshoe' subset of the data to illustrate imputation of miss-
ing values.

Source


References


Examples

```r
## Not run:
# Arciniegas-Alarcon et al use SVD and regression to estimate missing values.
# Partition the matrix X as a missing value xm, row vector xr1, column
# vector xc1, and submatrix X11
# X = [ xm  xr1 ]
#    [ xc1  X11 ] and let X11 = UDV'.
# Estimate the missing value xm = xr1 V D^(-1) U xc1

data(lavoranti.eucalyptus)
dat <- lavoranti.eucalyptus

libs(lattice)
levelplot(height~loc*gen, dat, main="lavoranti.eucalyptus - GxE heatmap")

dat <- droplevels(subset(dat, origin=="Ravenshoe"))
libs(reshape2)
dat <- acast(dat, gen~loc, value.var='height')

dat[,1,1] <- NA
x11 <- dat[-1,][,-1]
X11.svd <- svd(x11)
xc1 <- dat[-1,][,1]
xr1 <- dat[-1,][-1,]
xm <- xr1
xm # = 18.29, Original value was 17.4

## End(Not run)
```
Description
Uniformity trials of tea

Usage
data("laycock.tea.uniformity")

Format
A data frame with 54 observations on the following 4 variables.

- loc: location, L1 or L2
- row: row
- col: column
- yield: yield

Details
Actual physical dimensions for the tea shrubs are not given, so we use an estimate of four feet square for each shrub (which is similar to the eden.tea.uniformity experiment).

Location 1 (Laycock, page 108) is at the Research Station, Nyasaland. Plots were 10 by 15 bushes, harvested 23 times in 1942.

- Field length: 8 plots * 10 bushes * 4 feet = 320 feet.
- Field width: 4 plots * 15 bushes * 4 feet = 240 feet.

Location 2 (Laycock page 110) is at Mianga Estate, Nyasaland. Plots were 9 by 11 bushes, harvested 18 times in 1951/52.

- Field length: 9 plots * 9 bushes * 4 feet = 324 feet.
- Field width: 6 plots * 11 bushes * 4 feet = 264 feet.

Source

References

Examples
## Not run:
library(agridat)
data(laycock.tea.uniformity)
dat <- laycock.tea.uniformity
libs(desplot)
desplot(dat, yield ~ col*row|loc,
   flip=TRUE, aspect=322/252, # average of 2 locs
   main="laycock.tea.uniformity")

## End(Not run)

lee.potatoblight

Repeated measurements of resistance to potato blight

Description

Repeated measurements of resistance to potato blight.

Usage

data("lee.potatoblight")

Format

A data frame with 14570 observations on the following 7 variables.

year  planting year
gen  genotype / cultivar factor
col  column
row  row
rep  replicate block (numeric)
date  date for data collection
y  score 1-9 for blight resistance

Details

These data were collected from biennial screening trials conducted by the New Zealand Institute of Crop and Food Research at the Pukekohe Field Station. The trials evaluate the resistance of potato cultivars to late blight caused by the fungus *Phytophthora infestans*. In each trial, the damage to necrotic tissue was rated on a 1-9 scale at multiple time points during the growing season.

Lee (2009) used a Bayesian model that extends the ordinal regression of McCullagh to include spatial variation and sigmoid logistic curves to model the time dependence of repeated measurements on the same plot.

Data from 1989 were not included due to a different trial setup being used. All the trials here were laid out as latinized row-column designs with 4 or 5 reps. Each plot consisted of four seed tubers planted with two Ilam Hardy spread plants in a single row 2 meters long with 76 centimeter spacing between rows.
In 1997, 18 plots were lost due to flooding. In 2001, by the end of the season most plants were nearly dead.

Note, in plant-breeding, it is common to use a "breeder code" for each genotype, which after several years of testing is changed to a registered commercial variety name. For this R package, the Potato Pedigree Database, https://www.plantbreeding.wur.nl/potatopedigree/reverselookup.php, was used to change breeder codes (in early testing) to the variety names used in later testing. For example, among the changes made were the following:

Driver 287.12
Kiwitea 064/56
Gladiator 1308.66
Karaka 221.17
Kiwitea 064.56 maybe 064.54
Moonlight 511.1
Pacific 177.3
Red Rascal 1830.11
Rua 155.05
Summit 517.12
White Delight 1949.64

Used with permission of Arier Chi-Lun Lee and John Anderson.

Data retrieved from https://researchspace.auckland.ac.nz/handle/2292/5240.

Licensed via Open Database License 1.0. (allows sub-licensing). See: https://opendatacommons.org/licenses/dbcl/1.0/

Source


Examples

```r
## Not run:
library(agridat)
data(lee.potatoblight)
dat <- lee.potatoblight

# Common cultivars across years.
# Based on code from here: https://stackoverflow.com/questions/20709808
gg <- tapply(dat$gen, dat$year, function(x) as.character(unique(x)))
tab <- outer(1:11, 1:11,
Vectorize(function(a, b) length(Reduce(intersect, gg[c(a, b)]))))
head(tab) # Matches Lee page 27.
```

```
# [1,]  20  10   7   5   3   2   3   2   3   3   2
# [2,]  10  30  17   5   4   3   4   4   5   4   2
# [3,]  17  35  35   9   6   3   4   5   6   4   3
# [4,]   5   5  35  16   8   9  14  15  13  11
```
### [5,]  3  4  6 16 40 12 11 18 18 16 14

# Note the progression to lower scores as time passes in each year
skp <- c(rep(0,10),
         rep(0,7),1,1,1,
         rep(0,8),1,1,
         rep(0,6),1,1,1,1,
         rep(0,5),1,1,1,1,1,
         rep(0,5),1,1,1,1,1,1,
         rep(0,5),1,1,1,1,1,1,
         rep(0,5),1,1,1,1,1,1,1
)

libs(desplot)
desplot(dat, y ~ col*row|date,
         ylab="Year of testing", # unknown aspect
         layout=c(10,11),skip=as.logical(skp),
         main="lee.potatoblight - maps of blight resistance over time")

# 1983 only. I.Hardy succumbs to blight quickly
libs(lattice)
xyplot(y ~ date|gen, dat, subset=year==1983, group=rep,
       xlab="Date", ylab="Blight resistance score",
       main="lee.potatoblight 1983", as.table=TRUE,
       auto.key=list(columns=5),
       scales=list(alternating=FALSE, x=list(rot=90, cex=.7)))

## End(Not run)

lehmann.millet.uniformity

*Uniformity trial of millet in India*

**Description**

Uniformity trial of millet in India, 3 years on same land.

**Usage**

data("lehmann.millet.uniformity")

**Format**

A data frame with 396 observations on the following 5 variables.

- `year`
- `plot`
- `plot (row)`
range range (column)
yield grain yield (pounds)
total total crop yield (pounds)

Details

The 6th report has a map (only partially scanned in the pdf). The ninth report, pages 1-10 has comments. Pages 36-39 have data: Table 1 has grain yield, table 2 total yield.

Experiment farm near Bangalore. The plots are 1/10 acre, each 50 links wide and 200 links long. [6th report, p. 2]

The season of 1906 was abnormally wet compared with 1905 and 1907. [9th report]

Source


References

None

Examples

```r
## Not run:
library(agridat)
data(lehmann.millet.uniformity)
dat <- lehmann.millet.uniformity

libs(desplot)
dat$year = factor(dat$year)
desplot(dat, yield ~ range*plot|year,
aspect=(22*50)/(6*200),
main="lehmann.millet.uniformity",
flip=TRUE, tick=TRUE)
desplot(dat, total ~ range*plot|year,
aspect=(22*50)/(6*200),
main="lehmann.millet.uniformity",
flip=TRUE, tick=TRUE)

# libs(dplyr)
# group_by(dat, year)

## End(Not run)
```
Yield, white mold, and sclerotia for soybeans in Brazil

Usage

data("lehner.soybeanmold")

Format

A data frame with 382 observations on the following 9 variables.

- study: study number
- year: year of harvest
- loc: location name
- elev: elevation
- region: region
- trt: treatment number
- yield: crop yield, kg/ha
- mold: white mold incidence, percent
- sclerotia: weight of sclerotia g/ha

Details

Data are the mean of 4 reps.


Data included here via GPL3 license.

Source


References

Full commented code and analysis https://emdelponte.github.io/paper-white-mold-meta-analysis/
Examples

## Not run:

```r
library(agridat)
data(lehner.soybeanmold)
dat <- lehner.soybeanmold

if(0){
op <- par(mfrow=c(2,2))
    hist(dat$mold, main="White mold incidence")
    hist(dat$yield, main="Yield")
    hist(dat$sclerotia, main="Sclerotia weight")
    par(op)
}

library(lattice)
xyplot(yield ~ mold|study, dat, type=c("p","r"),
    main="lehner.soybeanmold")
# xyplot(sclerotia ~ mold|study, dat, type=c("p","r"))

# meta-analysis. Could use metafor package to construct the forest plot,
# but latticeExtra is easy; ggplot is slow/clumsy
library(latticeExtra, metafor)
# calculate correlation & confidence for each loc
cors <- split(dat, dat$study)
cors <- sapply(cors, FUN=function(X){
    res <- cor.test(X$yield, X$mold)
    c(res$estimate, res$parameter[1],
      conf.low=res$conf.int[1], conf.high=res$conf.int[2])
})
cors <- as.data.frame(t(as.matrix(cors)))
cors$study <- rownames(cors)
# Fisher Z transform
cors <- transform(cors, ri = cor)
cors <- transform(cors, ni = df + 2)
cors <- transform(cors, yi = 1/2 * log((1 + ri)/(1 - ri)),
    vi = 1/(ni - 3))
# Overall correlation across studies
overall <- rma.uni(yi, vi, method="ML", data=cors) # metafor package
# back transform
overall <- predict(overall, transf=transf.ztor)

# weight and size for forest plot
wi <- 1/sqrt(cors$vi)
size <- 0.5 + 3.0 * (wi - min(wi))/(max(wi) - min(wi))

# now the forest plot
# must use latticeExtra::layer in case ggplot2 is also loaded
segplot(factor(study) ~ conf.low+conf.high, data=cors,
    draw.bands=FALSE, level=size, centers=ri, cex=size,
    ...)```
# Meta-analyses are typically used when the original data is not available.
# Since the original data is available, a mixed model is probably better.

```r
libs(lme4)

m1 <- lmer(yield ~ mold # overall slope
  + (1+mold |study), # random intercept & slope per study
data=dat)

summary(m1)

## End(Not run)
```

---

### lessman.sorghum.uniformity

**Uniformity trial of sorghum**

**Description**

Uniformity trial of sorghum at Ames, Iowa, 1959.

**Usage**

```r
data("lessman.sorghum.uniformity")
```

**Format**

A data frame with 2640 observations on the following 3 variables.

- **row** row
- **col** column
- **yield** yield, ounces

**Details**

The uniformity trial was conducted at the Agronomy Farm at Ames, Iowa, in 1959. The field was planted to grain sorghum in rows spaces 40 inches apart, thinned to a stand of three inches between plants. The entire field was 48 rows (40 inches apart), each 300 feet long and harvested in 5-foot lengths. Threshed grain was dried to 8-10 percent moisture before weighing. Weights are ounces. Average yield for the field was 95.3 bu/ac.

Field width: 48 rows * 40 inches / 12in/ft = 160 feet
Field length: 60 plots * 5 feet = 300 feet
Plot yields from the two outer rows on each side of the field were omitted from the analysis.

CV values from this data do not quite match Lessman’s value. The first page of Table 17 was manually checked for correctness and there were no problems with the optical character recognition (other than obvious errors like 0/o).

Source


References

None.

Examples

```r
## Not run:
library(agridat)

data(lessman.sorghum.uniformity)
dat <- lessman.sorghum.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
       aspect=300/160, tick=TRUE, flip=TRUE, # true aspect
       main="lessman.sorghum.uniformity")

# Omit outer two columns (called 'rows' by Lessman)
dat <- subset(dat, col > 2 & col < 47)
nrow(dat)
var(dat$yield) # 9.09
sd(dat$yield)/mean(dat$yield) # CV 9.2

libs(reshape2)
libs(agricolae)
dmat <- acast(dat, row~col, value.var='yield')
index.smith(dmat,
           main="lessman.sorghum.uniformity",
           col="red") # Similar to Lessman Table 1
# Lessman said that varying the width of plots did not have an appreciable
# effect on CV, and optimal row length was 3.2 basic plots, about 15-20

## End(Not run)
```
li.millet.uniformity  Uniformity trial of millet

Description

Uniformity trial of millet at China in 1934.

Format

A data frame with 600 observations on the following 3 variables.

row  row
col  column
yield yield (grams)

Details

Crop date estimated to be 1934.
Field was 100 ft x 100 ft. Plots were 15 feet long by 1 foot wide.
Field width: 100 plots * 1 foot = 100 feet
Field length: 6 plots * 15 feet = 100 feet
Li found the most efficient use of land was obtained with plots 15 feet long and two rows wide. Also satisfactory would be one row 30 feet long.

Source

Li, HW and Meng, CJ and Liu, TN. 1936. Field Results in a Millet Breeding Experiment. Agronomy Journal, 28, 1-15. Table 1. https://doi.org/10.2134/agronj1936.0002196200280010001x

Examples

```r
## Not run:

library(agridat)

data(li.millet.uniformity)
dat <- li.millet.uniformity

mean(dat$yield) # matches Li et al.

libs(desplot)
desplot(dat, yield~col*row,
aspect=100/100, # true aspect
main="li.millet.uniformity")
```

## End(Not run)
Load multiple packages and install if needed

The 'agridat' package uses dozens of packages in the examples for each dataset. The 'libs' function provides a simple way to load multiple packages at once, and can install any missing packages on-the-fly. This is very similar to the 'pacman::p_load' function.

Description

Load multiple packages and install if needed

The 'agridat' package uses dozens of packages in the examples for each dataset. The 'libs' function provides a simple way to load multiple packages at once, and can install any missing packages on-the-fly.

This is very similar to the 'pacman::p_load' function.

Usage

libs(...)

Arguments

... Comma-separated unquoted package names

Value

None

Author(s)

Kevin Wright

References

None

Examples

## Not run:
libs(dplyr, reshape2)

## End(Not run)
lillemo.wheat

Multi-environment trial of wheat susceptible to powdery mildew

Description

Resistance of wheat to powdery mildew

Usage

data("lillemo.wheat")

Format

A data frame with 408 observations on the following 4 variables.

- gen  genotype, 24 levels
- env  environment, 13 levels
- score score
- scale scale used for score

Details

The data are means across reps of the original scores. Lower scores indicate better resistance to mildew.

Each location used one of four different measurement scales for scoring resistance to powdery mildew: 0-5 scale, 1-9 scale, 0-9 scale, percent.

Environment codes consist of two letters for the location name and two digits for the year of testing. Location names: CA=Cruz Alta, Brazil. Ba= Bawburgh, UK. Aa= As, Norway. Ha= Hamar, Norway. Ch= Choryn, Poland. Ce= Cerekwica, Poland. Ma= Martonvasar, Hungary. Kh= Kharkiv, Ukraine. BT= Bila Tserkva, Ukraine. Gl= Glevakha, Ukraine. Bj= Beijing, China.

Note, Lillemo et al. did not remove genotype effects as is customary when calculating Huehn’s non-parametric stability statistics.

In the examples below, the results do not quite match the results of Lillemo. This could easily be the result of the original data table being rounded to 1 decimal place. For example, environment 'Aa03' had 3 reps and so the mean for genotype 1 was probably 16.333, not 16.3.

Used with permission of Morten Lillemo.

Electronic data supplied by Miroslav Zoric.

Source

library(agridat)

data(lillemo.wheat)
dat <- lillemo.wheat

# Change factor levels to match Lillemo
dat$env <- as.character(dat$env)
dat$env <- factor(dat$env,
  levels=c("Bj03","Bj05","CA03","Ba04","Ma04",
    "Kh06","Gl05","BT06","Ch04","Ce04",
    "Ha03","Ha04","Ha05","Ha07","Aa03","Aa04","Aa05"))

# Interesting look at different measurement scales by environment
libs(lattice)
qqmath(~score|env, dat, group=scale,
  as.table=TRUE, scales=list(y=list(relation="free")),
  auto.key=list(columns=4),
  main="lillemo.wheat - QQ plots by environment")

# Not run:
# Change data to matrix format
libs(reshape2)
datm <- acast(dat, gen~env, value.var="Var(score)"

# Environment means. Matches Lillemo Table 3
apply(datm, 2, mean)

# Two different transforms within envts to approximate 0-9 scale
datt <- datm
   datt[,"CA03"] <- 1.8 * datt[,"CA03"]
ix <- c("Ba04","Kh06","Gl05","BT06","Ha03","Ha04","Ha05","Ha07","Aa03","Aa04","Aa05")
   datt[ix] <- apply(datt[ix],2,sqrt)

# Genotype means of transformed data. Matches Lillemo table 3.
   round(rowMeans(datt[,2]),2)

# Biplot of transformed data like Lillemo Fig 2
libs(gge)
biplot(gge(datt, scale=FALSE), main="lillemo.wheat")

# Median polish of transformed table
m1 <- medpolish(datt)

# Half-normal prob plot like Fig 1
libs(faraway)
   halfnorm(abs(as.vector(m1$resid)))
# Nonparametric stability statistics. Lillemo Table 4.
huehn <- function(mat) {
  # Gen in rows, Env in cols
  nenv <- ncol(mat)
  # Corrected yield. Remove genotype effects
  # Remove the following line to match Table 4 of Lillemo
  mat <- sweep(mat, 1, rowMeans(mat)) + mean(mat)
  # Ranks in each environment
  rmat <- apply(mat, 2, rank)
  # Mean genotype rank across envts
  MeanRank <- apply(rmat, 1, mean)

  # Huehn S1
  gfun <- function(x) {
    oo <- outer(x, x, "-")
    sum(abs(oo)) # sum of all absolute pairwise differences
  }
  S1 <- apply(rmat, 1, gfun) / (nenv * (nenv - 1))

  # Huehn S2
  S2 <- apply((rmat - MeanRank)^2, 1, sum) / (nenv - 1)

  out <- data.frame(MeanRank, S1, S2)
  rownames(out) <- rownames(mat)
  return(out)
}
round(huehn(datm), 2) # Matches table 4

# I do not think phenability package gives correct values for S1
# libs(phenability)
# nahu(datm)

## End(Not run)

lin.superiority

Multi-environment trial of 33 barley genotypes in 12 locations

Description

Multi-environment trial of 33 barley genotypes in 12 locations

Usage

data("lin.superiority")
**Format**

A data frame with 396 observations on the following 4 variables.

- **gen** genotype/cultivar
- **region** region
- **loc** location
- **yield** yield (kg/ha)

**Details**

Yield of six-row barley from the 1983 annual report of Eastern Cooperative Test in Canada.

The named cultivars Bruce, Conquest, Laurier, Leger are checks, while the other cultivars were tests.

**Source**


**References**


**Examples**

```r
## Not run:
library(agridat)
data(lin.superiority)
dat <- lin.superiority

libs(latticeExtra)
libs(reshape2)
# calculate the superiority measure of Lin & Binns 1988

dat2 <- acast(dat, gen ~ loc, value.var="yield")
locmean <- apply(dat2, 2, mean)
locmax <- apply(dat2, 2, max)
P <- apply(dat2, 1, function(x) {
  sum((x-locmax)^2)/(2*length(x))
})/1000
P <- sort(P)
```
round(P) # match Lin & Binns 1988 table 2, column Pi

# atlantic & quebec regions overlap
# libs(gge)
# m1 <- gge(dat, yield ~ gen*loc, env.group=region,
#   main="lin.superiority")
# biplot(m1)

# create a figure similar to Lin & Binns 1988

# add P, locmean, locmax back into the data
dat$locmean <- locmean[match(dat$loc, names(locmean))]
dat$locmax <- locmax[match(dat$loc, names(locmax))]
dat$P <- P[match(dat$gen, names(P))]
dat$gen <- reorder(dat$gen, dat$P)

xyplot(locmax ~ locmean|gen, data=dat,
   type=c('p','r'), as.table=TRUE, col="gray",
   main="lin.superiority - Superiority index",
   xlab="Location Mean",
   ylab="Yield of single cultivars (blue) & Maximum (gray)") +

xyplot(yield ~ locmean|gen, data=dat,
   type=c('p','r'), as.table=TRUE, pch=19)

## End(Not run)

---

**lin.unbalanced**  
*Multi-environment trial of 33 barley genotypes in 18 locations*

**Description**

Multi-environment trial of 33 barley genotypes in 18 locations

**Usage**

data("lin.unbalanced")

**Format**

A data frame with 405 observations on the following 4 variables.

- **gen** genotype/cultivar
- **loc** location
- **yield** yield (kg/ha)
- **region** region
Details

Yield of six-row barley from the 1986 Eastern Cooperative trial

The named cultivars Bruce, Laurier, Leger are checks, while the other cultivars were tests. Cultivar names use the following codes. "A" is for Atlantic-Quebec. "O" is for "Ontario". "S" is second-year. "T" is third-year.

Source


References

None

Examples

```r
# Not run:
library(agridat)
data(lin.unbalanced)
dat <- lin.unbalanced

# location maximum, Lin & Binns table 1
# aggregate(yield ~ loc, data=dat, FUN=max)

# location mean/index, Lin & Binns, table 1
dat2 <- subset(dat, is.element(dat$gen,
c('Bruce','Laurier','Leger','S1','S2','S3','S4','S5','S6','S7','T1','T2')))
aggregate(yield ~ loc, data=dat2, FUN=mean)

libs(reshape2)
dat3 <- acast(dat, gen ~ loc, value.var="yield")
libs(lattice)
lattice::levelplot(t(scale(dat3)), main="lin.unbalanced", xlab="loc", ylab="genotype")

# calculate the superiority measure of Lin & Binns 1988.
# lower is better
locmax <- apply(dat3, 2, max, na.rm=TRUE)
P <- apply(dat3, 1, function(x) {
  sum((x-locmax)^2, na.rm=TRUE)/(2*length(na.omit(x)))
})/1000
P <- sort(P)
round(P) # match Lin & Binns 1988 table 2, column P
```

## End(Not run)
Description

Multi-environment trial of wheat in Switzerland

Usage

data("linder.wheat")

Format

A data frame with 252 observations on the following 4 variables.

env  environment
block  block
gen  genotype
yield  yield, in 10 kg/ha

Details

An experiment of 9 varieties of wheat in 7 localities in Switzerland in 1960, RCB design.

Source


References

None.

Examples

library(agridat)
data(linder.wheat)
dat <- linder.wheat
## Not run:
libs(gge)
dat <- transform(dat, eb=paste0(env,block))
m1 <- gge(dat, yield~gen*eb, env.group=env)
biplot(m1, main="linder.wheat")

## End(Not run)
**Description**

Split-block experiment of sugar beets.

**Usage**

data("little.splitblock")

**Format**

A data frame with 80 observations on the following 6 variables.

- **row** row
- **col** column
- **yield** sugar beet yield, tons/acre
- **harvest** harvest date, weeks after planting
- **nitro** nitrogen, pounds/acre
- **block** block

**Details**

Four rates of nitrogen, laid out as a 4x4 Latin-square experiment.

Within each column block, the sub-plots are strips (across 4 rows) of 5 different harvest dates.

The use of sub-plots as strips necessitates care when determining the error terms in the ANOVA table.

Note, Little has yield value of 22.3 for row 3, column I-H3. This data uses 23.3 in order to match the marginal totals given by Little.

**Source**

Thomas M. Little, F. Jackson Hills. (1978) *Agricultural Experimentation*

**References**

None.
loesell.bean.uniformity

Uniformity trial of white pea beans

Description

Uniformity trial of white pea beans

Usage

data("loesell.bean.uniformity")
loesell.bean.uniformity

Format

A data frame with 1890 observations on the following 3 variables.

row  row ordinate
col  column ordinate
yield  yield, grams per plot

Details

Trial conducted at Michigan Agricultural Experiment Station, 1.75 acres. Beans were planted in rows 28 inches apart on 15 Jun 1932. Plants spaced 1 to 2 inches apart. After planting, an area 210 ft x 210 feet. This area was divided into 21 columns, each 10 foot wide, and each containing 90 rows.

Field length: 90 rows * 28 inches = 210 feet.
Field width: 21 series * 10 feet = 210 feet.

Author’s conclusion: Increasing the size of the plot by increasing its length was more efficient than increasing its width.

Note, the missing values in this dataset are a result of the PDF scan omitting corners of the table.

Source


Examples

## Not run:
require(agridat)
data(loesell.bean.uniformity)
dat <- loesell.bean.uniformity
require(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE, aspect=1, tick=TRUE,
       main="loesell.bean.uniformity")

## End(Not run)
Multi-environment trial of maize, half diallel

Description
Half diallel of maize

Usage
data("lonnquist.maize")

Format
A data frame with 78 observations on the following 3 variables.

- p1 parent 1 factor
- p2 parent 2 factor
- yield yield

Details
Twelve hybrids were selfed/crossed in a half-diallel design. The data here are means adjusted for block effects. Original experiment was 3 reps at 2 locations in 2 years.

Source

References

Examples
## Not run:
library(agridat)
data(lonnquist.maize)
dat <- lonnquist.maize
dat <- transform(dat,
  p1=factor(p1,
  p2=factor(p2,
### lonnquist.maize

```r
# Calculate the F1 means in Lonnquist, table 1
# libs(reshape2)
# mat <- acast(dat, p1~p2)
# mat[upper.tri(mat)] <- t(mat)[upper.tri(mat)] # make symmetric
# diag(mat) <- NA
# round(rowMeans(mat, na.rm=TRUE),1)
## C L M H G P B RM N K R2 K2
## 94.8 89.2 95.0 96.4 95.3 95.2 97.3 93.7 95.0 94.0 98.9 102.4

# Griffings method
# https://www.statforbiology.com/2021/stat_met_diallel_griffing/
# libs(lmDiallel)
# dat2 <- lonnquist.maize
# dat2 <- subset(dat2,
# is.element(p1, c("M","H","G","B","K","K2")) &
# is.element(p2, c("M","H","G","B","K","K2"))
# dat2 <- droplevels(dat2)
# dmod1 <- lm(yield ~ GCA(p1, p2) + tSCA(p1, p2),
# data = dat2)
# dmod2 <- lm.diallel(yield ~ p1 + p2,
# data = dat2, fct = "GRIFFING2")
# anova.diallel(dmod1, MSE=7.1, dfr=60)
## Response: yield
## Df Sum Sq Mean Sq F value   Pr(>F)
## GCA(p1, p2)  5 234.23  46.846  6.5980 5.923e-05 ***
## tSCA(p1, p2) 15 238.94  15.929  2.2436 0.01411  
## Residuals 60  7.100

# asreml4
# Mohring 2011 used 6 varieties to calculate GCA & SCA
# Matches Table 3, column 2
# d2 <- subset(dat, is.element(p1, c("M","H","G","B","K","K2")) &
# is.element(p2, c("M","H","G","B","K","K2")))
# d2 <- droplevels(d2)
# libs(asreml,lucid)
# m2 <- asreml(yield~ 1, data=d2, random = ~ p1 + and(p2))
# vc(m2)
## effect component std.error z.ratio  con
## p1!:p1.var 3.865 3.774  1 Positive
## R!variance 15.93 5.817  2.7 Positive

# Calculate GCA effects
```
m3 <- asreml(yield~ p1 + and(p2), data=d2)
coef(m3)$fixed-1.462
# Matches Gardner 1966, Table 5, Griffing method

## End(Not run)

lord.rice.uniformity  

Uniformity trial of rice

Description
Uniformity trial of rice in Ceylon, 1929.

Usage
data("lord.rice.uniformity")

Format
A data frame with 560 observations on the following 5 variables.

field  field
row  row
col  column
grain  grain weight, pounds per plot
straw  straw weight, pounds per plot

Details
In 1929, eight fields 1/5 acre in size were broadcast seeded with rice at the Anuradhapura Experiment Station in the northern dry zone of Ceylon. After broadcast, the fields were marked into 10 ft by 10 ft squares. At harvest, weights of grain and straw were recorded.

Fields 10-14 were on one side of a drain, and fields 26-28 on the other side.

Each field was surrounded by a bund. Plots next to the bunds had higher yields.

Field width: 5 plots * 10 feet = 50 feet
Field length: 14 plots * 10 feet = 140 feet

Conclusions: "It would appear that plots of about 1/87 acre are the most effective."

Source
Examples

```r
## Not run:

library(agridat)
data(lord.rice.uniformity)
dat <- lord.rice.uniformity

# match table on page 180
## libs(dplyr)
## dat
## field grain straw
## <chr> <dbl> <dbl>
## 1 10 590 732
## 2 11 502 600
## 3 12 315 488
## 4 13 291 538
## 5 14 489 670
## 6 26 441 560
## 7 27 451 629
## 8 28 530 718

# There are consistently high yields along all edges of the field
# libs(lattice)
# bwplot(grain ~ factor(col)|field,dat)
# bwplot(grain ~ factor(col)|field,dat)

# Heatmaps
libs(desplot)
desplot(dat, grain ~ col*row|field,
       flip=TRUE, aspect=140/50,
       main="lord.rice.uniformity")

# bivariate scatterplots
# xyplot(grain ~ straw|field, dat)

## End(Not run)
```

Description

Uniformity trial of cotton

References

None
Usage

data("love.cotton.uniformity")

Format

A data frame with 170 observations on the following 3 variables.

row  row
col  column
yield  yield

Details

Within each 100-foot row, the first 20 feet were harvested as a single plot, and then the rest of the row was harvested in 5-foot lengths.
Crop location is unknown, but likely Southeast Asia given Love’s work teaching in China. Date circa 1930.
Possibly more information would be in the collected papers of Harry Love at Cornell: https://rmc.library.cornell.edu/EAD/htmldocs/RMA00890.html
Cotton - Plot Technic Study 1930-1932. Box 3, Folder 34

Source


References

None.

Examples

## Not run:

library(agridat)

data(love.cotton.uniformity)
# omit first column which has 20-foot plots
dat <- subset(love.cotton.uniformity, col > 1)

libs(desplot)
desplot(dat, yield ~ col*row,
        flip=TRUE, aspect=20/80, # just a guess
        main="love.cotton.uniformity")

## End(Not run)
lu.stability

Multi-environment trial of maize, to illustrate stability statistics

Description
Multi-environment trial to illustrate stability statistics

Usage
data("lu.stability")

Format
A data frame with 120 observations on the following 4 variables.

yield yield
gen genotype factor, 5 levels
eenv environment factor, 6 levels
block block factor, 4 levels

Details
Data for 5 maize genotypes in 2 years x 3 sites = 6 environments.

Source

References

Examples
## Not run:
library(agridat)
data(lu.stability)
dat <- lu.stability

# GxE means. Match Lu 1995 table 1
libs(reshape2)
datm <- acast(dat, gen~env, fun=mean, value.var='yield')
round(datm, 2)
# Gen/Env means. Match Lu 1995 table 3
apply(datm, 1, mean)
apply(datm, 2, mean)

# Traditional ANOVA. Match Hwu table 2
# F value for gen,env
m1 = aov(yield~env+gen+Error(block:env+env:gen), data=dat)
summary(m1)
# F value for gen:env, block:env
m2 <- aov(yield ~ gen + env + gen:env + block:env, data=dat)
summary(m2)

# Finlay Wilkinson regression coefficients
# First, calculate env mean, merge in
libs(dplyr)
dat2 <- group_by(dat, env)
dat2 <- mutate(dat2, locmn=mean(yield))
m4 <- lm(yield ~ gen -1 + gen:locmn, data=dat2)
coef(m4) # Match Hwu table 4

# Table 6: Shukla’s heterogeneity test
dat2$ge = paste0(dat2$gen, dat2$env) # Create a separate ge interaction term
m6 <- lm(yield ~ gen + env + ge + ge:locmn, data=dat2)
m6b <- lm(yield ~ gen + env + ge + locmn, data=dat2)
anova(m6, m6b) # Non-significant difference

# Table 7 - Shukla stability
# First, environment means
emn <- group_by(dat2, env)
emn <- summarize(emn, ymn=mean(yield))
# Regress GxE terms on envt means
getab = (model.tables(m2,"effects")$tables)$"gen:env"
for (ll in 1:nrow(getab)){
  m7l <- lm(getab[ll, ] ~ emn$ymn)
  cat("\n************** Gen ",ll," **************\n")
  cat("Regression coefficient: ",round(coefficients(m7l)[2],5),"\n")
  print(anova(m7l))
} # Match Hwu table 7.

## End(Not run) # dontrun

lucas.switchback

Switchback experiment on dairy cattle, milk yield for 3 treatments

Description
Switchback experiment on dairy cattle, milk yield for 3 treatments
Usage

data("lucas.switchback")

Format

A data frame with 36 observations on the following 5 variables.

cow  cow factor, 12 levels
trt  treatment factor, 3 levels
period period factor, 3 levels
yield yield (FCM = fat corrected milk), pounds/day
block block factor

Details

Lucas says "because no data from feeding trials employing the present designs are yet available, uniformity data will be used".

Six cows were started together in block 1, then three cows in block 2 and three cows in block 3.

Source


References


Examples

library(agridat)

data(lucas.switchback)

dat <- lucas.switchback

# Create a numeric period variable
dat$per <- as.numeric(substring(dat$period, 2))

libs(lattice)

xyplot(yield ~ period|block, data=dat, group=cow, type=c('l','r'),
       auto.key=list(columns=6),
       main="lucas.switchback - (actually uniformity data)"

# Need to use 'terms' to preserve the order of the model terms
# Really, cow(block), per:cow(block), period(block)
m1 <- aov(terms(yield ~ block + cow:block + per:cow:block +
              period:block + trt, keep.order=TRUE), data=dat)
lyon.potato.uniformity

Uniformity trial of potatoes

Description
Uniformity trial of potatoes at Nebraska Experiment Station, 1909.

Format
A data frame with 204 observations on the following 3 variables.

row row
col column, section
yield yield, pounds

Details
In 1909, potatoes were harvested from uniform land at Nebraska Experiment Station.
There were 34 rows, 34 inches apart. Lyon, page 97 says "He harvested each row in six sections, each of which was seventy-two feet and seven inches long." It is not clear if each section is 72 feet long, or if the entire row is 72 feet long. Yield of potato is roughly 0.5 to 0.8 pounds per square foot, so it seems more plausible the entire row is 72 feet long (see calculations below).

Field width: 6 plots = 72 feet
Field length: 34 rows * 34 in / 12in/ft = 96 ft

Source
Lyon, T.L. (1911). Some experiments to estimate errors in field plat tests. Proc. Amer. Soc. Agron, 3, 89-114. Table III. https://doi.org/10.2134/agronj1911.00021962000300010016x

References
None.
### Examples

```r
## Not run:
library(agridat)

data(lyon.potato.uniformity)
dat <- lyon.potato.uniformity

# Yield per square foot, assuming 72 foot rows
sum(dat$yield)/(72*96) # 0.67 # seems about right
# Yield per square foot, assuming 72 foot plots
sum(dat$yield)/(6*72*96) # 0.11

libs(desplot)
desplot(dat, yield ~ col*row, 
  tick=TRUE, flip=TRUE, aspect=96/72, # true aspect
  main="lyon.potato.uniformity")

## End(Not run)
```

### lyons.wheat

**Multi-environment trial of winter wheat at 12 sites in 4 years.**

### Description

Yield of winter wheat at 12 sites in 4 years.

### Format

A data frame with 48 observations on the following 3 variables.

- `loc` location, 12 levels
- `year` year, numeric
- `yield` yield (kg)

### Details

Krzanowski uses this briefly for multi-dimensional scaling.

### Source


### References

Examples

```
library(agridat)

data(lyons.wheat)
dat <- lyons.wheat

libs(lattice)
xyplot(yield~factor(year), dat, group=loc,
     main="lyons.wheat",
     auto.key=list(columns=4), type=c("l","l"))
```

---

`magistad.pineapple.uniformity`

**Uniformity trial of pineapple**

Description

Uniformity trial of pineapple in Hawaii in 1932

Usage

```
data("magistad.pineapple.uniformity")
```

Format

A data frame with 137 observations on the following 6 variables.

- `field` field number
- `plat` plat number
- `row` row
- `col` column
- `number` number of fruits
- `weight` weight of fruits, grams

Details


"In this field, harvested in 1932, there were four rows per bed. A 300-foot bed was divided into four equal parts to form plats 1, 2, 3, and 4. The third [sic, second] bed from this was similarly divided to form plats 5 to 8, inclusive. In the same manner plats 9 to 24 were formed. In this way 24 plats each 75 feet long and 1 bed wide were formed."

Field length: 4 plats * 75 feet = 300 feet
Field width: 6 plats * 6.5 feet = 39 feet

Field 82. Pearl City.
"Eight beds, each separated by two beds, were selected and harvested. Beds were 8 feet center to center. Each bed was divided into three plats 76 feet long."

Note: Layout of plats into rows/columns assumes the same pattern as field 19.

Field length: 3 plats * 76 feet = 228 feet
Field width: 8 plats * 8 feet * 3 (every third bed) = 192 feet.


"In field 21, Kahuku, the experimental plan was of the Latin square type, having five beds of five plats each. The beds were 7.5 feet center to center. Each plat was approximately 60 feet long and each third bed was selected and harvested." Note: Layout of plats into rows/columns assumes the same pattern as field 19.

Field length: 5 plats * 60 feet = 300 feet
Field width: 5 plats * 7.5 feet * 3 (every third bed) = 112.5 feet

Field 1. Kunia.

"This experiment was another Latin square test having eight plats in each column and eight plats in each row. It was harvested in 1930. Each plat consisted of two beds 150 feet long. Beds were 6 feet center to center and consisted of three rows each. The entire experimental area occupied 2.85 acres."

Field length: 8 plats * 2 beds * 150 feet = 2400 feet
Field width: 8 plats * 6 feet = 48 feet
Total area: 2400*48/43560=2.64 acres

Source


References

None

Examples

## Not run:

library(agridat)

data(magistad.pineapple.uniformity)
dat <- magistad.pineapple.uniformity

# match table page 641
## dat
## summarize(number=mean(number),
## weight=mean(weight))
## field number weight
## 1 1 596.4062 2499.922
## 2 19 171.1667 2100.250
## 3 21 171.1600 2056.800
## Data

```r
# 4 82 220.7500 1264.500

libs(desplot)
desplot(dat, weight ~ col*row,
        subset=field==1,
        aspect=2400/48,
        main="magistad.pineapple.uniformity - field 1")
desplot(dat, weight ~ col*row,
        subset=field==19,
        aspect=300/39,
        main="magistad.pineapple.uniformity - field 19")
desplot(dat, weight ~ col*row,
        subset=field==82,
        aspect=228/192,
        main="magistad.pineapple.uniformity - field 82")
desplot(dat, weight ~ col*row,
        subset=field==21,
        aspect=300/112.5,
        main="magistad.pineapple.uniformity - field 21")

## End(Not run)
```

---

**masood.rice.uniformity**

*Uniformity trial of rice*

---

**Description**

Uniformity trial of rice at Lahore, Punjab, circa 2011.

**Usage**

```r
data("masood.rice.uniformity")
```

**Format**

A data frame with 288 observations on the following 3 variables.

- `row` row
- `col` column
- `yield` yield, kg/m^2

**Details**

Data by collected from the Rice Research Institute on a paddy yield trial. A single variety of rice was harvested in an area 12m x 24 m. Yield in kilograms was measured for each square meter. Masood et al report a low degree of similarity for neighboring plots.
Note, the Smith index calculations below match the results in the Pakistan Journal of Agricultural Research, but do not match the results in the American-Eurasian Journal, which seems to be the same paper and seems to refer to the same data. The results may simply differ by a scaling factor.

Field length: 24 plots x 1m = 24m.
Field width: 12 plots x 1m = 12m.

Used with permission of Asif Masood.

Source

References

Examples

```r
## Not run:
library(agridat)

data(masood.rice.uniformity)
dat <- masood.rice.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE, tick=TRUE, aspect=24/12, # true aspect
       main="masood.rice.uniformity - yield heatmap")

libs(agricolae)
libs(reshape2)
dmat <- acast(dat, row~col, value.var="yield")
index.smith(dmat,
           main="masood.rice.uniformity",
           col="red") # CVs match Table 3

## End(Not run)
```

---

**mcclelland.corn.uniformity**

*Uniformity trial of corn*

**Description**

Uniformity trial of corn at Arkansas Experiment Station, 1925.
Usage

```r
data("mcclelland.corn.uniformity")
```

Format

A data frame with 438 observations on the following 3 variables.

- `row` row
- `col` column
- `yield` yield

Details

A uniformity trial of corn in 1925 at the Arkansas Experimental Station.

Field width = 66 ft * 2 = 132 feet.

Field length = 219 rows * 44 inches / 12 inches/ft = 803 ft.

Note: In the source document, table 2, first 'west' column and second-to-last row (page 822), the value 1.40 is assumed to be a typographical error and was changed to 14.0 for this data.

Source


References

None

Examples

```r
## Not run:
library(agridat)

data(mcclelland.corn.uniformity)
dat <- mcclelland.corn.uniformity

# McClelland table 3, first row, gives 11.2
# Probable error = 0.67449 * sd(). Relative to mean.
# 0.67449 * sd(dat$yield)/mean(dat$yield) # 11.2

libs(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE,
       aspect=(219*44/12)/132, # true aspect, 219 rows x 44 inches x 132 feet
       main="mcclelland.corn.uniformity")

## End(Not run)
```
Description

RCB experiment of turnips, 2 treatments for planting date and density

Format

A data frame with 64 observations on the following 6 variables.

- gen genotype
- date planting date, levels 21Aug1990 28Aug1990
- density planting density, 1, 2, 4, 8 kg/ha
- block block, 4 levels
- yield yield

Details

This is a randomized block experiment with 16 treatments allocated at random to each of four blocks. The 16 treatments were combinations of two varieties, two planting dates, and four densities.


Piepho (2009) proposed an ordinary ANOVA using transformed data.

Used with permission of Kevin McConway.

Source


References


Examples

```r
## Not run:
library(agridat)
data(mcconway.turnip)
dat <- mcconway.turnip
dat$densf <- factor(dat$density)
```
# Table 2 of Lee et al.

```r
m0 <- aov( yield ~ gen * densf * date + block, dat )
summary(m0)
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## gen 1 84.0 83.95 8.753 0.00491 **
## densf 3 470.4 156.79 16.347 2.51e-07 ***
## date 1 233.7 233.71 24.367 1.14e-05 ***
## block 3 163.7 54.58 5.690 0.00216 **
## gen:densf 3 8.6 2.88 0.301 0.82485
## gen:date 1 36.5 36.45 3.800 0.05749 .
## densf:date 3 154.8 51.60 5.380 0.00299 **
## gen:densf:date 3 18.0 6.00 0.626 0.60224
## Residuals 45 431.6 9.59
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Boxplots suggest heteroskedasticity for date, density
```
libs("HH")
interaction2wt(yield ~ gen + date + densf + block, dat,
  x.between=0, y.between=0,
  main="mcconway.turnip - yield")
```

```r
libs(nlme)
# Random block model
m1 <- lme(yield ~ gen * date * densf, random= ~1|block, data=dat)
summary(m1)
anova(m1)
```

# Multiplicative variance model over densities and dates
```r
m2 <- update(m1,
  weights=varComb(varIdent(form=~1|densf),
                  varIdent(form=~1|date)))
summary(m2)
anova(m2)
```

# Unstructured variance model over densities and dates
```r
m3 <- update(m1, weights=varIdent(form=~1|densf*date))
summary(m3)
anova(m3)
```

# Table 3 of Piepho, using transformation
```r
m4 <- aov( yield^.235 ~ gen * date * densf + block, dat )
summary(m4)
```

## End(Not run)
Multi-environment trial of barley in South Canterbury with yield and yield components

Description

Yield and yield components for barley with different seeding rates.

Format

A data frame with 40 observations on the following 10 variables.

- year: year, numeric
- site: site factor
- rate: rate, numeric
- plants: plants per sq meter
- tillers: tillers per plant
- heads: heads per plant
- surviving: percent surviving tillers
- grains: grains per head
- weight: weight of 1000 grains
- yield: yield tons/hectare

Details

Trials were conducted at 5 sites, 3 years in South Canterbury. (not all sites in every year). Values are the average of 6 blocks. In 1974 there was a severe drought. The other years had favorable growing conditions.

Source


References


Examples

```r
library(agridat)

data(mcleod.barley)
dat <- mcleod.barley

# Table 3 of McLeod. Across-environment means by planting rate
```
d1 <- aggregate(cbind(plants, tillers, heads, surviving, grains, weight, yield) ~ rate, dat, FUN=mean)
# Calculate income based on seed cost of $280/ton, grain $140/ton.
d1 <- transform(d1, income=140*yield-280*rate/1000)
signif(d1,3)
## rate plants tillers heads surviving grains weight yield
## 50 112.12 5.22 4.36 83.95 21.25 46.11 3.97
## 75 162.75 4.04 3.26 80.89 19.95 45.10 4.26
## 100 202.62 3.69 2.73 74.29 19.16 44.66 4.38
## 125 239.00 3.28 2.33 71.86 18.45 43.45 4.41
## 150 293.62 2.90 2.00 69.54 17.94 42.77 4.47

# Even though tillers/plant, heads/plant, surviving tillers, 
# grains/head, weight/1000 grains are all decreasing as planting 
# rate increases, the total yield is still increasing. 
# But, income peaks around seed rate of 100.

libs(lattice)
xyplot(yield +income +surviving +grains +weight +plants +tillers +heads ~ rate, 
data=d1, outer=TRUE, type=c('p','l'), 
scales=list(y=list(relation="free")), 
xlab="Nitrogen rate", ylab="Trait value", 
main="mcleod.barley - nitrogen response curves" )

---

**mead.cauliflower**

*Leaves for cauliflower plants at different times*

**Description**

Leaves for cauliflower plants at different times in two years.

**Format**

A data frame with 14 observations on the following 4 variables.

- **year**  year factor
- **degdays**  degree days above 32F
- **leaves**  number of leaves

**Details**

Numbers of leaves for 10 cauliflower plants in each of two years, and temperature degree-days above 32F, divided by 100.

The year is 1956-57 or 1957-58.

Over the data range shown, the number of leaves is increasing linearly. Extrapolating backwards shows that a linear model is inappropriate, and so a glm is used.
**Source**


**References**


**Examples**

```r
## Not run:
library(agridat)
data(mead.cauliflower)
dat <- mead.cauliflower
dat <- transform(dat, year=factor(year))
m1 <- glm(leaves ~ degdays + year, data=dat, family=poisson)
coef(m1)
## (Intercept) degdays year1957
## 3.49492453 0.08512651 0.21688760
dat$pred <- predict(m1, type="response")
libs(lattice)
libs(latticeExtra)
xyplot(leaves~degdays, data=dat, groups=year, type=c("p"),
    auto.key=list(columns=2),
    main="mead.cauliflower - observed (symbol) & fitted (line)",
    xlab="degree days", ylab="Number of leaves",
    ) +
xyplot(pred~degdays, data=dat, groups=year, type=c("l"), col="black")

## End(Not run)
```

---

**mead.cowpea.maize**  
**Intercropping experiment of maize/cowpea**

**Description**

Intercropping experiment of maize/cowpea, multiple nitrogen treatments.

**Format**

A data frame with 72 observations on the following 6 variables.

<table>
<thead>
<tr>
<th>variable</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>block</td>
<td>block, 3 levels</td>
</tr>
<tr>
<td>nitro</td>
<td>nitrogen, 4 levels</td>
</tr>
</tbody>
</table>
cowpea  cowpea variety, 2 levels
maize  maize variety, 3 levels
cyield  cowpea yield, kg/ha
myield  maize yield, kg/ha

Details
An intercropping experiment conducted in Nigeria. The four nitrogen treatments were 0, 40, 80, 120 kg/ha.

Source

References

Examples
## Not run:
library(agridat)
data(mead.cowpea.maize)
dat <- mead.cowpea.maize

# Cowpea and maize yields are clearly in competition
libs("latticeExtra")
useOuterStrips(xyplot(myield ~ cyield|maize*cowpea, dat, group=nitro,
main="mead.cowpea.maize - intercropping",
xlab="cowpea yield",
ylab="maize yield", auto.key=list(columns=4)))

# Mead Table 2 Cowpea yield anova...strongly affected by maize variety.
anova(aov(cyield ~ block + maize + cowpea + nitro +
    maize:cowpea + maize:nitro + cowpea:nitro +
    maize:cowpea:nitro, dat))

# Cowpea mean yields for nitro*cowpea
aggregate(cyield ~ nitro+cowpea, dat, FUN=mean)

# Cowpea mean yields for each maize variety
aggregate(cyield ~ maize, dat, FUN=mean)

# Bivariate analysis
aov.c <- anova(aov(cyield/1000 ~ block + maize + cowpea + nitro +
    maize:cowpea + maize:nitro + cowpea:nitro +
    maize:cowpea:nitro, dat))
aov.m <- anova(aov(myield/1000 ~ block + maize + cowpea + nitro +
mead.germination

Seed germination with different temperatures/concentrations

Description

Seed germination with different temperatures/concentrations

Format

A data frame with 64 observations on the following 5 variables.

temp  temperature regimen
rep  replication factor (not blocking)
conc  chemical concentration
germs  number of seeds germinating
seeds  number of seeds tested = 50

Details

The rep factor is NOT a blocking factor.

Used with permission of Roger Mead, Robert Curnow, and Anne Hasted.
Source

References

Examples

library(agridat)
data(mead.germination)
dat <- mead.germination
dat <- transform(dat, concf=factor(conc))
library(lattice)
xyplot(germ~log(conc+.01)|temp, dat, layout=c(4,1),
    main="mead.germination", ylab="number of seeds germinating")
m1 <- glm(cbind(germ, seeds-germ) ~ 1, dat, family=binomial)
m2 <- glm(cbind(germ, seeds-germ) ~ temp, dat, family=binomial)
m3 <- glm(cbind(germ, seeds-germ) ~ concf, dat, family=binomial)
m4 <- glm(cbind(germ, seeds-germ) ~ temp + concf, dat, family=binomial)
m5 <- glm(cbind(germ, seeds-germ) ~ temp * concf, dat, family=binomial)
anova(m1,m2,m3,m4,m5)
## Resid. Df Resid. Dev Df Deviance
## 1 63 1193.80
## 2 60 430.11 3 763.69
## 3 60 980.10 0 -549.98
## 4 57 148.11 3 831.99
## 5 48 55.64 9 92.46

# Show logit and fitted values. T2 has highest germination
subset(cbind(dat, predict(m5), fitted(m5)), rep=="R1")

mead.lamb

Number of lambs born to 3 breeds on 3 farms

Description
Number of lambs born to 3 breeds on 3 farms

Usage
data("mead.lamb")
Format

A data frame with 36 observations on the following 4 variables.

farm farm: F1, F2, F3
breed breed: B1, B2, B3
lambclass lambing class: L0, L1, L2, L3
y count of ewes in class

Details

The data 'y' are counts of ewes in different lambing classes. The classes are number of live lambs per birth for 0, 1, 2, 3+ lambs.

Source


References

None

Examples

## Not run:

library(agridat)
data(mead.lamb)
dat <- mead.lamb

# farm 1 has more ewes in lambclass 3
d2 <- xtabs(y ~ farm+breed+lambclass, data=dat)
mosaicplot(d2, color=c("lemonchiffon1", "moccasin", "lightsalmon1", "indianred"),
           xlab="farm/lambclass", ylab="breed", main="mead.lamb")

names(dat) <- c('F', 'B', 'L', 'y') # for compactness
# Match totals in Mead example 14.6
libs(dplyr)
dat <- group_by(dat, F,B)
summarize(dat, y=sum(y))

# F  B  y
## <fctr> <fctr> <int>
## 1 F1 A 150
## 2 F1 B 46
## 3 F1 C 78
## 4 F2 A 72
## 5 F2 B 79
## 6 F2 C 28
## 7 F3 A 224
## 8 F3 B 129
# Models
m1 <- glm(y ~ F + B + F:B, data=dat,
          family=poisson(link=log))
m2 <- update(m1, y ~ F + B + F:B + L)
m3 <- update(m1, y ~ F + B + F:B + L + B:L)
m4 <- update(m1, y ~ F + B + F:B + L + F:L)
m5 <- update(m1, y ~ F + B + F:B + L + B:L + F:L)
AIC(m1, m2, m3, m4, m5) # Model 4 has best AIC

# Change contrasts for Miroslav
m4 <- update(m4,
              contrasts=list(F=contr.sum,B=contr.sum,L=contr.sum))
summary(m4)

# Match deviance table from Mead
libs(broom)
all <- do.call(rbind, lapply(list(m1, m2, m3, m4, m5), broom::glance))
all$model <- unlist(lapply(list(m1, m2, m3, m4, m5),
                           function(x) as.character(formula(x)[3])))
all[, c('model', 'deviance', 'df.residual')]

if(0){
  # Using MASS::loglm
  libs(MASS)
  # Note: without 'fitted=TRUE', devtools::run_examples has an error
  m4b <- MASS::loglm(y ~ F + B + F:B + L + F:L, data = dat, fitted=TRUE)
  # Table of farm * class interactions. Match Mead p. 360
  round(coef(m4b)$F.L,2)
  fitted(m4b)
  resid(m4b)
  # libs(vcd)
  # mosaic(m4b, shade=TRUE,
  # formula = ~ F + B + F:B + L + F:L,
  # residual_type="rstandard", keep_aspect=FALSE)
## RCB experiment of strawberry

**Description**

RCB experiment of strawberry

**Format**

A data frame with 32 observations on the following 5 variables.

- row row
- col column
- block block, 4 levels
- gen genotype, 8 levels
- yield yield, pounds

**Details**

A hedge along the right side (column 8) caused shading and lower yields.

R. Mead said (in a discussion of the Besag & Higdon paper), "the blocks defined (as given to me by the experimenter) are the entire horizontal rows...the design of the trial is actually (and unrecognized by me also) a checker-board of eight half-blocks with two groups of split-plot varieties".

The two sub-groups of genotypes are G, V, R1, F and Re, M, E, P.

**Source**

Unknown, but prior to 1968 according to Besag. Probably via R. Mead.

**References**


Examples

```r
## Not run:
library(agridat)
data(mead.strawberry)
dat <- mead.strawberry
dat$sub <- ifelse(is.element(dat$gen, c('G', 'V', 'R1', 'F')), "S1", "S2")
libs(desplot)
desplot(dat, yield~col*row,
    text=gen, cex=1, out1=block, out2=sub, # unknown aspect
    main="mead.strawberry")

## End(Not run)
```

---

`mead.turnip` Density/spacing experiment for turnips in 3 blocks.

Description

Density/spacing experiment for turnips in 3 blocks.

Usage

```r
data("mead.turnip")
```

Format

A data frame with 60 observations on the following 4 variables.

- `yield` log yield (pounds/plot)
- `block` block
- `spacing` row spacing, inches
- `density` density of seeds, pounds/acre

Details

An experiment with turnips, 3 blocks, 20 treatments in a factorial arrangement of 5 seeding rates (density) and 4 widths (spacing).

Source

References


Examples

```r
## Not run:
library(agridat)
data(mead.turnip)
dat <- mead.turnip
dat$ratef <- factor(dat$density)
dat$widthf <- factor(dat$spacing)

m1 <- aov(yield ~ block + ratef + widthf + ratef:widthf, data=dat)
anova(m1) # table 12.10 in Mead

# Similar to Piepho fig 10
libs(lattice)
xypplot(yield ~ log(spacing)|ratef, data=dat, auto.key=list(columns=5), main="mead.turnip - log(yield) for each density", group=ratef)

## End(Not run)
```

---

mercer.mangold.uniformity

*Uniformity trial of mangolds*

Description

Uniformity trial of mangolds at Rothamsted Experiment Station, England, 1910.

Usage

```r
data("mercer.mangold.uniformity")
```

Format

A data frame with 200 observations on the following 4 variables.

- **row**: row
- **col**: column
- **roots**: root yields, pounds
- **leaves**: leaf yields, pounds
Details

Grown in 1910.

Each plot was 3 drills, each drill being 2.4 feet wide. Plots were 1/200 acres, 7.2 feet by 30.25 feet long. The "length of the plots runs with the horizontal lines of figures [in Table I], this being also the direction of the drills across the field."

Field width: 10 plots * 30.25 ft = 302.5 feet

Field length: 20 plots * 7.25 ft = 145 feet

Source


References


Examples

```r
## Not run:

library(agridat)

data(mercer.mangold.uniformity)
dat <- mercer.mangold.uniformity

libs(desplot)
desplot(dat, leaves~col*row,
    aspect=145/302, # true aspect
    main="mercer.mangold.uniformity - leaves")

libs(desplot)
desplot(dat, roots~col*row,
    aspect=145/302, # true aspect
    main="mercer.mangold.uniformity - roots")

libs(lattice)
xypplot(roots~leaves, data=dat)

## End(Not run)```
Description

Uniformity trial of wheat at Rothamsted Experiment Station, England, 1910.

Format

A data frame with 500 observations on the following 4 variables.

- row
- col
- grain: grain yield, pounds
- straw: straw yield, pounds

Details

The wheat crop was grown in the summer of 1910 at Rothamsted Experiment Station (Harpenden, Hertfordshire, England). In the Great Knott, a seemingly uniform area of 1 acre was harvested in separate plots, each 1/500th acre in size. The grain and straw from each plot was weighed separately. McCullagh gives more information about the plot size.

Field width: 25 plots * 8 ft = 200 ft
Field length: 20 plots * 10.82 ft = 216 ft

D. G. Rossiter (2014) uses this data for an extensive data analysis tutorial.

Source


References


The 'spdep' package includes the grain yields (only) and spatial positions of plot centres in its example dataset 'wheat'.

Note, checked that all '4.03' values in this data match the original document.
Examples

```# Not run:

library(agridat)
data(mercer.wheat.uniformity)
dat <- mercer.wheat.uniformity

libs(desplot)
desplot(dat, grain ~ col*row, 
aspect=216/200, # true aspect
main="mercer.wheat.uniformity - grain yield")

libs(lattice)
xyplot(straw ~ grain, data=dat, type=c('p','r'),
main="mercer.wheat.uniformity - regression")

libs(hexbin)
hexbinplot(straw ~ grain, data=dat)

libs(sp, gstat)
plot.wid <- 2.5
plot.len <- 3.2
nr <- length(unique(dat$row))
nc <- length(unique(dat$col))
xy <- expand.grid(x = seq(plot.wid/2, by=plot.wid, length=nc),
    y = seq(plot.len/2, by=plot.len, length=nr))
dat.sp <- dat
coordinates(dat.sp) <- xy
# heatmap
spplot(dat.sp, zcol = "grain", cuts=8,
    cex = 1.6,
    col.regions = bpy.colors(8),
    main = "Grain yield", key.space = "right")

# variogram
# Need gstat::variogram to get the right method
vg <- gstat::variogram(grain ~ 1, dat.sp, cutoff = plot.wid * 10, width = plot.wid)
plot(vg, plot.numbers = TRUE,
    main="mercer.wheat.uniformity - variogram")

# Not run)```

miguez.biomass

Biomass of 3 crops in Greece
Description

Biomass of 3 crops in Greece

Usage

```
data("miguez.biomass")
```

Format

A data frame with 212 observations on the following 5 variables.

- **doy**: day of year
- **block**: block, 1-4
- **input**: management input, Lo/Hi
- **crop**: crop type
- **yield**: yield tons/ha

Details

Experiment was conducted in Greece in 2009. Yield values are destructive Measurements of above-ground biomass for fiber sorghum, maize, sweet sorghum.

Hi management refers to weekly irrigation and high nitrogen applications. Lo management refers to bi-weekly irrigation and low nitrogen.

The experiment had 4 blocks.

Crops were planted on DOY 141 with 0 yield.

Source


References


Examples

```r
## Not run:
library(agridat)
data(miguez.biomass)
dat <- miguez.biomass
dat <- subset(dat, doy > 141)

library(lattice)
xyplot(yield ~ doy | crop*input, data = dat,
       main="miguez.biomass",
       panel=function(x,y,...){
         panel.xyplot(x,y,...)
         panel.abline(a=0,b=1,col=4)
       })
```
groups = crop,
        type=c('p','smooth'),
auto.key=TRUE)

# ----------
# Archontoulis et al fit some nonlinear models.
# Here is a simple example which does NOT account for crop/input
# Slow, so dont run
if(0){
  dat2 <- transform(dat, eu = paste(block, input, crop))
  dat2 <- groupedData(yield ~ doy | eu, data = dat2)
  fit.lis <- nlsList(yield ~ SSfpl(doy, A, B, xmid, scal),
                      data = dat2,
                      control=nls.control(maxiter=100))
  print(plot(intervals(fit.lis)))
}

libs(nlme)
# use all data to get initial values
inits <- getInitial(yield ~ SSfpl(doy, A, B, xmid, scal), data = dat2)
inits
xvals <- 150:325
y1 <- with(as.list(inits), SSfpl(xvals, A, B, xmid, scal))
plot(yield ~ doy, dat2)
lines(xvals,y1)

# must have groupedData object to use augPred
dat2 <- groupedData(yield ~ doy|eu, data=dat2)
plot(dat2)

# without 'random', all effects are included in 'random'
m1 <- nlme(yield ~ SSfpl(doy, A, B, xmid, scale),
            data= dat2,
            fixed= A + B + xmid + scale ~ 1,
            random = B ~ 1|eu, # to make only B random
            random = A + B + xmid + scale ~ 1|eu,
            start=inits)
fixef(m1)
summary(m1)
plot(augPred(m1, level=0:1),
     main="miguez.biomass - observed/predicted data") # only works with groupedData object
}

## End(Not run)
Description

This is monthly weather summaries for the 6 sites where barley yield trials were conducted.

Format

A data frame with 719 observations on the following 8 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>site</td>
<td>site, 6 levels</td>
</tr>
<tr>
<td>year</td>
<td>year, 1927-1936</td>
</tr>
<tr>
<td>mo</td>
<td>month, 1-12, numeric</td>
</tr>
<tr>
<td>cdd</td>
<td>monthly cooling degree days, Fahrenheit</td>
</tr>
<tr>
<td>hdd</td>
<td>monthly heating degree days, Fahrenheit</td>
</tr>
<tr>
<td>precip</td>
<td>monthly precipitation, inches</td>
</tr>
<tr>
<td>min</td>
<td>monthly average daily minimum temp, Fahrenheit</td>
</tr>
<tr>
<td>max</td>
<td>monthly average daily maximum temp, Fahrenheit</td>
</tr>
</tbody>
</table>

Details

When the weather data was extracted from the National Climate Data Center, the following weather stations were chosen, based on availability of weather data in the given time frame (1927-1936) and the proximity to the town (site) for the barley data.

<table>
<thead>
<tr>
<th>site</th>
<th>station name</th>
<th>station</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morris</td>
<td>MORRIS WC EXPERIMENTAL STATION</td>
<td>USC00215638</td>
</tr>
<tr>
<td>StPaul</td>
<td>MINNEAPOLIS WEATHER BUREAU DOWNTOWN</td>
<td>USC00215433</td>
</tr>
<tr>
<td>Crookston</td>
<td>CROOKSTON NW EXPERIMENTAL STATION</td>
<td>USC00211891</td>
</tr>
<tr>
<td>GrandRapids</td>
<td>GRAND RAPIDS FRS LAB</td>
<td>USC00213303</td>
</tr>
<tr>
<td>Waseca</td>
<td>WASECA EXPERIMENTAL STATION</td>
<td>USC00218692</td>
</tr>
<tr>
<td>Duluth</td>
<td>SUPERIOR</td>
<td>USC00478349</td>
</tr>
</tbody>
</table>

'cdd' are cooling degree days, which is the number of degree days with a temperature _above_ 65 Fahrenheit.

'hdd' are heating degree days, _below_ 65 Fahrenheit.

No data is available for Duluth in Dec, 1931.

Source


References

Examples

## Not run:

```r
library(agridat)
data(minnesota.barley.yield)
dat <- minnesota.barley.yield
data( minnesota.barley.weather)
datw <- minnesota.barley.weather

# Weather trends over time
libs(latticeExtra)
useOuterStrips(xyplot(cdd~mo|year*site, datw, groups=year,
 main="minnesota.barley",
 xlab="month", ylab="Cooling degree days",
 subset=(mo > 3 & mo < 10),
 scales=list(alternating=FALSE),
 type='l', auto.key=list(columns=5))

# Total cooling/heating/precip in Apr-Aug for each site/yr
ww <- subset(datw, mo>=4 & mo<=8)
ww <- aggregate(cbind(cdd,hdd,precip)~site+year, data=ww, sum)

# Average yield per each site/env
yy <- aggregate(yield~site+year, dat, mean)
minn <- merge(ww, yy)

# Higher yields generally associated with cooler temps, more precip
libs(reshape2)
me <- melt(minn, id.var=c('site','year'))
meY <- subset(me, variable=='yield')
names(meY) <- c('site','year','value')
mec <- subset(me, variable=='covar')
names(mec) <- c('site','year','covar','x')
mecY <- merge(mec, meY)
mecY$yr <- factor(mecY$year)

foo <- xyplot(y~x|covar*site, data=mecY, groups=yr, cex=1, ylim=c(5,65),
 par.settings=list(superspose.symbol=list(pch=substring(levels(mecY$yr),4))),
 xlab="", ylab="yield", main="minnesota.barley",
 panel=function(x,y,...) {
   panel.lmline(x,y,..., col="gray")
   panel.superpose(x,y,...)
 },
 scales=list(x=list(relation="free")))
libs(latticeExtra)
foo <- useOuterStrips(foo, strip.left = strip.custom(par.strip.text=list(cex=.7)))
combineLimits(foo, margin.x=2L) # Use a common x axis for all rows
```
minnesota.barley.yield

Multi-environment trial of barley in Minnesota at 6 sites in 1927-1936.

Description
This is an expanded version of the barley data that is often used to illustrate dot plots.

Format
A data frame with 647 observations on the following 4 variables.

- yield yield in bu/ac
- gen genotype (variety) factor
- year year
- site site factor, 6 levels

Details
The lattice package contains a smaller version of this data for the years 1931 and 1932.
The following comments are in reference to the mentioned source documents.
Notes about Immer (1934).
The University Farm location is at the city of Saint Paul.
This source provides the yield data for each of the three blocks at each location in 1931 and 1932.
The following registration numbers and names are given:

<table>
<thead>
<tr>
<th>C.I. number</th>
<th>Variety name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minn 184</td>
<td>Manchuria</td>
</tr>
<tr>
<td>Minn 445</td>
<td>Glabron</td>
</tr>
<tr>
<td>Minn 440</td>
<td>Svansota</td>
</tr>
<tr>
<td>Minn 447</td>
<td>Velvet</td>
</tr>
<tr>
<td>Minn 448</td>
<td>Trebi</td>
</tr>
<tr>
<td>Minn 457</td>
<td>Manchuria x Smooth Awn</td>
</tr>
<tr>
<td>Minn 462</td>
<td>Smooth Awn x Manchuria</td>
</tr>
<tr>
<td>Minn 452</td>
<td>Peatland</td>
</tr>
<tr>
<td>Minn 475</td>
<td>Svanhals x Lion</td>
</tr>
<tr>
<td>Minn 529</td>
<td>Wisconsin No 38</td>
</tr>
</tbody>
</table>

Notes about Harlan (1935):
The 1931 yields match the average values of Immer (1934).
The Minnesota 474 and 475 cultivars are both 'Svanhals x Lion' crosses.
No yields are reported at Crookston in 1928 because of a crop failure. (Page 20)
Also, in the report for North Dakota it says "the zero yields at Williston, ND in 1931 were caused by drought'. (Page 31)

Notes about Wiebe (1940):
The 1932 data generally match the average values from Immer (1934) with the following notes.
The data for Glabron at St Paul in 1932 are missing, but given as 36.8 in Immer (1934). This value is treated as missing in this R dataset.
The data for Svansota at Morris in 1932 are missing, but given as 35.0 in Immer (1934). This value is treated as missing in this R dataset.
The yield for 'Wisconsin 38' at St Paul in 1932 is shown as 3.80, but 38 in Immer (1934). The latter value is used in this R dataset.
The yields for No475 in 1932 are not reported in Wiebe (1940), but are reported in Immer (1934).

Notes about Hayes (1942).
This sources gives the block-level yield data for 5 cultivars at 4 sites in 1932 and 1935. Cultivar 'Barbless' is the same as 'Wisconsin No38'.

Source

References

Examples

library(agridat)

data(minnesota.barley.yield)
dat <- minnesota.barley.yield
dat$yr <- factor(dat$year)

# Drop Dryland, Jeans, CompCross, MechMixture because they have less than 5
# year-loc values
dat <- droplevels(subset(dat, !is.element(gen, c("CompCross","Dryland","Jeans","MechMixture"))))

# 1934 has huge swings from one loc to the next
montgomery.wheat.uniformity

 uneven distribution

405

libs(lattice)
dotplot(gen~yield|site, dat, groups=yr,
   main="minnesota.barley.yield",
   auto.key=list(columns=5), scales=list(y=list(cex=.5)))

montgomery.wheat.uniformity

Uniformity trial of wheat, 2 years on the same land

Description

Uniformity trial of wheat at Nebraska Experiment Station, 1909 & 1911.

Usage

data("montgomery.wheat.uniformity")

Format

A data frame with 448 observations on the following 3 variables.

year year
col column
row row
yield yield, grams

Details

Experiments were conducted by the Nebraska Experiment Station.
A field was sown to Turkey winter wheat in the fall of 1908 and harvested in 1909. The drill, 5.5 feet wide, was driven across the first series of 14 blocks, the boundaries of the blocks being later established. Each series was sown the same way, no space was allowed between the blocks. Each block was 5.5 ft square.
The experiment was done 3 times with harvests in 1909, 1910, 1911. A simple heatmap of the 3 years' yields are shown in Montgomery (1912), figure 3, p. 178.
The 1909 data are given by Montgomery (1913), figure 10, page 37. NOTE: North is at the right side of this diagram (as determined by comparing yield values with the fertility map in Montgomery 1912, p. 178).
The 1910 data are not available.
The 1911 data are given by Montgomery (1912), figure 1, page 165. NOTE: North is at the top of this diagram.
Surface & Pearl (1916) give a simple method for adjusting yield due to fertility effects using the 1909 data.
Field width: 14 plots * 5.5 feet
Field length: 16 blocks * 5.5 feet
Source


References


Examples

```r
## Not run:

library(agridat)

data(montgomery.wheat.uniformity)
dat <- Montgomery.wheat.uniformity
dat09 <- subset(dat, year==1909)
dat11 <- subset(dat, year==1911)

# Match the figures of Montgomery 1912 Fig 3, p. 178
libs(desplot)
desplot(dat09, yield ~ col*row,
       aspect=1, # true aspect
       main="montgomery.wheat.uniformity - 1909 yield")
desplot(dat, yield ~ col*row, subset= year==1911,
       aspect=1, # true aspect
       main="montgomery.wheat.uniformity - 1911 yield")

# Surface & Pearl adjust 1909 yield for fertility effects.
# They calculate smoothed yield as (row sum)*(column sum)/(total)
# and subtract this from the overall mean to get 'deviation'.
# We can do something similar with a linear model with rows and columns
# as factors, then predict yield to get the smooth trend.
# Corrected yield = observed - deviation = observed - (smooth-mean)

m1 <- lm(yield ~ factor(col) + factor(row), data=dat09)
dev1 <- predict(m1) - mean(dat09$yield)
# Corrected. Similar (but not exact) to Surface, fig 2.
dat09$correct <- round(dat09$yield - dev1,0)

libs(desplot)
desplot(dat09, yield ~ col*row,
        shorten="none", text=yield,
        main="montgomery.wheat.uniformity 1909 observed")
desplot(dat09, correct ~ col*row, text=correct,
        cex=0.8, shorten="none",
```

```

# Corrected yields are slightly shrunk toward overall mean
plot(correct~yield,dat09, xlim=c(350,1000), ylim=c(350,1000))
abline(0,1)

## End(Not run)
```

---

**moore.uniformity**  
*Uniformity trials of pole beans, bush beans, sweet corn, carrots, spring and fall cauliflower*

---

**Description**

Uniformity trials of pole beans, bush beans, sweet corn, carrots, spring and fall cauliflower at Washington, 1952-1955.

**Format**

Each data frame has the following columns at a minimum. Some datasets have an additional trait column.

- `row` row
- `col` column
- `yield` yield

**Details**

All trials were grown on sandy loam soil in the Puyallup valley of Washington. In most experiments a gradient in soil fertility was evident. Moore & Darroch appear to have assigned 4 treatments to the plots and used the residual variation to calculate a CV. In the examples below a 'raw' CV is calculated and is always higher than the CV given by Moore & Darroch.

**Blue Lake Pole Beans.**

Conducted 1952. Seven pickings were made at about 5-day intervals. Table 26.

Field width: 12 rows x 5 feet = 60 feet.
Field length: 12 ranges x 10 feet = 120 feet.

**Bush Beans.**

Conducted in 1955. Two harvests. Table 27.

Field width: 24 rows x 3 feet = 72 feet.
Field length: 24 ranges x 5 feet = 120 feet.

**Sweet Corn.**

Conducted 1952. Table 28-29.

Field width: 24 rows x 3 feet = 72 feet.
Field length: 12 ranges x 10 feet = 120 feet.
Carrot.
Conducted 1952. Table 30.
Field width: 24 rows * 1.5 feet = 36 feet.
Field length: 12 ranges * 5 feet = 60 feet.

Spring Cauliflower.
Conducted spring 1951. Five harvests. Table 31-32.
Field width: 12 rows x 3 feet = 36 feet.
Field length: 10 plants * 1.5 feet * 20 ranges = 300 feet.

Fall Cauliflower.
Conducted fall 1951. Five harvests. Table 33-34.
Field width: 12 rows x 3 feet = 36 feet.
Field length: 10 plants * 1.5 feet * 20 ranges = 300 feet.

Source
Moore, John F and Darroch, JG. (1956). Field plot technique with Blue Lake pole beans, bush
beans, carrots, sweet corn, spring and fall cauliflower. Washington Agricultural Experiment Sta-
tions, Institute of Agricultural Sciences, State College of Washington.

References
None.

Examples
## Not run:

```r
library(agridat)

# Pole Bean
data(moore.polebean.uniformity)
cv(moore.polebean.uniformity$yield) # 8.00. Moore says 6.73.
desplot(moore.polebean.uniformity, yield~col*row,
        flip=TRUE, tick=TRUE, aspect=120/60, # true aspect
        main="moore.polebean.uniformity - yield")

# Bush bean
data(moore.bushbean.uniformity)
cv(moore.bushbean.uniformity$yield) # 12.1. Moore says 10.8
        desplot(moore.bushbean.uniformity, yield~col*row,
        flip=TRUE, tick=TRUE, aspect=120/72, # true aspect
        main="moore.bushbean.uniformity - yield")

# Sweet corn
```
nagai.strawberry.uniformity

Uniformity trial of strawberry

```r
library(desplot)
data(moore.sweetcorn.uniformity)
cv(moore.sweetcorn.uniformity$yield) # 17.5. Moore says 13.6

graph <- desplot(moore.sweetcorn.uniformity, yield~col*row,
                  flip=TRUE, tick=TRUE, aspect=120/72, # true aspect
                  main="moore.sweetcorn.uniformity - yield")

graph <- desplot(moore.sweetcorn.uniformity, ears~col*row,
                 flip=TRUE, tick=TRUE, aspect=120/72, # true aspect
                 main="moore.sweetcorn.uniformity - ears")

library(lattice)
xyplot(yield ~ ears, moore.sweetcorn.uniformity)

# Carrot

data(moore.carrot.uniformity)
cv(moore.carrot.uniformity$yield) # 33.4. Moore says 27.6

graph <- desplot(moore.carrot.uniformity, yield~col*row,
                 flip=TRUE, tick=TRUE, aspect=60/36, # true aspect
                 main="moore.carrot.uniformity - yield")

library(lattice)
xyplot(yield ~ ears, moore.carrot.uniformity)

# Spring cauliflower

data(moore.springcauliflower.uniformity)
cv(moore.springcauliflower.uniformity$yield) # 21. Moore says 19.5

graph <- desplot(moore.springcauliflower.uniformity, yield~col*row,
                  flip=TRUE, tick=TRUE, aspect=300/36, # true aspect
                  main="moore.springcauliflower.uniformity - yield")

graph <- desplot(moore.springcauliflower.uniformity, heads~col*row,
                 flip=TRUE, tick=TRUE, aspect=300/36, # true aspect
                 main="moore.springcauliflower.uniformity - heads")

library(lattice)
xyplot(yield ~ heads, moore.springcauliflower.uniformity)

# Fall cauliflower

data(moore.fallcauliflower.uniformity)
cv(moore.fallcauliflower.uniformity$yield) # 17.7. Moore says 17.0

graph <- desplot(moore.fallcauliflower.uniformity, yield~col*row,
                  flip=TRUE, tick=TRUE, aspect=300/36, # true aspect
                  main="moore.fallcauliflower.uniformity - yield")

graph <- desplot(moore.fallcauliflower.uniformity, heads~col*row,
                 flip=TRUE, tick=TRUE, aspect=300/36, # true aspect
                 main="moore.fallcauliflower.uniformity - heads")

library(lattice)
xyplot(yield ~ heads, moore.fallcauliflower.uniformity)

## End(Not run)
```
Description
Uniformity trial of strawberry in Brazil.

Usage
data("nagai.strawberry.uniformity")

Format
A data frame with 432 observations on the following 3 variables.

row row
col column
yield yield grams/plot

Details
A uniformity trial of strawberry, at Jundiai, Brazil, in April 1976.
The spacing between plants and rows was 0.3 m. Test area was 233.34 m^2. There were 18 rows of 144 plants. Each plat consisted of 6 consecutive plants. There were 432 plats, each 0.54 m^2.
Field length: 18 rows * 0.3 m = 5.4 m.
Field width: 24 columns * 6 plants * 0.3 m = 43.2 m.

Source

References
None

Examples
## Not run:
library(agridat)

data(nagai.strawberry.uniformity)
dat <- nagai.strawberry.uniformity

# CV matches Nagai
# with(dat, sd(yield)/mean(yield))
# 23.42

library(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE, aspect=(5.4)/(43.2), # true aspect
Description
Uniformity trial of turmeric in India, 1984.

Usage
data(nair.turmeric.uniformity)

Format
A data frame with 864 observations on the following 3 variables.

row row ordinate
col column ordinate
yield yield per plot, kg

Details
An experiment conducted at the College of Horticulture, Vellanikkara, India, in 1984. The crop was grown in raised beds.
The gross experimental area was 74.2 m long x 15.2 m wide. Small elevated beds 0.6 m x 1.5 m were raised providing channels of 0.4 m around each bed. One row of beds all around the experiment was discarded to eliminate border effects. After discarding the borders, there were 432 beds in the experiment. At the time of harvest, each bed was divided into equal plots of size .6 m x .75 m, and the yield from each plot was recorded.
Field map on page 64 of Nair.
Field length: 14 plots * .6 m + 13 alleys * .4 m = 13.6 m
Field width: 72 plots * .75 m + 35 alleys * .4 m = 68 m
Data found in the appendix.

Source

References
None.
narain.sorghum.uniformity

## Examples

### Not run:

```r
library(agridat)
data(nair.turmeric.uniformity)
dat <- gopakumaran.turmeric.uniformity
libs(desplot)
desplot(dat, yield ~ col*row, 
  flip=TRUE, aspect=13.6/68, 
  main="nair.turmeric.uniformity")

### End(Not run)
```

---

### Description

Uniformity trial of sorghum in Pakistan, 1936.

### Usage

```r
data("narain.sorghum.uniformity")
```

### Format

A data frame with 160 observations on the following 3 variables.

- **row**
- **col**
- **yield** yield, maunds per 1/40 acre

### Details

A uniformity trial with chari (sorghum) at Rawalpindi Agricultural Station (Pakistan) in kharif (monsoon season) in 1936. Each plot was 36 feet by 30.25 feet. The source document does not describe the orientation of the plots, but the fertility map shown in Narain figure 1 shows the plots are taller than wide.

- Field width: 10 plots * 30.25 feet
- Field length: 16 plots * 36 feet

### Source

nass.corn

References
None

Examples

## Not run:

```r
library(agridat)

data(narain.sorghum.uniformity)
dat <- narain.sorghum.uniformity

# Narain figure 1
libs(desplot)
desplot(dat, yield ~ col*row,
       flip=TRUE, aspect=(16*36)/(10*30.25),
       main="narain.sorghum.uniformity")
```

## End(Not run)

nass.corn  U.S. historical crop yields by state

Description

Yields and acres harvested in each state for the major agricultural crops in the United States, from approximately 1900 to 2011. Crops include: barley, corn, cotton, hay, rice, sorghum, soybeans, wheat.

Usage

nass.barley
nass.corn	nass.cotton
nass.hay
nass.sorghum
nass.wheat
nass.rice
nass.soybean

Format

year  year
state  state factor
acres  acres harvested
yield  average yield
Details

Be cautious with yield values for states with small acres harvested.

Yields are in bushels/acre, except: cotton pounds/acre, hay tons/acre, rice pounds/acre.

Each crop is in a separate dataset: nass.barley, nass.corn, nass.cotton, nass.hay, nass.sorghum, nass.wheat, nass.rice, nass.soybean.

Source


Examples

library(agridat)

data(nass.corn)
dat <- nass.corn

# Use only states that grew at least 100K acres of corn in 2011
keep <- drop.levels(subset(dat, year == 2011 & acres > 100000))$state

dat <- drop.levels(subset(dat, is.element(state, keep)))

# Acres of corn grown each year
libs(lattice)

xyplot(acres ~ year|state, dat, type='l', as.table=TRUE,
  main="nass.corn: state trends in corn acreage")

## Plain levelplot, using only states
## libs(reshape2)
## datm <- acast(dat, year~state, value.var='yield')
## redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
## levelplot(datm, aspect=.7, col.regions=redblue,
##   main="nass.corn",
##   scales=list(x=list(rot=90, cex=.7)))

# Model the rate of genetic gain in Illinois as a piecewise regression
# Breakpoints define periods of open-pollinated varieties, double-cross,
# single-cross, and transgenic hybrids.

dil <- subset(nass.corn, state="Illinois" & year >= 1900)
m1 <- lm(yield ~ pmin(year,1932) + pmax(1932, pmin(year, 1959)) +
  pmax(1959, pmin(year, 1995)) + pmax(1995, year), dil)
signif(coef(m1)[-1],3) # Rate of gain for each segment
plot(yield ~ year, dil, main="nass.corn: piecewise linear model of Illinois corn yields")

lines(dil$year, fitted(m1))
abline(v=c(1932,1959,1995), col="wheat")
Description

Nebraska farm income in 2007 by county

Format

A data frame with 93 observations on the following 4 variables.

county  county
crop    crop income, thousand dollars
animal  livestock and poultry income, thousand dollars
area    area of each county, square miles

Details

The variables for each county are:
Value of farm products sold - crops (NAICS) 2007 (adjusted)
Value of farm products sold - livestock, 2007 (adjusted).
Area in square miles.
Note: Cuming county is a very important beef-producing county. Some counties are not reported
to protect privacy. Western Nebraska is dryer and has lower income. South-central Nebraska is
irrigated and has higher crop income per square mile.

Source


Examples

## Not run:

library(agridat)

data(nebraska.farmincome)
dat <- nebraska.farmincome

libs(maps, mapproj, latticeExtra)
# latticeExtra for mapplot

dat$stco <- paste0('nebraska.', dat$county)
# Scale to million dollars per county
dat <- transform(dat, crop=crop/1000, animal=animal/1000)

# Raw, county-wide incomes. Note the outlier Cuming county
```r
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
mapplot(stco ~ crop + animal, data = dat, colramp = redblue,
  main = "nebraska.farmincome",
  xlab = "Farm income from animals and crops (million $ per county)",
  scales = list(draw = FALSE),
  map = map('county', 'nebraska', plot = FALSE, fill = TRUE,
    projection = "mercator")
)

# Now scale to income/mile^2
dat <- within(dat, {
  crop.rate <- crop/area
  animal.rate <- animal/area
})
# And use manual breakpoints.
mapplot(stco ~ crop.rate + animal.rate, data = dat, colramp = redblue,
  main = "nebraska.farmincome: income per square mile (percentile breaks)",
  xlab = "Farm income (million $ / mi^2) from animals and crops",
  scales = list(draw = FALSE),
  map = map('county', 'nebraska', plot = FALSE, fill = TRUE,
    projection = "mercator"),
  # Percentile break points
  # breaks = quantile(c(dat$crop.rate, dat$animal.rate),
  #   c(0,.1,.2,.4,.6,.8,.9,1), na.rm=TRUE)
  # Fisher-Jenks breakpoints via classInt package
  # breaks = classIntervals(na.omit(c(dat$crop.rate, dat$animal.rate)),
  #   n = 7, style = 'fisher')$brks
  breaks = c(0, .049, .108, .178, .230, .519, .958, 1.31))
```

---

**nonnecke.peas.uniformity**

*Uniformity trial of canning peas*

**Description**

Uniformity trial of canning peas in southern Alberta, 1957.

**Usage**

```r
data("nonnecke.peas.uniformity")
```

**Format**

A data frame with 540 observations on the following 5 variables.

- block  block factor
- row  row

---
Details

Width of basic plot was 10 feet, length was 5 feet, as limited by the viner. At each of two blocks/locations, planting consisted of 18 rows (only 15 rows were harvested) that were 10 feet wide and 90 feet long. Rows were separated by 7 foot bare ground to facilitate harvesting. Nonnecke 1960 shows a map of one block.

Plots were harvested with a five foot mower. Vines from each plot were weighed, then shelled. The two blocks/locations were side by side and combined by Nonnecke. The optimum plot size was found to be 5 feet long and 10 feet wide.

Field width: 15 rows * 10 ft/row + 14 gaps * 7 ft/gap = 248 feet
Field length: 18 plots * 5 ft/plot = 90 feet

Source


References


Examples

```r
## Not run:

library(agridat)

data(nonnecke.peas.uniformity)
dat <- nonnecke.peas.uniformity

libs(desplot)
desplot(dat, vines~col*row|block,
   tick=TRUE, flip=TRUE, aspect=248/90, # true aspect
   main="nonnecke.peas.uniformity - vines")
desplot(dat, peas~col*row|block,
   tick=TRUE, flip=TRUE, aspect=248/90, # true aspect
   main="nonnecke.peas.uniformity - peas")

libs(lattice)
xyplot(peas~vines|block, dat,
   xlab="vine weight", ylab="shelled pea weight",
   main="nonnecke.peas.uniformity")
```
nonnecke.sweetcorn.uniformity

Uniformity trial of sweet corn

Description

Uniformity trials of sweet corn in Alberta, 1956.

Usage

data("nonnecke.sweetcorn.uniformity")

Format

A data frame:

- loc location
- row row
- col column
- yield yield of marketable ears, pounds

Details

Experiments were conducted at three locations in Southern Alberta at Lethbridge, Vauxhall, and Cranford in 1956. Plot layout was 32 rows, each 179 feet long, allowing 18 ten-foot plots per row. Rows were 3 feet apart, thinned to one foot between plants. A double guard row surrounded the entire plot. The same two persons were assigned to harvest the corn from all locations. All 576 plots were harvested in one day. Optimal plot sizes were found to be 10ft x 6ft or 20ft by 3ft. The R data uses row/column for plot/row.

Field width: 18 plots * 10 ft = 180 feet
Field length: 32 rows * 3 ft = 96 feet

Source


References

Examples

```r
## Not run:

library(agridat)

# Corn 1
data(nonnecke.sweetcorn.uniformity)
dat <- nonnecke.sweetcorn.uniformity
libs(desplot)
desplot(dat, yield~col*row|loc,
       flip=TRUE, tick=TRUE, aspect=96/180, # true aspect
       main="nonnecke.sweetcorn.uniformity")

## End(Not run)
```

---

obsi.potato.uniformity

*Uniformity trial of potato in Africa 2001*

Description

Uniformity trial of potato in Africa in 2001

Usage

```r
data("obsi.potato.uniformity")
```

Format

A data frame with 2569 observations on the following 4 variables.

- `loc` location, 2 levels
- `row` row
- `col` column
- `yield` yield

Details

Data collected from potato uniformity trials at Hollota (L1) and Kulumsa (L2). Each field was 0.15 hectares. In each field, 75cm between rows and 60cm between plants. The basic units harvested were 1.2m x 1.5m. It is not clear which way the plots are oriented in the field with respect to the rows and columns.

At location L1, plot (10,7) was 22.5 in the source document, but was changed to 2.25 for this electronic data.

Note the horizontal banding of 8 or 9 rows at location L1.
odland.soybean.uniformity

Source


References

None.

Examples

```r
library(agridat)
data(obsi.potato.uniformity)
dat <- obsi.potato.uniformity

# libs(dplyr)
# dat <- group_by(dat, loc)
# summarize(dat, yield=mean(yield))
## loc  yield
## <fct> <dbl>
## 1 L1  2.54 # Obsi says 2.55
## 2 L2  5.31 # Obsi says 5.36

## Not run:
libs(desplot)
desplot(dat, yield ~ col*row, subset=loc=="L1",
        main="obsi.potato.uniformity - loc L1",
        flip=TRUE, tick=TRUE)
desplot(dat, yield ~ col*row, subset=loc=="L2",
        main="obsi.potato.uniformity - loc L2",
        flip=TRUE, tick=TRUE)

## End(Not run)
```

odland.soybean.uniformity

Uniformity trials of soy hay and soybeans

Description

Uniformity trials of soy hay and soybeans at Virginia Experiment Station, 1925-1926.

Format

Data frames with 3 variables.

row row

col column

yield yield: hay in tons, beans in bushels
Details

Grown at West Virginia Experiment Station in 1925 & 1926. Soybeans were planted in rows 30 inches apart.

Soy Hay

In 1925 the crop was harvested for forage, 42 rows, each 200 feet long. Yields of 8-foot plats recorded.

Field width: 42 plots * 30 in / 12in/ft = 105 ft
Field length: 24 plots * 8 feet = 192 feet + border = total 200 feet.

Soy Beans

In 1926 the crop was harvested for seed, 55 rows, each 232 feet long. Yields of 8-foot plats were recorded. In 1926, data for the last row on page 96 seems to be missing.

Field width: 55 plots * 30 in / 12in/ft = 137.5 feet
Field length: 28 plots * 8 feet = 224 feet + border = total 232 feet.

Odland and Garber provide no agronomic context for the yield variation.

Source


Examples

```r
## Not run:

library(agridat)

libs(desplot)

    data(odland.soyhay.uniformity)
    dat1 <- odland.soyhay.uniformity
    desplot(dat1, yield ~ col*row,
            flip=TRUE, aspect=200/105, # true aspect
            main="odland.soyhay.uniformity")

    data(odland.soybean.uniformity)
    dat2 <- odland.soybean.uniformity
    desplot(dat2, yield ~ col*row,
            flip=TRUE, aspect = 232/137,
            main="odland.soybean.uniformity")

## End(Not run)
```
Multi-environment trial of sorghum, 6 environments

Usage

data("omer.sorghum")

Format

A data frame with 432 observations on the following 4 variables.

- env  environment
- rep  replication
- gen  genotype factor
- yield  yield, kg/ha

Details

Trials were conducted in Sudan, 3 years at 2 locations, 4 reps in RCBD at each location. The year and location have been combined to form 6 environments. Only environments are given in the data, not the individual year and location.

Source

Siraj Osman Omer, Abdel Wahab Hassan Abdalla, Mohammed Hamza Mohammed, Murari Singh (2015). Bayesian estimation of genotype-by-environment interaction in sorghum variety trials Communications in Biometry and Crop Science, 10 (2), 82-95.

Electronic data provided by Siraj Osman Omer.

References

None.

Examples

library(agridat)

data(omer.sorghum)
dat <- omer.sorghum

# REML approach
libs(lme4)
libs(lucid)
# 1 loc, 2 years. Match Omer table 1.
m1 <- lmer(yield ~ 1 + env + (1|env:rep) + (1|gen) + (1|gen:env),
           data=subset(dat, is.element(env, c('E2','E4'))))
vc(m1)
## grp  var1  var2  vcov   sdcor
## gen:env (Intercept) <NA> 17050 130.6
## gen (Intercept) <NA> 2760 52.54
## env:rep (Intercept) <NA> 959.1 30.97
## Residual <NA> <NA> 43090 207.6

# 1 loc, 3 years. Match Omer table 1.
m2 <- lmer(yield ~ 1 + env + (1|env:rep) + (1|gen) + (1|gen:env),
           data=subset(dat, is.element(env, c('E2','E4','E6'))))
vc(m2)
## grp  var1  var2  vcov   sdcor
## gen:env (Intercept) <NA> 22210 149
## gen (Intercept) <NA> 9288 96.37
## env:rep (Intercept) <NA> 1332 36.5
## Residual <NA> <NA> 40270 200.7

# all 6 locs. Match Omer table 3, frequentist approach
m3 <- lmer(yield ~ 1 + env + (1|env:rep) + (1|gen) + (1|gen:env),
           data=dat)
vc(m3)
## grp  var1  var2  vcov   sdcor
## gen:env (Intercept) <NA> 21340 146.1
## env:rep (Intercept) <NA> 1152 33.95
## gen (Intercept) <NA> 1169 34.2
## Residual <NA> <NA> 24660 157

onofri.winterwheat  Multi-environment trial of winter wheat, 7 years

Description
Multi-environment trial of winter wheat, 7 years, 8 gen

Usage
data("onofri.winterwheat")

Format
A data frame with 168 observations on the following 5 variables.
year  year, numeric
block  block, 3 levels
424

ortiz.tomato
plot plot, numeric
gen genotype, 7 levels
yield yield for each plot

Details
Yield of 8 durum winter wheat varieties across 7 years with 3 reps.
Downloaded electronic version from here Nov 2015: https://www.casaonofri.it/Biometry/index.html
Used with permission of Andrea Onofri.
Source
Andrea Onofri, Egidio Ciriciofolo (2007). Using R to Perform the AMMI Analysis on Agriculture
References
F. Mendiburu. AMMI. https://tarwi.lamolina.edu.pe/~fmendiburu/AMMI.htm
Examples
library(agridat)
data(onofri.winterwheat)
dat <- onofri.winterwheat
dat <- transform(dat, year=factor(dat$year))
m1 <- aov(yield ~ year + block:year + gen + gen:year, dat)
anova(m1) # Matches Onofri figure 1
## Not run:
libs(agricolae)
m2 <- AMMI(dat$year, dat$gen, dat$block, dat$yield)
plot(m2)
title("onofri.winterwheat - AMMI biplot")
## End(Not run)

ortiz.tomato

Multi-environment trial of tomato in Latin America, weight/yield and
environmental covariates

Description
Multi-environment trial of tomato in Latin America, weight/yield and environmental covariates


Usage

```r
data("ortiz.tomato.covs")
data("ortiz.tomato.yield")
```

Format

The `ortiz.tomato.covs` data frame has 18 observations on the following 18 variables.

- `env`  environment
- `Day`  degree days (base 10)
- `Dha`  days to harvest
- `Driv`  drivings (0/1)
- `ExK`  extra potassium (kg / ha)
- `ExN`  extra nitrogen (kg / ha)
- `ExP`  extra phosphorous (kg / ha)
- `Irr`  irrigation (0/1)
- `K`  potassium (me/100 g)
- `Lat`  latitude
- `Long`  longitude
- `MeT`  mean temperature (C)
- `MnT`  min temperature (C)
- `MxT`  max temperature (C)
- `OM`  organic matter (percent)
- `P`  phosphorous (ppm)
- `pH`  soil pH
- `Prec`  precipitation (mm)
- `Tri`  trimming (0/1)

The `ortiz.tomato.yield` data frame has 270 observations on the following 4 variables.

- `env`  environment
- `gen`  genotype
- `yield`  marketable fruit yield t/ha
- `weight`  fruit weight, g

Details

The environment locations are:

- E04  Estanzuela, Guatemala
- E05  Baja Verapaz, Guatemala
- E06  Cogutepeque, El Salvador
- E07  San Andres, El Salvador
E11  Comayagua, Honduras
E14  Valle de Sabaco, Nicaragua
E15  San Antonio de Belen, Costa Rica
E20  San Cristobal, Dominican Republic
E21  Constanza, Dominican Republic
E27  Palmira, Colombia
E40  La Molina, Peru
E41  Santiago, Chile
E42  Chillan, Chile
E43  Curacavi, Chile
E44  Colina, Chile
E50  Belem, Brazil
E51  Caacupe, Paraguay
E53  Centeno, Trinidad Tobago

Used with permission of Rodomiro Ortiz.

Source


Examples

```r
## Not run:
library(agridat)
data(ortiz.tomato.covs)
data(ortiz.tomato.yield)

libs(pls, reshape2)
# Double-centered yield matrix
Y <- acast(ortiz.tomato.yield, env ~ gen, value.var='Var yield')
Y <- sweep(Y, 1, rowMeans(Y, na.rm=TRUE))
Y <- sweep(Y, 2, colMeans(Y, na.rm=TRUE))

# Standardized covariates
X <- ortiz.tomato.covs
rownames(X) <- X$env
"ExN", "ExP", "ExK", "Trim", "Driv", "Irr", "Dha")]
X <- scale(X)

# Now, PLS relating the two matrices.
# Note: plsr deletes observations with missing values
m1 <- pls(M~X)
# Inner-product relationships similar to Ortiz figure 1.
```
**Description**

Yields of 18 soybean genotypes at 11 environments in Brazil.

**Format**

- **gen** genotype, 18 levels
- **env** environment, 11 levels
- **yield** yield, kg/ha

**Details**

In each environment was used an RCB design with 3 reps. The means of the reps are shown here. Used with permission of Robert Pacheco.

**Source**


**Examples**

```r
## Not run:
library(agridat)
data(pacheco.soybean)
dat <- pacheco.soybean

# AMMI biplot similar to Fig 2 of Pacheco et al.
libs(agricolae)
m1 <- with(dat, AMMI(env, gen, REP=1, yield))
bip <- m1$biplot[,1:3]
# Fig 1 of Pacheco et al.
with(bip, plot(yield, PC1, cex=0.0,
    text(yield,PC1,labels=row.names(bip), col="blue"),
    xlim=c(1000,3000),main="pacheco.soybean - AMMI biplot",frame=TRUE))
with(bip[19:29,], points(yield, PC1, cex=0.0,
    text(yield,PC1,labels=row.names(bip[19:29,]),
    col="darkgreen"))
```
paez.coffee.uniformity

### Description

Uniformity trial of coffee in Caldas Colombia

### Usage

```r
data("paez.coffee.uniformity")
```

### Format

A data frame with 4190 observations on the following 5 variables.

- **plot**: plot number
- **row**: row
- **col**: column
- **year**: year
- **yield**: yield per tree, kilograms

### Details

The field map on Paez page 56, has plots 1 to 838. The data tables on page 79-97 have data for plots 1 to 900.

Note: The 'row' ordinate in this data would imply that the rows and columns are perpendicular. But the field map on page 56 of Paez shows that the rows are not at a 90-degree angle compared to the columns, but only at a 60-degree angle compared to the columns. In other words, the columns are vertical, and the rows are sloping up and right at about 30 degrees.

Paez looks at blocks that are 1,2,...,36 trees in size. Page 30 shows annual CV.

### Source


### References

None
panse.cotton.uniformity

**Examples**

```r
## Not run:

library(agridat)
data(panse.coffee.uniformity)
dat <- paez.coffee.uniformity

libs(reshape2, corrgram)
datt <- acast(dat, plot ~ year)
corrgram(datt, lower.panel=panel.pts,
  main="paez.coffee.uniformity")

# Not quite right. The rows are not actually horizontal. See notes above.
libs(desplot)
desplot(dat, yield ~ col*row,subset=year="Y1",
  tick=TRUE, aspect=1,
  main="paez.coffee.uniformity - Y1")
desplot(dat, yield ~ col*row,subset=year="Y2",
  tick=TRUE, aspect=1,
  main="paez.coffee.uniformity - Y2")
desplot(dat, yield ~ col*row,subset=year="Y3",
  tick=TRUE, aspect=1,
  main="paez.coffee.uniformity - Y3")
desplot(dat, yield ~ col*row,subset=year="Y4",
  tick=TRUE, aspect=1,
  main="paez.coffee.uniformity - Y4")
desplot(dat, yield ~ col*row,subset=year="Y5",
  tick=TRUE, aspect=1,
  main="paez.coffee.uniformity - Y5")

## End(Not run)
```

---

**panse.cotton.uniformity**

*Uniformity trial of cotton*

**Description**

Uniformity trial of cotton in India in 1934.

**Usage**

```r
data("panse.cotton.uniformity")
```
Format

A data frame with 1280 observations on the following 3 variables.

row  row
col  column
yield  total yield per plot, grams

Details

A uniformity trial of cotton at the Institute of Plant Industry, Indore, India.
The trial consisted of 128 rows of cotton with a spacing of 14 in between rows and length 186 feet 8 in.
Each harvested plot was 4 rows wide and 4 ft 8 in long, measuring 1/2000 acre.
Four pickings were made between Nov 1933 and Jan 1934. The data here are the total yields.
The fertility map shows appreciable variation, not following any systematic pattern.
Field length: 40 plots * 4 feet 8 inches = 206 feet 8 inches
Field width: 32 plots * 4 rows/plot * 14 inches/row = 150 feet
Conclusions: Lower error was obtained when the plots were long rows instead of across the rows.

Source


References


Examples

```r
## Not run:
library(agridat)
data(panse.cotton.uniformity)
dat <- panse.cotton.uniformity

# match the CV of Panse 1954
# sd(dat$yield)/mean(dat$yield) * 100
# 32.1

# match the fertility map of Hutchinson, fig 1
libs(desplot)
desplot(dat, yield ~ col*row,
```
parker.orange.uniformity

Description
Uniformity trial of oranges at Riverside, CA, 1921-1927.

Usage
data("parker.orange.uniformity")

Format
A data frame with 1364 observations on the following 4 variables.

- year  year
- row  row
- col  column
- yield  yield, pounds/tree for plot

Details
An orchard of naval oranges was planted in 1917 at the University of California Citrus Experiment Station at Riverside. The orchard was maintained under uniform conditions for 10 years.

Eight Washington Navel orange trees in a single row constituted a plot. The planting distance is 20 feet between trees within the row and 24 feet between rows. Every other row was a guard row, so row 2 and row 4 were observational units, while row 3 was a guard row. For example, from row 2 to row 4 is 2*24 = 48 feet. Another way to think of this is that each plot was 48 feet wide, but only the middle 24 feet was harvested. At each end of the plot was one guard tree. Including guard trees at the row ends, each row plot was 10 trees * 20 feet = 200 feet long.

Field width (west-east) 10 plots * 200 feet = 2000 feet.
Field length (north-south) 27 plots * 48 feet = 1296 feet.

An investigation into the variability between plots included systematic soil surveys, soil moisture, soil nitrates, and inspection for differences in infestation of the citrus nematode. None of these factors was considered to be the primary cause of the variations in yield.

After the 7 years of uniformity trials, different treatments were applied to the plots.
Parker et al. state that soil heterogeneity is considerable and first-year yields are not predictive of future yields.

Table 25 has mean top volume per tree for each plot in 1926. Table 26 has mean area of trunk cross section.
Source


References


Examples

```r
## Not run:
library(agridat)

data(parker.orange.uniformity)
dat <- parker.orange.uniformity

# Parker fig 2, field plan
libs(desplot)
dat$year <- factor(dat$year)
# 27 rows * 48 ft x 10 cols * 200 feet
desplot(dat, yield ~ col*row|year, 
   flip = TRUE, aspect = 27*48/(10*200), # true aspect
   main = "parker.orange.uniformity")

# CV across plots in each year. Similar to Parker table 11
cv <- function(x) {
  x <- na.omit(x)
  sd(x)/mean(x)
}
round(100*tapply(dat$yield, dat$year, cv),2)

# Correlation of plot yields across years. Similar to Parker table 15.
# Parker et al may have calculated correlation differently.
libs(reshape2)
libs(corrgram)
dat2 <- acast(dat, row+col ~ year, value.var = 'yield')
round(cor(dat2, use = 'pair'),3)
corrgram(dat2, lower = panel.pts, upper = panel.conf, 
            main="parker.orange.uniformity")

# Fertility index. Mean across years (ignoring 1921). Parker table 16
dat3 <- aggregate(yield ~ row+col, data = subset(dat, year !=1921 ), 
            FUN = mean, na.rm = TRUE)
round(acast(dat3, row ~ col, value.var = 'yield'),0)

libs(desplot)
desplot(dat3, yield ~ col*row,
```

patterson.switchback

Switchback experiment on dairy cattle, milk yield for 4 treatments

Description

Switchback experiment on dairy cattle, milk yield for 4 treatments

Usage

data("patterson.switchback")

Format

A data frame with 36 observations on the following 4 variables.

y  response, milk FCM
trt treatment factor, 4 levels
period period factor, 3 levels
cow cow factor, 12 levels

Details

There are three periods. Each cow is assigned to one treatment cycle like T1-T2-T1, where T1 is
the treatment in period P1 and P3, and T2 is the treatment in period P2.
There are four treatments.
All 4*3 = 12 treatment cycles are represented.
Data were extracted from Lowry, page 70.

Source

Agricultural Experimental Station.

References

Examples

```r
## Not run:

library(agridat)
data(patterson.switchback)
dat <- patterson.switchback

# Create groupings for first treatment, second treatment
datp1 <- subset(dat, period=="P1")
datp2 <- subset(dat, period=="P2")
dat$p1trt <- datp1$trt[match(dat$cow, datp1$cow)]
dat$p2trt <- datp2$trt[match(dat$cow, datp2$cow)]

library(latticeExtra)
useOuterStrips(xyplot(y ~ period|p1trt*p2trt, data=dat,
group=cow, type=c("l", "r"),
auto.key=list(columns=5),
main="patterson.switchback",
xislab="First/Third period treatment",
ylab="Second period treatment"))

# Create a numeric period variable
dat$per <- as.numeric(substring(dat$period,2))

# Need to use 'terms' to preserve the order of the model terms
m1 <- aov(terms(y ~ cow + per:cow + period + trt, keep.order=TRUE), data=dat)
anova(m1) # Match table 2 of Lowry
```

## Analysis of Variance Table

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cow</td>
<td>11</td>
<td>3466.0</td>
<td>315.091</td>
<td>57.1773</td>
<td>2.258e-06 ***</td>
</tr>
<tr>
<td>cow:per</td>
<td>12</td>
<td>953.5</td>
<td>79.455</td>
<td>14.4182</td>
<td>0.0004017 ***</td>
</tr>
<tr>
<td>period</td>
<td>1</td>
<td>19.7</td>
<td>19.740</td>
<td>3.5821</td>
<td>0.0950382 .</td>
</tr>
<tr>
<td>trt</td>
<td>3</td>
<td>58.3</td>
<td>19.418</td>
<td>3.5237</td>
<td>0.0685092 .</td>
</tr>
<tr>
<td>Residuals</td>
<td>8</td>
<td>44.1</td>
<td>5.511</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## End(Not run)

---

**Description**

Apple tree yields for 6 treatments with covariate of previous yield.

**Format**

A data frame with 24 observations on the following 4 variables.
block block factor, 4 levels
trt treatment factor, 6 levels
prev previous yield in boxes
yield yield per plot

Details

Treatment 'S' is the standard practice in English apple orchards of keeping the land clean in the summer.
The previous yield is the number of boxes of fruit, for the four seasons previous to the application of the treatments.

Source


References


Examples

```r
## Not run:
library(agridat)
data(pearce.apple)
dat <- pearce.apple

libs(lattice)
xyplot(yield~prev|block, dat, main="pearce.apple", xlab="previous yield")

m1 <- lm(yield ~ trt + block + prev, data=dat)
# Predict values, holding the covariate at its overall mean of 8.3
newdat <- expand.grid(trt=c('A','B','C','D','E','S'), block=c('B1','B2','B3','B4'), prev=8.308333)
newdat$pred <- predict(m1, newdata=newdat)
# Average across blocks to get the adjusted mean, Booth et al. Table 1
tapply(newdat$pred, newdat$trt, mean)
# A   B   C   D   E   S
# 280.4765 266.5666 274.0666 281.1370 300.9175 251.3357

m2 <- lmer(yield ~ trt + (1|block) + prev, data=dat)
newdat$pred2 <- predict(m2, newdata=newdat)
```
pearl.kernels

Counts of yellow/white and sweet/starchy maize kernels by 15 observers

Description

Counts of yellow/white and sweet/starchy kernels on each of 4 maize ears by 15 observers.

Format

A data frame with 59 observations on the following 6 variables.

ear ear, 8-11
obs observer, 1-15
ys number of yellow starchy kernels
yt yellow sweet
ws white starchy
wt white sweet

Details

An ear of white sweet corn was crossed with an ear of yellow starchy corn. The F1 kernels of
the cross were grown and a sample of four ears was harvested. The F2 kernels of these ears were
classified by each of 15 observers into white/yellow and sweet/starchy.

By Mendelian genetics, the kernels should occur in the ratio 9 yellow starch, 3 white starch, 3
yellow sweet, 1 white sweet.

The observers had the following positions:

1 Plant pathologist
2 Asst plant pathologist
3 Prof agronomy
4 Asst prof agronomy
5 Prof philosophy
6 Biologist
7 Biologist
8 Asst biologist
9 Computer
10 Farmer
11 Prof plant physiology
pederson.lettuce.repeated

Repeated measurements of lettuce growth

Description
Repeated measurements of lettuce growth for 3 treatments.

Usage
data("pederson.lettuce.repeated")

Format
A data frame with 594 observations on the following 4 variables.

plant plant number
day day of observation
trt treatment
weight weight

Details
Experiment conducted in a greenhouse in Silver Bay, Minnesota. Plants were grown hydroponically. Treatment 1 had 9 plants per raft. Treatment 2 had 18 plants, treatment 3 had 36 plants. The response variable is weight of plant, roots, soil, cup, and water. The plants were measured repeatedly beginning Dec 1, and ending Jan 9, when the plants were harvested.

Source

References
None

Examples
library(agridat)
data(pederson.lettuce.repeated)
dat <- pederson.lettuce.repeated

libs(lattice)
dat <- dat[order(dat$day),]
xyplot(weight ~ day|trt, dat, type='l', group=plant, layout=c(3,1),
main="pederson.lettuce.repeated")

# Not run:
# Pederson used this SAS MIXED model for unstructured covariance
# proc mixed data=Project.Spacingdata;
# class trt plant day;
# model weight=trt day trt*day;
# repeated day / subject=plant type=un r rcorr;

# This should give the same results as SAS, but does not.
libs(nlme)
dat <- transform(dat, plant=factor(plant), day=factor(day))
datg <- groupedData(weight ~ day|plant, data=dat)
un1 <- gls(weight ~ trt * day, data=datg,
correlation=corSymm(value=rep(.6,55), form = ~ 1 | plant),
control=lmeControl(opt="optim", msVerbose=TRUE,
maxIter=500, msMaxIter=500))
logLik(un1)*2 # nlme has 1955, SAS had 1898.6

# Comparing the SAS results in Pederson (page 16) and the nlme results, we notice
# the SAS correlations in table 5.2 are unusually low for the first

Description

Usage
data("perry.springwheat")

Format
A data frame with 560 observations on the following 6 variables.

yield yield, kg/ha
gen genotype/cultivar factor, 28 levels
eenv environment factor, 20 levels
site site factor
year year, 1979-1982
yor year of release, 1860-1982

Details
Twenty-eight of the most significant wheat cultivars of the past century in Western Australia, were grown in 20 field trials over 4 years in the Central and Eastern wheat-belt of Australia.
At the Wongan Hills site there were separate early and late sown trials in 1979 and 1980. Later sowing dates generally have lower yields.
Note: Although not indicated by the original paper, it may be that the Merredin site in 1979 also had early/late sowing dates.
Used with permission of Mario D’Antuono and CSIRO Publishing.

Source
Examples

## Not run:

```r
library(agridat)
data(perry.springwheat)
dat <- perry.springwheat

libs(lattice)
xyplot(yield~yor|env, dat, type=c('p','r'), xlab="year of release",
main="perry.springwheat")

# Show a table of sites*year
# libs(latticeExtra)
# useOuterStrips(xyplot(yield~yor|site+factor(year), dat,
# type=c('p','r')))

# Perry reports a rate of gain of 5.8 kg/ha/year. No model is given.
# We fit a model with separate intercept/slope for each env
m1 <- lm(yield ~ env + yor + env:yor, data=dat)
# Average slope across environments
mean(c(coef(m1)[21], coef(m1)[21]+coef(m1)[22:40]))
## [1] 5.496781

# -------

# Now a mixed-effects model. Fixed overall int/slope. Random env int/slope.

# First, re-scale response so we don't have huge variances
dat$y <- dat$yield / 100

libs(lme4)
# Use || for uncorrelated int/slope. Bad model. See below.
# m2 <- lmer(y ~ 1 + yor + (1|yor||env), data=dat)
# Warning messages:
# 1: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :  
# Model failed to converge with max|grad| = 0.55842 (tol = 0.002, component 1)
# 2: In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
# Model is nearly unidentifiable: very large eigenvalue
# - Rescale variables?; Model is nearly unidentifiable: large eigenvalue ratio
# - Rescale variables?

# Looks like lme4 is having trouble with variance of intercepts
# There is nothing special about 1800 years, so change the
# intercept -- 'correct' yor by subtracting 1800 and try again.
dat$yorc <- dat$yor - 1800
m3 <- lmer(y ~ 1 + yorc + (1+yorc||env), data=dat)

# Now lme4 succeeds. Rate of gain is 100*0.0549 = 5.49
fixef(m3)
## (Intercept)  yorc
```
## 5.87492444 0.05494464

# asreml3 & asreml4
libs(asreml,lucid)
m3a <- asreml(y ~ 1 + yorc, data=dat, random = ~ env + env:yorc)
vc(m3)
## grp var1 var2 vcov sdcor
## env (Intercept) <NA> 11.61 3.407
## env.1 yorc <NA> 0.00063 0.02511
## Residual <NA> <NA> 3.551 1.884
vc(m3a)
## effect component std.error z.ratio con
## env!env.var 11.61 4.385 2.6 Positive
## env:yorc!env.var 0.00063 0.000236 2.7 Positive
## R!variance 3.551 0.2231 16 Positive

## End(Not run)

---

**petersen.sorghum.cowpea**

*Intercropping experiment of sorghum/cowpea*

**Description**

Intercropping experiment of sorghum/cowpea.

**Usage**

data("petersen.sorghum.cowpea")

**Format**

A data frame with 18 observations on the following 5 variables.

- **block** block
- **srows** sorghum rows
- **crows** cowpea rows
- **syield** sorghum yield, kg/ha
- **cyield** cowpea yield, kg/ha

**Details**

An intercropping experiment in Tanzania. The treatments consisted of four ratios of sorghum rows to cowpea rows as 1:4, 2:3, 3:2, 4:1.

The sole-crop yields with 5 rows per crop are also given (not part of the blocks).
piepho.barley.uniformity

Uniformity trial of barley

Description

Uniformity trial of barley in Germany

Usage

data("piepho.barley.uniformity")

Format

A data frame with 1080 observations on the following 5 variables.

row  row ordinate
col  column ordinate
yield  yield per plot
Details

Uniformity trial of barley at Ihinger Hof farm, conducted by the University of Hohenheim, Germany, in 2007.

Note: The paper by Piepho says "The trial had 30 rows and 36 columns. Plot widths were 1.90 m along rows and 3.73 m along columns." However, the SAS code supplement to the paper, called "PBR_1654_sm_example1.sas", has row=1-36, col=1-30. We cannot determine which dimension is "row" and which is "column", and therefore cannot determine the actual dimensions of the field.

Source


References

None

Examples

```r
## Not run:
data(piepho.barley.uniformity)
dat <- piepho.barley.uniformity
libs(desplot)
desplot(yield ~ col*row, dat,
tick=TRUE, # aspect unknown
       main="piepho.barley.uniformity.csv")

libs(asreml,dplyr)
dat <- mutate(dat, x=factor(col), y=factor(row))
dat <- arrange(dat, x, y)

# Piepho AR1xAR1 model (in random term, NOT residual)
m1 <- asreml(data=dat,
yield ~ 1,
             random = ~ x + y + ar1(x):ar1(y),
             residual = ~ units,
             na.action=na.method(x="keep") )
m1 <- update(m1)
# Match Piepho table 3, footnote 4: .9671, .9705 for col,row correlation
# Note these parameters are basically at the boundary of the parameter
# space. Questionable fit.
libs(lucid)
lucid::vc(m1)

## End(Not run)
```
piepho.cocksfoot

Multi-environment trial of cock’s foot, heading dates for 25 varieties in 7 years

Description

Multi-environment trial of cock’s foot, heading dates for 25 varieties in 7 years

Usage

data("piepho.cocksfoot")

Format

A data frame with 111 observations on the following 3 variables.

- gen: genotype factor, 25 levels
- year: year, numeric
- date: heading date (days from April 1)

Details

These data are heading dates (days from April 1 to heading) of 25 cock’s foot *Dactylis glomerata* varieties in trials at Hannover, Germany, repeated over seven years. Values are means over replications.

Piepho fits a model similar to Finlay-Wilkinson regression, but with genotype and environment swapped.

Source


Examples

```r
## Not run:
library(agridat)
data(piepho.cocksfoot)
dat <- piepho.cocksfoot
dat$year <- factor(dat$year)

library(lattice)
# Gaussian, not gamma distn
densityplot(~date|year, data=dat, main="piepho.cocksfoot - heading date")

library(mumm) # The mumm package can reproduce Piepho's results
```
Uniformity trial of safflower

Description
Uniformity trial of safflower at Farmington, Utah, 1962.

Usage
data("polson.safflower.uniformity")

Format
A data frame with 1716 observations on the following 3 variables.

row row
col column
yield yield (grams)
Details

A uniformity trial of safflower at the Utah State University field station in Farmington, Utah, in 1962. The field was approximately 0.5 acres in size, 110 x 189 feet. A four-row planter was used, 22 inches between rows. Four rows on either side and 12 feet on both ends were removed before harvesting.

Yield of threshed grain was recorded in grams.

Field width: (52 rows + 8 border rows) * 22 in = 110 ft
Field length: 33 sections * 5 ft + 2 borders * 12 ft = 189 ft

Source

David Polson. 1964. Estimation of Optimum Size, Shape, and Replicate Number of Safflower Plots for Yield Trials. Utah State University, All Graduate Theses and Dissertations, 2979. Table 6. https://digitalcommons.usu.edu/etd/2979

References

None.

Examples

```r
## Not run:

library(agridat)

data(polson.safflower.uniformity)
dat <- polson.safflower.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
    flip=TRUE, aspect=189/110, # true aspect
    main="polson.safflower.uniformity")

libs(agricolae)
libs(reshape2)
dmat <- acast(dat, row~col, value.var="yield")
# Similar to Polson fig 4.
tab <- index.smith(dmat, col="red",
    main="polson.safflower.uniformity - Smith Index",
    xlab="Plot size in number of basic plots")

# Polson p. 25 said CV decreased from 14.3 to 4.5
# for increase from 1 unit to 90 units. Close match.
tab <- data.frame(tab$uniformity)

# Polson only uses log(Size) < 2 in his Fig 5, obtained slope -0.63
coef(lm(log(Vx) ~ log(Size), subset(tab, Size <= 6))) # -0.70

# Polson table 2 reported labor for
# K1, number of plots, 133 hours 75
```
# K2, size of plot, 43.5 hours 24
# Optimum plot size
# X = b K1 / ((1-b) K2)
# Polson suggests optimum plot size 2.75 to 11 basic plots

## End(Not run)

---

**Description**

Onion yields for different densities at two locations

**Format**

This data frame contains the following columns:

- **density**: planting density (plants per square meter)
- **yield**: yield (g / plant)
- **loc**: location, Purnong Landing or Virginia

**Details**

Spanish white onions.

**Source**


**References**


**Examples**

```r
# Not run:
library(agridat)
data(ratkowsky.onions)
dat <- ratkowsky.onions

# Model inverse yield as a quadratic. Could be better...
libs(lattice)
dat <- transform(dat, iyield = 1/yield)
m1 <- lm(iyield ~ I(density^2)*loc, dat)
```
dat$pred <- predict(ml)

libs(latticeExtra)
foo <- xyplot(iyield ~ density, data=dat, group=loc, auto.key=TRUE,
     main="ratkowski.onions",ylab="Inverse yield")
foo + xyplot(pred ~ density, data=dat, group=loc, type='l')

## End(Not run)

---

**reid.grasses**

Yields of four grasses for a wide range of nitrogen fertilizer

**Description**

Yields of four grasses for a wide range of nitrogen fertilizer, conducted over 3 years.

**Usage**

data("reid.grasses")

**Format**

A data frame with 210 observations on the following 5 variables.

- nitro: nitrogen, 21 numeric levels
- year: Y1, Y2, or Y3
- gen: genotype
- drymatter: dry matter content
- protein: protein content

**Details**

Experiment at the Hannah Research Institute, Ayr.

Single plots were planted to 4 different kinds of grasses. Within each plot, 21 nitrogen treatments were randomized.

Reid modeled the dry matter yield with four-parameter logistic curves of the form $y = a - b \exp(-cx^d)$.

**Source**

D. Reid (1985). A comparison of the yield responses of four grasses to a wide range of nitrogen application rates. *J. Agric. Sci.*, 105, 381-387. Table 1 & 3. [https://doi.org/10.1017/S0021859600056434](https://doi.org/10.1017/S0021859600056434)

**References**

None
Examples

## Not run:

```r
library(agridat)
data(reid.grasses)
dat <- reid.grasses

libs(latticeExtra)
foo <- xyplot(drymatter + protein ~ nitro|year, dat, group=gen,
  auto.key=list(columns=4),
  as.table=TRUE, type=c('p','l'),
  main="reid.grasses",ylab="drymatter/protein trait value",
  scales=list(y=list(relation="free")))
combineLimits(foo)

# devtools::run_examples does NOT like groupedData
if (0){
  libs(nlme)
  dat2 <- dat
  dat2$indiv <- paste(dat$year, dat$gen) # individual year+genotype curves

  # use all data to get initial values
  inits <- getInitial(drymatter ~ SSfpl(nitro, A, B, xmid, scal), data = dat2)
inits
  ## A   B xmid   scal
  ## -4.167902 12.139796 68.764796 128.313106
  xvals <- 0:800
  y1 <- with(as.list(inits), SSfpl(xvals, A, B, xmid, scal))
  plot(drymatter ~ nitro, dat2)
  lines(xvals,y1)

  # must have groupedData object to use augPred
  dat2 <- groupedData(drymatter ~ nitro|indiv, data=dat2)
  plot(dat2)

  # without 'random', all effects are included in 'random'
m1 <- nlme(drymatter ~ SSfpl(nitro, A, B, xmid,scale),
    data= dat2,
    fixed= A + B + xmid + scale ~ 1,
    random = A + B + xmid + scale ~ 1|indiv,
    start=inits)
  fixef(m1)
  summary(m1)
  plot(augPred(m1, level=0:1),
    main="reid.grasses - observed/predicted data") # only works with groupedData object
} # if(0)

## End(Not run)
```

---

**ridout.appleshoots**

*Root counts for propagated columnar apple shoots.*
Description

Root counts for propagated columnar apple shoots.

Usage

data("ridout.appleshoots")

Format

A data frame with 270 observations on the following 4 variables.

roots number of roots per shoot
trtn number of shoots per treatment combination
photo photoperiod, 8 or 16
bap BAP concentration, numeric

Details

There were 270 micropropagated shoots from the columnar apple cultivar Trajan. During the rooting period, shoot tips of length 1.0-1.5 cm were cultured on media with different concentrations of the cytokinin BAP in two growth chambers with 8 or 16 hour photoperiod.

The response variable is the number of roots after 4 weeks at 22 degrees C.

Almost all of the shoots in the 8 hour photoperiod rooted. Under the 16 hour photoperiod only about half rooted.

High BAP concentrations often inhibit root formation of apples, but perhaps not for columnar varieties.

Used with permission of Martin Ridout.

Source


References

SAS. Fitting Zero-Inflated Count Data Models by Using PROC GENMOD. support.sas.com/rnd/app/examples/stat/GENMODZIP/roots.pdf

Examples

## Not run:

library(agridat)
data(ridout.appleshoots)
dat <- ridout.appleshoots

# Change photo and bap to factors
dat <- transform(dat, photo=factor(photo), bap=factor(bap))
# histogram(~roots, dat, breaks=0:18-0.5)

# For photo=8, Poisson distribution looks reasonable.
# For photo=16, half of the shoots had no roots
# Also, photo=8 has very roughly 1/45 as many zeros as photo=8,
# so we anticipate prob(zero) is about 1/45=0.22 for photo=8.

histogram(~roots|photo, dat, breaks=0:18-0.5, main="ridout.appleshoots")

# histogram(~roots|photo*bap, dat, breaks=0:18-0.5, type="density",
# xlab="Number of roots for photoperiod 8, 16",
# ylab="Density for BAP levels",
# main="ridout.appleshoots")

foo.obs <- histogram(~roots|photo*bap, dat, breaks=0:18-0.5, type="density",
                   xlab="Number of roots for photoperiod 8, 16",
                   ylab="Density for BAP levels",
                   main="ridout.appleshoots")

useOuterStrips(foo.obs)

# Ordinary (non-ZIP) Poisson GLM
m1 <- glm(roots ~ bap + photo + bap:photo, data=dat,
          family="poisson")
summary(m1) # Appears to have overdispersion

# ----- Fit a Zero-Inflated Poisson model -----

# Use SAS contrasts to match SAS output
oo <- options(contrasts=c('contr.SAS','contr.poly'))

# There are unequal counts for each trt combination, which obviously affects
# the distribution of counts, so use log(trtn) as an offset.
dat$ltrtn <- log(dat$trtn)

# Ordinary Poisson GLM: 1 + bap*photo.
# Zero inflated probability depends only on photoperiod: 1 + photo

m2 <- zeroinfl(roots ~ 1 + bap*photo | 1 + photo, data=dat,
               dist="poisson", offset=ltrtn)

logLik(m2) # -622.2283 matches SAS Output 1
-2 * logLik(m2) # 1244.457 Matches Ridout Table 2, ZIP, H*P, P
summary(m2) # Coefficients match SAS Output 3.

exp(coef(m2, "zero")) # Photo=8 has .015 times as many zeros as photo=16

# Get predicted _probabilities_

# Prediction data
newdat <- expand.grid(photo=c(8,16), bap=c(2.2, 4.4, 8.8, 17.6))
newdat <- aggregate(trtn~bap+photo, dat, FUN=mean)
newdat$ltrtn <- log(newdat$trtn)

# The predicted (Poisson + Zero) probabilities
d2 <- cbind(newdat[,c('bap','photo')], predict(m2, newdata=newdat, type="prob"))
robinson.peanut.uniformity

Uniformity trial of peanuts

Description

Uniformity trial of peanuts in North Carolina in 1939, 1940.

Usage

data("robinson.peanut.uniformity")

Format

A data frame with 1152 observations on the following 4 variables.

row row
col column
yield yield in grams/plot
year year

Details

Two crops of peanuts were grown in North Carolina in 1939 and 1940. A different field was used each year.

A block of 36 rows 3 feet wide and 200 feet long were harvested in 12.5 foot lengths.
Field length: 36 plots * 12.5 feet = 200 feet
Field width: 16 plots * 3 feet = 48 feet

Widening the plot was not as effective as increasing the plot length in order to reduce error. This agrees with the results of other uniformity studies.

Assuming 30 percent of the total cost of an experiment is proportional to the size of the plots used, the optimum plot size is approximately 3.2 units.
rothamsted.brussels

Source


References

None

Examples

```r
## Not run:

library(agridat)

data(robinson.peanut.uniformity)
dat <- robinson.peanut.uniformity

# Mean yield per year. Robinson has 703.9, 787.3
tapply(dat$yield, dat$year, mean)
# 1939 1940
# 703.7847 787.8125

libs(desplot)
desplot(dat, yield ~ col*row|year,
flip=TRUE, tick=TRUE, aspect=200/48,
main="robinson.peanut.uniformity")

## End(Not run)
```

rothamsted.brussels  

RCB experiment of brussels sprouts, 9 fertilizer treatments

Description

RCB experiment of brussels sprouts, 9 fertilizer treatments

Format

A data frame with 48 observations on the following 5 variables.

row row
col column
yield yield of saleable sprouts, pounds
trt treatment, 9 levels
block block, 4 levels
Details

The block numbers are arbitrary, and may not match the original source.
Plots were 10 yards x 14 yards. Plot orientation is not clear.

Source


References


Examples

```r
## Not run:

library(agridat)
data(rothamsted.brussels)
dat <- rothamsted.brussels

libs(lattice)
bwplot(yield~trt, dat, main="rothamsted.brussels")

libs(desplot)
desplot(dat, yield~col*row,
    num=trt, out1=block, cex=1, # aspect unknown
    main="rothamsted.brussels")

## End(Not run)
```

---

rothamsted.oats  \hspace{1cm} RCB experiment of oats, straw and grain, 9 fertilizer treatments

Description

RCB experiment of oats, straw and grain, 9 fertilizer treatments

Usage

data("rothamsted.oats")
rothamsted.oats

Format

A data frame with 96 observations on the following 6 variables.

- block: block
- trt: fertilizer treatment with 9 levels
- grain: grain, pounds per plot
- straw: straw, pounds per plot
- row: row
- col: column

Details

Oats (Grey Winter) grown at Rothamsted, Long Hoos field 1926.

Values of grain and straw are actual weights in pounds. Each plot was 1/40 acre. The plot dimensions are not given, but the Rothamsted report shows the field being square.

The treatment codes are: OA,OB,OC,OD = No top dressing. E/L = Early/late application. S/M = Sulphate or muriate of ammonia. 1/2 = Single or double dressing.

Source


References


Examples

```r
## Not run:

library(agridat)
data(rothamsted.oats)
dat <- rothamsted.oats

libs(desplot)
desplot(dat, grain~col*row,
   out1=block, text=trt, cex=1, shorten=FALSE,
   aspect=1,
   main="rothamsted.oats")
desplot(dat, straw~col*row,
   out1=block, text=trt, cex=1, shorten=FALSE,
   aspect=1,
   main="rothamsted.oats")
```
```r
# RCB experiment of groundnut, wet and dry yields

desc <- "RCB experiment of groundnut, wet and dry yields"

r <- readRDS(system.file("extdata", "ryder.groundnut.RData", package="agricolae"))

head(r, n=3)

## block row col straw grain
## 1 1 1 50.3 78.9
## 2 2 2 53.7 77.7
## 3 3 3 44.1 75.4
```

---

**Description**

RCB experiment of groundnut, wet and dry yields

**Format**

A data frame with 24 observations on the following 6 variables.

- block block
- row row
- col column
gen  genotype factor
wet  wet yield, kg/plot
dry  dry yield, kg/plot

Details
Ryder (1981) uses this data to discuss the importance of looking at the field plan for an experiment. Based on analysis of the residuals, he suggests that varieties A and B in block 3 may have had their data swapped.

Source
https://doi.org/10.1017/S0014479700011601

Examples

```r
# Not run:
library(agridat)
data(ryder.groundnut)
dat <- ryder.groundnut

# RCB model
m1 <- lm(dry~block+gen, dat)
dat$res1 <- resid(m1)

# Table 3 of Ryder. Scale up from kg/plot to kg/ha
round(dat$res1 * 596.6, 0)

# Visually. Note largest positive/negative residuals are adjacent
libs(desplot)
desplot(dat, res1 ~ col + row,
       text=gen, # aspect unknown
       main="ryder.groundnut - residuals")

libs(desplot)
# Swap the dry yields for two plots and re-analyze
dat[dat$block == "B3" & dat$gen == "A", "dry"] <- 2.8
dat[dat$block == "B3" & dat$gen == "B", "dry"] <- 1.4
m2 <- lm(dry~block+gen, dat)
dat$res2 <- resid(m2)
desplot(dat, res2 ~ col+row,
       # aspect unknown
       text=gen, main="ryder.groundnut")
```

## End(Not run)
salmon.bunt  

Fungus infection in varieties of wheat

Description

Fungus infection in varieties of wheat

Format

A data frame with 400 observations on the following 4 variables.

bunt  bunt factor, 20 levels
pct  percent infected
rep  rep factor, 2 levels
gen  genotype factor, 10 levels

Details

Note: Salmon (1938) gives results for all 69 types of bunt, not just the 20 shown in the paper.

H. A. Rodenhiser and C. S. Holton (1937) say that races from two different species of bunt were used, Tilletia tritici and T. levis.

This data gives the results with 20 types of bunt (fungus) for winter wheat varieties at Kearneysville, W. Va., in 1935. Altogether there were 69 types of bunt included in the experiment, of which the 20 in this data are representative. Each type of wheat was grown in a short row (5 to 8 feet), the seed of which had been inoculated with the spores of bunt. The entire seeding was then repeated in the same order.

Infection was recorded as a percentage of the total number of heads counted at or near harvest. The number counted was seldom less than 200 and sometimes more than 400 per row.

Source


References

Salmon says the data came from:

Examples

```r
## Not run:

library(agridat)
data(salmon.bunt)
dat <- salmon.bunt
d2 <- aggregate(pct~bunt+gen, dat, FUN=mean) # average reps
d2$gen <- reorder(d2$gen, d2$pct)
d2$bunt <- reorder(d2$bunt, d2$pct)
# Some wheat varieties (Hohenheimer) are resistant to all bunts, and some (Hybrid128)
# are susceptible to all bunts. Note the groups of bunt races that are similar,
# such as the first 4 rows of this plot. Also note the strong wheat*bunt interaction.
libs(lattice)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(pct~gen+bunt,d2, col.regions=redblue,
main="salmon.bunt percent of heads infected",
  xlab="Wheat variety", ylab="bunt line")

# We don't have individual counts, so use beta regression
libs(betareg)
dat$y <- dat$pct/100 + .001 # Beta regression does not allow 0
dat$gen <- reorder(dat$gen, dat$pct) # For a prettier dot plot
m1 <- betareg(y ~ gen + bunt + gen:bunt, data=dat)

# Construct 95 percent confidence intervals
p1 <- cbind(dat,
  lo = predict(m1, type='quantile', at=.025),
  est = predict(m1, type='quantile', at=.5),
  up = predict(m1, type='quantile', at=.975))
p1 <- subset(p1, rep=="R1")

# Plot the model intervals over the original data
libs(latticeExtra)
dotplot(bunt~y|gen, data=dat, pch='x', col='red',
  main="Observed data and 95 pct intervals for bunt infection") +
  segplot(bunt~lo+up|gen, data=p1, centers=est, draw.bands=FALSE)

# To evaluate wheat, we probably want to include bunt as a random effect...
```

## End(Not run)

---

**sawyer.multi.uniformity**

*Uniformity trials of wheat, swedes, oats, 3 years on the same land*
Description


Usage

data("sawyer.multi.uniformity")

Format

A data frame with 48 observations on the following 7 variables.

- year
- crop
- row
- col
- grain: wheat/oats grain weight, pounds
- straw: wheat/oats straw weight, pounds
- leafwt: swedes leaf weight, pounds
- rootwt: swedes root weight, pounds
- rootct: swedes root count

Details


Row 6, column 1 was not planted in any year.

Wheat was harvested in 1925. Row 1, column 1 had partially missing data for the wheat values in 1925 and was not used in the Rothamsted summary statistics on page 155.

Swedes were harvested in 1926.

Oats were harvested in 1927.

Note the summaries statistics at the bottom of the page in each report are calibrated to ACRES.

Field width: 8 plots * 22 feet = 528 feet

Field length: 6 plots * 22 feet = 396 feet

The field is 8 plots wide, 6 plots long. The plots are drawn in the source documents as squares 0.098 acres each (1 chain = 66 feet on each side).

Eden & Maskell (page 165) say the field was clover, and ploughed in the autumn of 1924. The field was laid out uniformly in lands of one chain width and each plot width made to coincide with the land width from ridge to ridge. The length of each plot was also one chain and from the point of view of yield data the trial comprised 47 plots in 8x6 except that the run of the hedge only allowed a rank of five plots at one of the ends.

Source


References


Examples

```r
## Not run:
library(agridat)
data("sawyer.multi.uniformity")
dat <- sawyer.multi.uniformity

libs(desplot)
# The field plan shows square plots
desplot(dat, grain~col*row,
  subset= year==1925,
  main="sawyer.multi.uniformity - 1925 wheat grain yield",
  aspect=(6)/(8)) # true aspect

desplot(dat, rootwt~col*row,
  subset= year==1926,
  main="sawyer.multi.uniformity - 1926 root weight of swedes",
  aspect=(6)/(8))

desplot(dat, grain~col*row, subset= year==1927,
  main="sawyer.multi.uniformity - 1927 oats grain yield",
  aspect=(6)/(8))

# This plot shows the "outlier" in the wheat data reported by Mackenzie.
libs(lattice)
xyplot(grain ~ straw, data=subset(dat, year==1925))

round(cor(dat[,7:9], use="pair"),2) # Matches McCullagh p 2121
## leafwt rootwt rootct
## leafwt 1.00 0.66 0.47
## rootwt 0.66 1.00 0.43
## rootct 0.47 0.43 1.00

## pairs(dat[,7:9],
## main="sawyer.multi.uniformity")
```

## End(Not run)
Description

Uniformity trial of sugarcane in India, 1933 & 1934.

Usage

data("sayer.sugarcane.uniformity")

Format

A data frame with 2056 observations on the following 4 variables.

- row row
- col column
- yield yield, pounds/plot
- year year

Details

1933 Experiment
First experiment was conducted in 1933 at Meghaul (Monghyr). A road was cut through the field, creating blocks 480 ft x 315 ft and 480 ft x 93 ft. (See Plate XLI). There were 136 rows, 3 feet apart, 480 feet long each. It required 16 days to harvest the 1088 plots. Each plot was 1/242 acre. The authors conclude that long narrow plots of 12/242 to 16/242 acre would be best.

Field width: 8 plots * 60 feet = 480 feet
Field length: 136 rows * 3 feet = 408 feet

1934 Experiment
This experiment was conducted at the New Area, Pusa. The experiment was laid out in 6 blocks, each separated by a 3-foot bund. The cutting of the canes began in Jan 1934, taking 24 days. (An earthquake 15 January delayed harvesting). Conclusion: Variation is reduced by increasing the plot size up to 9/242 acre.

Field width: 8 plots * 60 feet = 480 feet
Field length: 121 rows * 3 feet = 363 feet

Source

1933 Data
1934 data


References

None

Examples

```r
## Not run:
library(agridat)

data(sayer.sugarcane.uniformity)
dat33 <- subset(sayer.sugarcane.uniformity, year==1933)
dat34 <- subset(sayer.sugarcane.uniformity, year==1934)

b1 <- subset(dat33, row<31)
b2 <- subset(dat33, row > 30 & row < 61)
b3 <- subset(dat33, row > 60 & row < 91)
b4 <- subset(dat33, row > 105 & row < 136)
mean(b1$yield) # 340.7 vs Sayer 340.8
mean(b2$yield) # 338.2 vs Sayer 338.6
mean(b3$yield) # 331.3 vs Sayer 330.2
mean(b4$yield) # 295.4 vs Sayer 295.0

mean(dat34$yield) # 270.83 vs Sayer 270.83

libs(desplot)
desplot(dat33, yield ~ col*row, 
  flip=TRUE, aspect=408/480, # true aspect
  main="sayer.sugarcane.uniformity 1933")

desplot(dat34, yield ~ col*row, 
  flip=TRUE, aspect=363/480, # true aspect
  main="sayer.sugarcane.uniformity 1934")

## End(Not run)
```

senshu.rice

Multi-environment trial of rice, with solar radiation and temperature

Description

Response of rice to solar radiation and temperature
Format

A data frame with 40 observations on the following 7 variables.

country  country
loc  location
year  year of planting, last two digits
month  month of planting
rad  solar radiation
mint  minimum temperature
yield  yield t/ha

Details

Minimum temperature is the average across 30 days post flowering.
Opinion: Fitting a quadratic model to this data makes no sense.

Source


References


Examples

```r
## Not run:
library(agridat)
data(senshu.rice)dat <- senshu.rice

# Model 1 of Senshu & Cady
m1 <- lm(yield ~ 1 + rad + mint + I(mint^2), dat)
coef(m1)

# Use Fieller to calculate conf int around optimum minimum temp
# See: Piegorsch & Bailer, p. 31.
# Calculation derived from vegan::fieller.MOStest
m2 <- lm(yield ~ 1 + mint + I(mint^2), dat)
b1 <- coef(m2)[2]
b2 <- coef(m2)[3]
v <- vcov(m2)
sig1 <- vc[2,2]
sig2 <- vc[2,3]
sig3 <- vc[3,3]
u <- -b1/2/b2
```
shafi.tomato.uniformity <- qt(1-.05/2, nrow(dat)-3)
gam <- tval^2 * sig22 / b2^2
x <- u + gam * sig12 / (2 * sig22)
f <- tval / (-2*b2)
sq <- sqrt(sig11 + 4*u*sig12 + 4*u^2*sig22 - gam * (sig11 - sig12^2 / sig22) )
    ci <- (x + c(1,-1)*f*sq) / (1-gam)
    plot(yield ~ mint, dat, xlim=c(17, 32),
         main="senshu.rice: Quadratic fit and Fieller confidence interval",
        xlab="Minimum temperature", ylab="Yield")
    lines(17:32, predict(m2, new=data.frame(mint=17:32)))
    abline(v=ci, col="blue")

## End(Not run)

---

shafi.tomato.uniformity

Uniformity trial of tomato

Description

Uniformity trial of tomato in India.

Usage

data("shafi.tomato.uniformity")

Format

A data frame with 200 observations on the following 3 variables.

- row row ordinate
- col column ordinate
- yield yield, kg/plot

Details

The original data was collected on 1m x 1m plots. The data here are aggregated 2m x 2m plots.
Field length: 20 row * 2 m = 40 m
Field width: 10 col * 2 m = 20 m

Source

References


Examples

```r
## Not run:
library(agridat)
data(shafi.tomato.uniformity)
shafi.tomato.uniformity <- dat

libs(desplot)
desplot(dat, yield ~ col*row,
    aspect=40/20, # true aspect
    main="shafi.tomato.uniformity")

## End(Not run)
```

shafii.rapeseed  Multi-environment trial of rapeseed in U.S.

Description

Rapeseed yield multi-environment trial, 6 genotypes, 3 years, 14 loc, 3 rep

Format

A data frame with 648 observations on the following 5 variables.

- **year**: year, numeric: 87, 88, 89
- **loc**: location, 14 levels
- **rep**: rep, 3 levels
- **gen**: genotype, 6 levels
- **yield**: yield, kg/ha

Details

The data are from the U.S. National Winter Rapeseed trials conducted in 1986, 1987, and 1988. Trial locations included Georgia (GGA, TGA), Idaho (ID), Kansas (KS), Mississippi (MS), Montana (MT), New York (NY), North Carolina (NC), Oregon (OR), South Carolina (SC), Tennessee (TN), Texas (TX), Virginia (VA), and Washington (WA).

SAS codes for the analysis can be found at https://webpages.uidaho.edu/cals-statprog/ammi/index.html

Electronic version from: https://www.uiweb.uidaho.edu/ag/statprog/ammi/yld.data

Used with permission of Bill Price.
shafii.rapeseed

Source


References


Examples

```r
## Not run:
library(agridat)
data(shafii.rapeseed)
dat <- shafii.rapeseed
dat$gen <- with(dat, reorder(gen, yield, mean))
dat$loc <- with(dat, reorder(loc, yield, mean))
dat$yield <- dat$yield/1000

dat <- transform(dat, rep=factor(rep), year=as.factor(as.character(year)))
dat$locyr = paste(dat$loc, dat$year, sep="")

# The 'means' of reps
datm <- aggregate(yield~gen+year+loc+locyr, data=dat, FUN=mean)
datm <- datm[order(datm$gen),]
datm$gen <- as.character(datm$gen)
datm$gen <- factor(datm$gen, levels=c("Bienvenu","Bridger","Cascade", "Dwarf","Glacier","Jet"))
dat$locyr <- reorder(dat$locyr, dat$yield, mean)

libs(lattice)
# This picture tells most of the story
dotplot(loc~yield|gen,group=year,data=dat,
auto.key=list(columns=3),
par.settings=list(superpose.symbol=list(pch = c('7','8','9'))),
main="shafii.rapeseed",ylab="Location")

# AMMI biplot. Remove gen and locyr effects.
ml.lm <- lm(yield ~ gen + locyr, data=datm)
datm$res <- resid(ml.lm)
# Convert to a matrix
libs(reshape2)
dm <- melt(datm, measure.var="res", id.var=c('gen', 'locyr'))
dmat <- acast(dm, gen~locyr)
# AMMI biplot. Figure 1 of Shafii (1998)
biplot(prcomp(dmat), main="shafii.rapeseed - AMMI biplot")
```
Description

Multi-environment trial of oats in India, 13 genotypes, 3 year, 2 loc, 5 reps

Usage

data("shaw.oats")

Format

A data frame with 390 observations on the following 5 variables.

- env  environment, 2 levels
- year year, 3 levels
- block block, 5 levels
- gen genotype variety, 13 levels
- yield yield of oats, pounds per plot

Details

An oat trial in India of 11 hybrid oats compared to 2 established high-yielding varieties, labeled L and M. The trial was conducted at 2 locations. The size and exact locations of the plots varied from year to year.

At Pusa, the crop was grown without irrigation. At Karnal the crop was given 2-3 irrigations. Five blocks were used, each plot 1000 square feet. In 1932, variety L was high-yielding at Pusa, but low-yielding at Karnal.

Shaw used this data to illustrate ANOVA for a multi-environment trial.

Source


References

None
Examples

```r
library(agridat)
data(shaw.oats)
dat <- shaw.oats
# sum(dat$yield) # 16309 matches Shaw p. 125
# sum( (dat$yield-mean(dat$yield))^2) # total SS matches Shaw p. 141

dat$year <- factor(dat$year)
libs(lattice)
dotplot(yield ~ gen|env, data=dat, groups=year,
   main="shaw.oats",
   par.settings=list(superpose.symbol=list(pch=c(2,3,4))),
   panel=function(x,y,...){
   panel.dotplot(x,y,...)
   panel.superpose(x,y,..., panel.groups=function(x,y,col.line,...) {
   dd<-aggregate(y~x,data.frame(x,y),mean)
   panel.xyplot(x=dd$x, y=dd$y, col=col.line, type="l")
   })},
   auto.key=TRUE)

# Shaw & Bose meticulously calculate the ANOVA table, p. 141
m1 <- aov(yield ~ year*env*block*gen - year:env:block:gen, dat)
anova(m1)
```

silva.cotton

Number of cotton bolls for different levels of defoliation.

Description

Number of cotton bolls, nodes, plant height, and plant weight for different levels of defoliation.

Usage

```r
data("silva.cotton")
```

Format

A data frame with 125 observations on the following 4 variables.

- stage  growth stage
- defoliation  level of defoliation, 0, 25, 50, 75, 100
- plant  plant number
- rep  replicate
reproductive number of reproductive structures
bolls number of bolls
height plant height
nodes number of nodes
weight weight of bolls

Details
Data come from a greenhouse experiment with cotton plants. Completely randomized design with 5 replicates, 2 plants per pot.
Artificial defoliation was used at levels 0, 25, 50, 75, 100 percent.
Data was collected per plant at five growth stages: vegetative, flower-bud, blossom, fig and cotton boll.
The primary response variable is the number of bolls. The data are counts, underdispersed, correlated.
Zeviana et al. used this data to compared Poisson, Gamma-count, and quasi-Poisson GLMs.
Bonat & Zeviani used this data to fit multivariate correlated generalized linear model.
Used with permission of Walmes Zeviani.
Electronic version from: https://www.leg.ufpr.br/~walmes/data/desfolha_algodao.txt

Source

References

Examples
```r
## Not run:
library(agridat)
data(silva.cotton)
dat <- silva.cotton
dat$stage <- ordered(dat$stage,
```
levels=c("vegetative","flowerbud","blossom","boll","bollopen"))

# make stage a numeric factors
dat <- transform(dat,
       stage = factor(stage, levels = unique(stage),
       labels = 1:nlevels(stage)))

# sum data across plants, 1 pot = 2 plants
dat <- aggregate(cbind(weight,height,bolls,nodes) ~
       stage+defoliation+rep, data=dat, FUN=sum)

# all traits, plant-level data
libs(latticeExtra)
foo <- xyplot(weight + height + bolls + nodes ~
       stage, data = dat, outer=TRUE,
       xlab="Defoliation percent", ylab="", main="silva.cotton",
       as.table = TRUE, jitter.x = TRUE, type = c("p", "smooth"),
       scales = list(y = "free"))
combineLimits(useOuterStrips(foo))

if(0){

# poisson glm with quadratic effect for defoliation
m0 <- glm(bolls ~ 1, data=dat, family=poisson)
m1 <- glm(bolls ~ defoliation+I(defoliation^2), data=dat, family=poisson)
m2 <- glm(bolls ~ stage:defoliation+I(defoliation^2), data=dat, family=poisson)
m3 <- glm(bolls ~ stage:(defoliation+I(defoliation^2)), data=dat, family=poisson)
par(mfrow=c(2,2)); plot(m3); layout(1)
anova(m0, m1, m2, m3, test="Chisq")

# predicted values
preddat <- expand.grid(stage=levels(dat$stage),
        defoliation=seq(0,100,length=20))
preddat$pred <- predict(m3, newdata=preddat, type="response")

# Zeviani figure 3
libs(latticeExtra)
xyplot(bolls ~ jitter(defoliation)|stage, dat,
       as.table=TRUE,
       main="silva.cotton - observed and model predictions",
       xlab="Defoliation percent",
       ylab="Number of bolls") +
xyplot(pred ~ defoliation|stage, data=preddat,
       as.table=TRUE,
       type='smooth', col="black", lwd=2)
}

if(0){
# ----- mcglm -----
dat <- transform(dat, deffac=factor(defoliation))

libs(car)
vars <- c("weight","height","bolls","nodes")
splom(~dat[vars], data=dat,
groups = stage,
  auto.key = list(title = "Growth stage",
                  cex.title = 1,
                  columns = 3),
  par.settings = list(superpose.symbol = list(pch = 4)),
  as.matrix = TRUE)

splom(~dat[vars], data=dat,
  groups = defoliation,
  auto.key = list(title = "Artificial defoliation",
                  cex.title = 1,
                  columns = 3),
  as.matrix = TRUE)

# multivariate linear model.
ml <- lm(cbind(weight, height, bolls, nodes) ~ stage * deffac,
          data = dat)
anova(ml)
summary.aov(ml)

r0 <- residuals(ml)

# Checking the models assumptions on the residuals.
car::scatterplotMatrix(r0,
                        gap = 0, smooth = FALSE, reg.line = FALSE, ellipse = TRUE,
                        diagonal = "qqplot")

}

## End(Not run)

---

### sinclair.clover

*Clover yields in a factorial fertilizer experiment*

**Description**

Clover yields in a factorial fertilizer experiment

**Usage**

```r
data("sinclair.clover")
```

**Format**

A data frame with 25 observations on the following 3 variables.

```r
yield    yield t/ha
```
Details

A phosphorous by sulfur factorial experiment at Dipton in Southland, New Zealand. There were 3 reps. Plots were harvested repeatedly from Dec 1992 to Mar 1994. Yields reported are the total dry matter across all cuttings.

Source


References


Examples

library(agridat)

data(sinclair.clover)
dat <- sinclair.clover

libs(lattice)

xyplot(yield~P|factor(S), dat, layout=c(5,1),
       main="sinclair.clover - Yield by sulfur levels",
       xlab="Phosphorous")

# Dodds fits a two-dimensional Mitscherlich-like model:
# z = a*(1+b*(s+t*x)/(x+1))^y) * (1+d*((th+r*y)/(y+1))^x)

# First, re-scale the problem to a more stable part of the parameter space
dat <- transform(dat, x=P/10, y=S/10)

z0m <- 5
zm0 <- 5
zmm <- 10.5

# The parameters are somewhat sensitive to starting values.
# I had to try a couple different initial values to match the paper by Dodds
m1 <- nls(yield ~ alpha*(1 + beta*{(sig+tau*x)/(x+1)}^y) * (1 + del*{(th+rho*y)/(y+1})^x),
          data=dat, # trace=TRUE,
          start=list(alpha=zmm, beta=(zm0/zmm)-1, del=(z0m/zmm)-1,
                     sig=.51, tau=.6, th=.5, rho=.7))

summary(m1) # Match Dodds Table 2
## Parameters:
## Estimate  Std. Error   t value Pr(>|t|)
## alpha 11.15148  0.66484  16.773 1.96e-12 ***
## beta -0.61223 0.03759 -16.286 3.23e-12 ***
## del -0.48781 0.04046 -12.057 4.68e-10 ***
## sig 0.26783 0.16985 1.577 0.13224
## tau 0.68030 0.06333 10.741 2.94e-09 ***
## th 0.59656 0.16716 3.569 0.00219 **
## rho 0.83273 0.06204 13.421 8.16e-11 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.5298 on 18 degrees of freedom

# Not run:

pred <- expand.grid(x=0:17, y=0:9)
pred$z <- predict(m1, pred)
# 3D plot of data with fitted surface. Matches Dodds figure 2.
libs(rgl)
bg3d(color = "white")
clear3d()
spheres3d(dat$x, dat$y, dat$yield,
  radius=.2, col = rep("navy", nrow(dat)))
surface3d(seq(0, 17, by = 1), seq(0, 9, by = 1), pred$z,
  alpha=0.9, col="wheat",
  front="fill", back="fill")
axes3d()
title3d("sinclair.clover - yield", "", xlab="Phosphorous/10",
  ylab="Sulfur/10", zlab="", line=3, cex=1.5)
view3d(userMatrix=matrix(c(.7,.2,-.7,0, -.7,.2,-.6,0,0,.9,3,0, 0,0,0,1),ncol=4))
  # snapshot3d(file, "png")
rgl.close()

## End(Not run)

---

## Description

Uniformity trials of beans at California, 1954-1955, 2 species in 2 years

## Usage

data("smith.beans.uniformity")
smith.beans.uniformity

Format

A data frame with 912 observations on the following 4 variables.

- `expt` experiment
- `row` row
- `col` column
- `yield` yield, kg

Details

Trials were conducted in California.
In 1955 plots were twice as wide and twice as long as in 1954. Red Kidney is a bush variety bean, Standard Pink is a viny variety.

Smith randomly assigned A,B,C,D to plots and used these as 'varieties' for calculating ANOVA tables. Plots were combined side-by-side and end-to-end to make larger plots. Decreasing LSDs were observed for increases in plot sizes. LSDs were seldom above 200, which was considered to be a noticeable difference for the farmers.

There are four datasets:

1954

Experiment 1: Red Kidney. Experiment 2: Standard Pink
Field width: 18 plots * 30 inches = 45 ft
Field length: 12 plots * 15 ft = 180 ft

1955

Experiment 3: Red Kidney. Experiment 4: Standard Pink
Field width: 16 plots * 2 rows * 30 in = 80 ft
Field length: 15 plots * 30 ft = 450 ft

Source


References

None.

Examples

```r
## Not run:
library(agridat)

data(smith.beans.uniformity)
dat1 <- subset(smith.beans.uniformity, expt="E1")
dat2 <- subset(smith.beans.uniformity, expt="E2")
```
dat3 <- subset(smith.beans.uniformity, expt=="E3")
dat4 <- subset(smith.beans.uniformity, expt=="E4")

cv <- function(x) { sd(x)/mean(x) }
cv(dat1$yield)
cv(dat2$yield) # Does not match Smith. Checked all values by hand.
cv(dat3$yield)
cv(dat4$yield)

libs("desplot")
desplot(dat1, yield ~ col*row,
    aspect=180/45, flip=TRUE, # true aspect
    main="smith.beans.uniformity, expt 1 (true aspect)"
)
desplot(dat2, yield ~ col*row,
    aspect=180/45, flip=TRUE, # true aspect
    main="smith.beans.uniformity, expt 2 (true aspect)"
)
desplot(dat3, yield ~ col*row,
    aspect=450/80, flip=TRUE, # true aspect
    main="smith.beans.uniformity, expt 3 (true aspect)"
)
desplot(dat4, yield ~ col*row,
    aspect=450/80, flip=TRUE, # true aspect
    main="smith.beans.uniformity expt 4, (true aspect)"
)

## End(Not run)

---

**smith.corn.uniformity**  
*Uniformity trial of corn, 3 years on same ground*

**Description**

Uniformity trial of corn, 3 years on same ground, 1895-1897, in Illinois.

**Format**

A data frame with 360 observations on the following 5 variables.

- **row** row
- **col** column
- **plot** plot number, consistent across years
- **year** year. Last two digits of 1895, 1896, 1897
- **yield** yield, bushels / acre
Details

Data come from the Illinois Experiment Station.

The data values are from Smith (1910) and the field map is from Harris (1920). Each plot was 1/10 acre, but the dimensions are not given. Note that 1/10 acre is also the area of a square 1 chain (66 feet) on a side.

The following text is abridged from Smith (1910).

How much variability may we reasonably expect in land that is apparently uniform? Some data among the records of the soil plots at the Illinois Experiment station furnish interesting material for study in this connection.

A field that had lain sixteen years in pasture was broken up in 1895 and laid out into plots to be subsequently used for soil experiments. The land is slightly rolling but otherwise quite uniform in appearance. There are in the series to be considered in this connection 120 one-tenth acre plots. These plots were all planted to corn for three consecutive years without any soil treatment, so that the records offer a rather exceptional opportunity for a study of this kind.

The yields of all these plots in terms of bushels of shelled corn per acre for the three years are given in the accompanying table.

A study of this data reveals some very striking variations. It will be noticed in the first place that there is a tremendous difference in production in the different years. The first year, 1895, was an extremely unfavorable one for corn and the yields are exceptionally low. The weather records show that the season was not only unusually dry, but also cool in the early part. The following year we have an exceptionally favorable corn season, and the yields run unusually high. The third year was also a good one, and the yields are perhaps somewhat above the normal for this locality.

It will be observed that certain plots appear to be very abnormal. Thus plots 117, 118, 119, and 120 give an abnormally high yield in the first season and an abnormally low one in the two following years. This is to be accounted for in the topography of the land. These plots lie in a low spot which was favorable in the dry year of 1895, but unfavorable in 1896 and 1897. For this reason these four plots were rejected from further consideration in this study, as were also plots 616, 617, 618, 619, and 620. This leaves 111 plots whose variations are apparently unaccounted for and which furnish the data from which the following results are taken.

It is noticeable that the variability as measured by the standard deviation becomes less in each succeeding year. This suggests the question as to whether continued cropping might not tend to induce uniformity. The records of a few of these plots which were continued in corn for three years longer, however, do not support such a conclusion.

It seems reasonable to expect greater variability in seasons very unfavorable for production, such as that of 1895, because so much may depend upon certain critical factors of production coming into play and this suggestion may be the explanation of the high standard deviation in this first year. Results extending over a longer series of years would be extremely interesting in this connection.

If we consider the total range of variation in any single year, we find differences as follows: Plots lying adjoining have shown the following maximum variations: 18 bushels in 1895; 11 bushels in 1896; 8 bushels in 1897.

The above results give us a conception of the unaccountable plot variations which we have to deal with in field tests. The possibility remains that a still closer study might detect some abnormal factors at play to account for these variations in certain cases, but the study certainly suggests the importance of conservatism in arriving at conclusions based upon plot tests.
The particular value that the writer has derived from this study is the strengthening of his conviction that the only dependence to be placed upon variety tests and other field experiments is from records involving the average of liberal numbers and extending over long periods of time.

Source

Smith, L.H. 1910. Plot arrangement for variety experiments with corn. Agronomy Journal, 1, 84–89. Table 1. https://books.google.com/books?id=mQT0AAAAAMAAJ&pg=PA84


Examples

```r
## Not run:
library(agridat)
data(smith.corn.uniformity)
dat <- smith.corn.uniformity
dat = transform(dat, year=factor(year))
libs(desplot)
desplot(dat, yield~col*row|year, 
  layout=c(2,2), aspect=1, 
  main="smith.corn.uniformity: yield across years 1895-1987")

## # Outliers are obvious
##libs(lattice)
## xyplot(yield~row|factor(col), dat, groups=year, 
##    auto.key=list(columns=3), main="smith.corn.uniformity")

libs(rgl)
# A few odd pairs of outliers in column 6
# black/gray dots very close to each other
plot3d(dat$col, dat$row, dat$yield, col=dat$year, 
  xlab="col",ylab="row",zlab="yield")
rgl.close()

## End(Not run)
```

---

**snedecor.asparagus**

*Asparagus yields for different cutting treatments*

**Description**

Asparagus yields for different cutting treatments, in 4 years.
**Format**

A data frame with 64 observations on the following 4 variables.

- **block**: block factor, 4 levels
- **year**: year, numeric
- **trt**: treatment factor of final cutting date
- **yield**: yield, ounces

**Details**

Planted in 1927. Cutting began in 1929. Yield is the weight of asparagus cuttings up to Jun 1 in each plot. Some plots received continued cuttings until Jun 15, Jul 1, and Jul 15.

In the past, repeated-measurement experiments like this were sometimes analyzed as if they were a split-plot experiment. This violates some independence assumptions.

**Source**


**References**


**Examples**

```r
## Not run:

library(agridat)
data(snedecor.asparagus)
dat <- snedecor.asparagus
dat <- transform(dat, year=factor(year))
dat$trt <- factor(dat$trt, levels=c("Jun-01", "Jun-15", "Jul-01", "Jul-15"))

# Continued cutting reduces plant vigor and yield
libs(lattice)
dotplot(yield ~ trt|year, data=dat,
       xlab="Cutting treatment", main="snedecor.asparagus")

# Split-plot
if(0){
  libs(lme4)
  ml <- lmer(yield ~ trt + year + trt:year +
              (1|block) + (1|block:trt), data=dat)
}

# ---------
```
libs(asreml,lucid) # asreml4

# Split-plot with asreml
m2 <- asreml(yield ~ trt + year + trt:year, data=dat,
             random = ~ block + block:trt)
# vc(m2)
## effect component std.error z.ratio bound
## block 354.3 405 0.87 P 0.1
## block:trt 462.8 256.9 1.8 P 0
## units!R 404.7 82.6 4.9 P 0

## # Antedependence with asreml. See O'Neill (2010).
dat <- dat[order(dat$block, dat$trt), ]
m3 <- asreml(yield ~ year * trt, data=dat,
             random = ~ block,
             residual = ~ block:trt:ante(year,1),
             max=50)

## # Extract the covariance matrix for years and convert to correlation
## covmat <- diag(4)
## covmat[upper.tri(covmat,diag=TRUE)] <- m3R.param$'block:trt:year'\$year$initial
## covmat[lower.tri(covmat)] <- t(covmat)[lower.tri(covmat)]
## round(cov2cor(covmat),2) # correlation among the 4 years
## # [1,] 1.00 0.45 0.39 0.31
## # [2,] 0.45 1.00 0.86 0.69
## # [3,] 0.39 0.86 1.00 0.80
## # [4,] 0.31 0.69 0.80 1.00

## # We can also build the covariance Sigma by hand from the estimated
## # variance components via: Sigma^-1 = U D^-1 U'
## vv <- vc(m3)
## print(vv)
## ## effect component std.error z.ratio constr
## ## block!block.var 86.56 156.9 0.55 pos
## ## R!variance 1 NA NA fix
## ## R!year.1930:1930 0.00233 0.00106 2.2 uncon
## ## R!year.1931:1930 -0.7169 0.4528 -1.6 uncon
## ## R!year.1931:1931 0.00116 0.00048 2.4 uncon
## ## R!year.1932:1931 -1.139 0.1962 -5.8 uncon
## ## R!year.1932:1932 0.00208 0.00085 2.4 uncon
## ## R!year.1933:1932 -0.6782 0.1555 -4.4 uncon
## ## R!year.1933:1933 0.00201 0.00083 2.4 uncon

## U <- diag(4)
## Dinv <- diag(c(vv[3,2], vv[5,2], vv[7,2], vv[9,2]))
## # solve(U
## solve(crossprod(t(U), tcrossprod(Dinv, U)) )
## ## [1,] 428.4310 307.1478 349.8152 237.2453
## ## [2,] 307.1478 1083.9717 1234.5516 837.2751
## ## [3,] 349.8152 1234.5516 1886.5150 1279.4378
## Description

Infection in wheat by different strains of Fusarium.

## Format

A data frame with 204 observations on the following 4 variables.

- `gen` wheat genotype
- `strain` fusarium strain
- `year` year
- `y` percent infected

## Details

The data are the percent of leaf area affected by Fusarium head blight, averaged over 4-5 reps, for 17 winter wheat genotypes.

Van Eeuwijk fit a generalized ammi-2 model to this data. It is a generalized model in the sense that a link function is used, and is a non-linear AMMI model in that there are main effects for variety and year-strain, but additional multiplicative effects for the interactions.

Note, the value for strain F348 in 1988, gen SVP75059-32 should be 28.3 (as shown in VanEeuwijk 1995) and not 38.3 (as shown in Snijders 1991).

Used with permission of Fred van Eeuwijk.

## Source

Snijders, CHA and Van Eeuwijk, FA. 1991. Genotype x strain interactions for resistance to Fusarium head blight caused by Fusarium culmorum in winter wheat. Theoretical and Applied Genetics, 81, 239–244. Table 1. https://doi.org/10.1007/BF00215729

## References

Examples

```r
library(agridat)

data(snijders.fusarium)
dat <- snijders.fusarium

aggregate(y ~ strain + year, dat, FUN=mean) # Match means in Snijders table 1

dat <- transform(dat, y=y/100, year=factor(year), yrstr=factor(paste0(year,"-",strain)))

# Strain F329 shows little variation across years. F39 shows a lot.
libs(lattice)
dotplot(gen~y|strain, data=dat, group=year,
   main="snijders.fusarium : infection by strain",
   xlab="Fraction infected", ylab="variety",
   auto.key=list(columns=3))

# Logit transform
dat <- transform(dat, logit=log(y/(1-y)))
m1 <- aov(logit ~ yrstr + gen, data=dat) # Match SS in VanEeuwijk table 4
anova(m1) # Match SS in VanEeuwijk table 4
m2 <- aov(logit ~ year*strain + gen + gen:year + gen:strain, data=dat)
anova(m2) # Match to VanEeuwijk table 5

## Not run:
# GLM on untransformed data using logit link, variance mu^2(1-mu)^2
libs(gnm) # for 'wedderburn' family
m2 <- glm(y ~ yrstr + gen, data=dat, family="wedderburn")
anova(m2) # Main effects match VanEeuwijk table 6

# Generalized AMMI-2 model. Matches VanEeuwijk table 6
bilin2 <- gnm(y ~ yrstr + gen + instances(Mult(yrstr, gen), 2),
   data=dat, family = wedderburn)
# plot(bilin2,1) # Resid vs fitted plot matches VanEeuwijk figure 3c
## anova(bilin2)
## Df Deviance Resid. Df Resid. Dev
## NULL 203 369.44
## yrstr 11 150.847 192 218.60
## gen 16 145.266 176 73.33
## Mult(yrstr, gen, inst = 1) 26 26.128 150 47.20
## Mult(yrstr, gen, inst = 2) 24 19.485 126 27.72

# Manually extract coordinates for biplot
cof <- coef(bilin2)
y1 <- cof[29:40]
g1 <- cof[41:57]
y2 <- cof[58:69]
g2 <- cof[70:86]
g12 <- cbind(g1,g2)
rownames(g12) <- substring(rownames(g12), 29)
```
Uniformity trial of sorghum silage

Description

Uniformity trial of sorghum silage at Chillicothe, Texas, 1915.

Format

A data frame with 2000 observations on the following 3 variables.

- row: row
- col: column / rod
- yield: yield, ounces

Details

Grown near Chillicothe, TX in 1915. Rows 40 inches apart. Each row harvested in 1-rod (16.5 ft) lengths. East side higher yielding than west side. Yields are weight (ounces) of green forage each rod-row. Total area harvested: 100*40/12 = 333.33 feet by 20*16.5=330 feet.

Field width: 20 plots * 16.5 ft (1 rod) = 330 feet.

Field length: 100 plots * 40 in = 333 feet

Source

Examples

```r
## Not run:
library(agridat)

data(stephens.sorghum.uniformity)
dat <- stephens.sorghum.uniformity

dat <- subset(dat, row>2 & row<99) # omit outer two rows
# mean(dat$yield) # 180.27
# range(dat$yield) # 75,302 matches Stephens

# densityplot(~dat$yield) # Stephens figure 3

# Aggregate 4 side-by-side rows.
d4 <- dat
d4$row2 <- ceiling((d4$row-2)/4)
d4 <- aggregate(yield ~ row2+col, data=d4, FUN=sum)
d4$row2 <- 25-d4$row2 # flip horizontally

libs(desplot)
grays <- colorRampPalette(c("#d9d9d9","#252525"))
desplot(d4, yield ~ row2*col,
     aspect=333/330, flip=TRUE, # true aspect
     main="stephens.sorghum.uniformity",
     col.regions=grays(3),
     at=c(500,680,780,1000))

# Similar to Stephens Figure 7. North at top. East at right.

## End(Not run)
```

Description

Phenotypic and genotypic data for a barley population of Steptoe x Morex. There were 150 doubled haploid crosses, evaluated at 223 markers. Phenotypic data was collected on 8 traits at 16 environments.

Usage

```r
data("steptoe.morex.pheno")
```
steptoe.morex.pheno

Format

steptoe.morex.pheno is a data.frame of phenotypic data with 2432 observations on 10 variables:

- **gen**: genotype factor with parents Steptoe and Morex, and 150 crosses SM1, SM2, ..., SM200. Not all 200 numbers were used.
- **env**: environment, 16 levels
- **amylase**: alpha amylase (20 Deg Units)
- **diapow**: diastatic power (degree units)
- **hddate**: heading date (julian days)
- **lodging**: lodging (percent)
- **malt**: malt extract (percent)
- **height**: plant height (centimeters)
- **protein**: grain protein (percent)
- **yield**: grain yield (Mt/Ha)

steptoe.morex.geno is a cross object from the qtl package with genotypic data of the 223 markers for the 150 crosses of Steptoe x Morex.

Details

As described by Hayes et al (1993), a population of 150 barley doubled haploid (DH) lines was developed by the Oregon State University Barley Breeding Program for the North American Barley Genome Mapping Project. The parentage of the population is Steptoe / Morex.

Steptoe is the dominant feed barley in the northwestern U.S.

Morex is the spring U.S. malting quality standard.

Seed from a single head of each parent was used to create the F1, from which a set of 150 lines was developed.

Phenotypic values for the parents Steptoe and Morex are here: https://wheat.pw.usda.gov/ggpages/SxM/parental_values.html

There are 16 locations. The average across locations is in column 17. Not all traits were collected at every location. At each location, all 150 lines were included in block 1, a random subset of 50 lines was used in block 2.

The traits are: Alpha Amylase (20 Deg Units), Diastatic Power (Deg Units), Heading Date (Julian Days), Lodging (percent), Malt Extract (percent), Grain Protein (percent), Grain Yield (Mt/Ha).

Phenotypic values of the 150 lines in the F1 population are here: https://wheat.pw.usda.gov/ggpages/SxM/phenotypes.html

Each trait is in a different file, in which each block of numbers represents one location.

The 223-markers Steptoe/Morex base map is here: https://wheat.pw.usda.gov/ggpages/SxM/smbasev2.map

The data for these markers on the 150 lines is https://wheat.pw.usda.gov/ggpages/SxM/smbasev2.mrk

These were hand-assembled (e.g. marker distances were cumulated to marker positions) into a .csv file which was then imported into R using qtl::read.cross. The class was manually changed from c('bc','cross') to c('dh','cross').

The marker data is coded as A = Steptoe, B = Morex, - = missing.

The pedigrees for the 150 lines are found here: https://wheat.pw.usda.gov/ggpages/SxM/pedigrees.html
Source


Data provided by the United States Department of Agriculture.

References


Examples

```r
# Not run:
library(agridat)
data(steptoe.morex.pheno)
dat <- steptoe.morex.pheno

# Visualize GxE of traits
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
levelplot(amyrase~env*gen, data=dat, col.regions=redblue,
scales=list(x=list(rot=90)), main="amyrase")
# levelplot(diapow~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="diapow")
# levelplot(hddate~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="hddate")
# levelplot(lodging~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="lodging")
# levelplot(malt~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="malt")
# levelplot(height~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="height")
# levelplot(protein~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="protein")
# levelplot(yield~env*gen, data=dat, col.regions=redblue,
## scales=list(x=list(rot=90)), main="yield")
```
# Calculate avg yield for each loc as in Romagosa 1996, table 3
# t(t(round(tapply(dat$yield, dat$env, FUN=mean),2)))
# SKo92,5Kg92 means in table 3 are switched. Who is right, him or me?

# Draw marker map
libs(qtl)
data(steptoe.morex.geno)
datg <- steptoe.morex.geno
plot.map(datg, main="steptoe.morex.pheno")  # or just use plot()

# This is a very rudimentary example.

# Fit a simple multi-environment mixed model
libs(asreml)
m1 <- asreml(yield ~ env, data=dat, random=~gen)

libs(wgaim)
qtl::plotMissing(datg)
wgaim::linkMap(datg)
# Create an interval object for wgaim
dati <- wgaim::cross2int(datg, id="gen")

# Whole genome qtl
q1 <- wgaim::wgaim(m1, intervalObj=dati, merge.by="gen", na.action=na.method(x="include"))
#wgaim::linkMap(q1, dati) # Visualize
wgaim::outStat(q1, dati) # outlier statistic
summary(q1, dati)  # Table of important intervals
# Chrom Left Marker dist(cM) Right Marker dist(cM) Size  Pvalue
# 3   ABG399     52.6  BCD828     56.1   0.254  0.000  45.0
# 5   MWG912     148   ABG387A   151.2   0.092  0.001   5.9
# 6   ABC169B    64.8   CDO497    67.5  -0.089  0.001   5.6

## End(Not run)
Format

A data frame with 1600 observations on the following 4 variables.

- `expt` experiment
- `row` row
- `col` col
- `yield` yield

Details


Experiment E1.
Field width = 20 units * 14 inches = 23.3 ft.
Field length = 20 units * 10 feet = 200 feet.

Experiment E2-E3.
Field width = 20 units * 44 inches = 73 feet
Field length = 20 units * 5 ft = 100 feet.

Source


References

None.

Examples

```r
# Not run:

library(agridat)

data(stickler.sorghum.uniformity)
dat <- stickler.sorghum.uniformity

dat1 <- subset(dat, expt=="E1")
dat2 <- subset(dat, expt!="E1")

lib(desplot)
desplot(dat, yield ~ col*row|expt, subset=expt=="E1",
  cex=1,text=yield, shorten="none",
  xlab="row",ylab="range",
  flip=TRUE, tick=TRUE, aspect=(20*10)/(20*14/12),
  main="stickler.sorghum.uniformity: expt E1")
```
stirret.borers

Corn borer control by application of fungal spores.

Description

Corn borer control by application of fungal spores.

Format

A data frame with 60 observations on the following 4 variables.

block block, 15 levels
trt treatment, 4 levels
count1 count of borers on August 18
count2 count of borers on October 19

Details

Experiment conducted in 1935, Ottawa. European corn borer infestation was established by application of egg masses to plants. Treatments were applied on July 8 and July 19 at two levels, 0 and 40 grams per acre. The number of borers per plot were counted on Aug 18 and Oct 19.

Source

Examples

```r
## Not run:
library(agridat)
data(stirret.borers)
dat <- stirret.borers

lib(lattice)
xypplot(count2~count1|trt, dat,
   main="stirret.borers - by treatment",
   xlab="Early count of borers", ylab="Late count")

# Even though the data are counts, Normal distribution seems okay
# qmath(~count1|trt, dat, main="stirret.borers")

m1 <- lm(count1 ~ -1 + trt + block, dat)
anova(m1)

# predicted means = main effect + average of 15 block effects
# note block 1 effect is 0
# coef(m1)[1:4] + sum(coef(m1)[-c(1:4)])/15
## trtBoth trtEarly trtLate trtNone
## 47.8667 62.93333 40.93333 61.13333

## End(Not run)
```

---

**streibig.competition**  
*Competition experiment between barley and sinapis.*

**Description**

Competition experiment between barley and sinapis, at different planting rates.

**Format**

A data frame with 135 observations on the following 8 variables.

- **pot** pot number
- **bseeds** barley seeds sown
- **sseeds** sinapis seeds sown
- **block** block
- **bfwt** barley fresh weight
- **sfwt** sinapis fresh weight
- **bdwt** barley dry weight
- **sdwt** sinapis dry weight
Details

The source data (in McCullagh) also contains a count of plants harvested (not included here) that sometimes is greater than the number of seeds planted.

Used with permission of Jens Streibig.

Source


References


Examples

```r
## Not run:

library(agridat)

data(streibig.competition)
dat <- streibig.competition

# See Schabenberger and Pierce, pages 370+
# Consider only the mono-species barley data (no competition from sinapis)
d1 <- subset(dat, sseeds<1)
d1 <- transform(d1, x=bseeds, y=bdwt, block=factor(block))

# Inverse yield looks like it will be a good fit for Gamma's inverse link
libs(lattice)
xyplot(1/y~x, data=d1, group=block, auto.key=list(columns=3),
       xlab="Seeding rate", ylab="Inverse yield of barley dry weight",
       main="streibig.competition")

# linear predictor is quadratic, with separate intercept and slope per block
m1 <- glm(y ~ block + block:x + x+I(x^2), data=d1,
           family=Gamma(link="inverse"))

# Predict and plot
newdf <- expand.grid(x=seq(0,120,length=50), block=factor(c('B1','B2','B3')))  
newdf$pred <- predict(m1, new=newdf, type='response')
plot(y~x, data=d1, col=block, main="streibig.competition - by block",
     xlab="Barley seeds", ylab="Barley dry weight")
for(bb in 1:3){
  newbb <- subset(newdf, block==c('B1','B2','B3')[bb])
  lines(pred~x, data=newbb, col=bb)
}
## End(Not run)
```
strickland.apple.uniformity

Uniformity trial in apple

Description
Uniformity trial in apple in Australia

Usage
data("strickland.apple.uniformity")

Format
A data frame with 198 observations on the following 3 variables.
row row
col column
yield yield per tree, pounds

Details
Some recently re-worked trees were removed from the data.
The distance between trees is uncertain, but likely in the range 20-30 feet.

Source

References
None

Examples
## Not run:
library(agridat)
data(strickland.apple.uniformity)
dat <- strickland.apple.uniformity
libs(desplot)
desplot(dat, yield ~ col*row,
       main="strickland.apple.uniformity",
       flip=TRUE, aspect=(18/11))

## End(Not run)
Description
Uniformity trial of grape in Australia

Usage
data("strickland.grape.uniformity")

Format
A data frame with 155 observations on the following 3 variables.
row  row
col  column
yield  yield per vine, pounds

Details
Yields of individual grape vines, planted 8 feet apart in rows 10 feet apart. Grown in Rutherglen, North-East Victoria, Australia, 1930.
Certain sections were omitted because of missing vines.

Source

References
None

Examples
## Not run:
library(agridat)
data(strickland.grape.uniformity)
dat <- strickland.grape.uniformity
libs(desplot)
desplot(dat, yield ~ col*row, 
  main="strickland.grape.uniformity", 
  flip=TRUE, aspect=(31*8)/(5*10) )
# CV 43.4
sd(dat$yield, na.rm=TRUE)/mean(dat$yield, na.rm=TRUE)

# anova like Strickland, appendix 1
anova(aov(yield ~ factor(row) + factor(col), data=dat))

# numbers ending in .5 much more common than .0
# table(substring(format(na.omit(dat$yield)),4,4))
#  0  5
# 25 100

## End(Not run)

---

**strickland.peach.uniformity**

*Uniformity trial of peach*

**Description**

Uniformity trial of peach trees in Australia.

**Usage**

data("strickland.peach.uniformity")

**Format**

A data frame with 144 observations on the following 3 variables.

- **row** row
- **col** column
- **yield** yield, pounds per tree

**Details**

Yields are the weight of peaches per individual tree in pounds.

**Source**


**References**

None
Examples

```r
## Not run:
library(agridat)

data(strickland.peach.uniformity)
dat <- strickland.peach.uniformity

mean(dat$yield) # 131.3, Strickland has 131.3
sd(dat$yield)/mean(dat$yield) # 31.1, Strickland has 34.4

libs(desplot)
desplot(dat, yield ~ col*row,
        main="strickland.peach.uniformity",
        flip=TRUE, aspect=1)

## End(Not run)
```

strickland.tomato.uniformity

*Uniformity trial of tomato*

Description

Uniformity trial of tomato in Australia

Usage

```r
data("strickland.tomato.uniformity")
```

Format

A data frame with 180 observations on the following 3 variables.

- **row** row
- **col** column
- **yield** yield per plot, pounds

Details

Tomato plants were placed 2 feet apart in rows 4 feet apart. Each plot contained 6 plants.

Field dimensions are not given, but the most likely design is:

- Field length: 6 plots * 6 plants * 2 feet = 72 feet
- Field width: 30 plots * 4 feet = 120 feet
Source


References

None

Examples

```r
## Not run:

library(agridat)
data(strickland.tomato.uniformity)
dat <- strickland.tomato.uniformity

mean(dat$yield)
sd(dat$yield)

libs(desplot)
desplot(dat, yield ~ col*row,
        main="strickland.tomato.uniformity",
        flip=TRUE, aspect=(6*12)/(30*4))

## End(Not run)
```

<table>
<thead>
<tr>
<th>stroup.nin</th>
<th>RCB experiment of wheat at the Nebraska Intrastate Nursery</th>
</tr>
</thead>
</table>

Description

The yield data from an advanced Nebraska Intrastate Nursery (NIN) breeding trial conducted at Alliance, Nebraska, in 1988/89.

Format

- gen genotype, 56 levels
- rep replicate, 4 levels
- yield yield, bu/ac
- col column
- row row
Details

Four replicates of 19 released cultivars, 35 experimental wheat lines and 2 additional triticale lines were laid out in a 22 row by 11 column rectangular array of plots. The varieties were allocated to the plots using a randomised complete block (RCB) design. Each plot was sown in four rows 4.3 m long and 0.3 m apart. Plots were trimmed down to 2.4 m in length before harvest. The orientation of the plots is not clear from the paper, but the data in Littel et al are given in meters and make the orientation clear.

Field length: 11 plots * 4.3 m = 47.3 m
Field width: 22 plots * 1.2 m = 26.4 m

All plots with missing data are coded as being gen = "Lancer". (For ASREML, missing plots need to be included for spatial analysis and the level of 'gen' needs to be one that is already in the data.) These data were first analyzed by Stroup et al (1994) and subsequently by Littell et al (1996, page 321), Pinheiro and Bates (2000, page 260), and Butler et al (2004).

This version of the data give the yield in bushels per acre. The yield values published in Stroup et al (1994) are expressed in kg/ha. For wheat, 1 bu/ac = 67.25 kg/ha.

Some of the gen names are different in Stroup et al (1994). (Sometimes an experimental genotype is given a new name when it is released for commercial use.) At a minimum, the following differences in gen names should be noted:

<table>
<thead>
<tr>
<th>stroup.nin</th>
<th>Stroup et al</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE83498</td>
<td>Rawhide</td>
</tr>
<tr>
<td>KS831374</td>
<td>Karl</td>
</tr>
</tbody>
</table>

Some published versions of the data use long/lat instead of col/row. To obtain the correct value of 'long', multiply 'col' by 1.2. To obtain the correct value of 'lat', multiply 'row' by 4.3.

Relatively low yields were clustered in the northwest corner, which is explained by a low rise in this part of the field, causing increased exposure to winter kill from wind damage and thus depressed yield. The genotype 'Buckskin' is a known superior variety, but was disadvantaged by assignment to unfavorable locations within the blocks.

Note that the figures in Stroup 2002 claim to be based on this data, but the number of rows and columns are both off by 1 and the positions of Buckskin as shown in Stroup 2002 do not appear to be quite right.

Source


References


**See Also**

Identical data (except for the missing values) are available in the `nlme` package as `Wheat2`.

**Examples**

```r
## Not run:
libraryagridat
data(stroup.nin)
dat <- stroup.nin

# Experiment layout. All "Buckskin" plots are near left side and suffer
# from poor fertility in two of the reps.
libs(desplot)
desplot(dat, yield~col*row,
    aspect=47.3/26.4, out1="rep", num=gen, cex=0.6, # true aspect
    main="stroup.nin - yield heatmap (true shape)")

# Dataframe to hold model predictions
preds <- data.frame(gen=levels(dat$gen))

# -----
# nlme
libs(nlme)
# Random RCB model
lme1 <- lme(yield ~ 0 + gen, random=-1|rep, data=dat, na.action=na.omit)
preds$lme1 <- fixef(lme1)

# Linear (Manhattan distance) correlation model
lme2 <- gls(yield ~ 0 + gen, data=dat,
    correlation = corLin(form = ~ col + row, nugget=TRUE),
    na.action=na.omit)
preds$lme2 <- coef(lme2)

# Random block and spatial correlation.
# Note: corExp and corSpher give nearly identical results
lme3 <- lme(yield ~ 0 + gen, data=dat,
    random = ~ 1 | rep,
    correlation = corExp(form = ~ col + row),
    na.action=na.omit)
preds$lme3 <- fixef(lme3)

# AIC(lme1,lme2,lme3) # lme2 is lowest
## df   AIC
## lme1 58 1333.702
## lme2 59 1189.135
```
## lme3

# *****
# SpATS
libs(SpATS)

dat <- transform(dat, yf = as.factor(row), xf = as.factor(col))

# what are colcode and rowcode???
sp1 <- SpATS(response = "yield",
             spatial = ~ SAP(col, row, nseg = c(10,20), degree = 3, pord = 2),
             genotype = "gen",
             #fixed = ~ colcode + rowcode,
             random = ~ yf + xf,
             data = dat,
             control = list(tolerance = 1e-03))
#plot(sp1)
preds$spats <- predict(sp1, which="gen")$predicted.value

# *****
# Template Model Builder
# See the ar1ar1 example:
# https://github.com/kaskr/adcomp/tree/master/TMB/inst/examples
# This example uses dpois() in the cpp file to model a Poisson response
# with separable AR1xAR1. I think this example could be used for the
# stroup.nin data, changing dpois() to something Normal.

# *****
# asreml4
libs(asreml,lucid)

# RCB analysis
as1 <- asreml(yield ~ gen, random = ~ rep, data=dat,
              na.action=na.method(x="omit"))
preds$asrem1 <- predict(as1, data=dat, classify="gen")$pvals$predicted.value

# Two-dimensional AR1xAR1 spatial model
dat <- transform(dat, xf=factor(col), yf=factor(row))
dat <- dat[order(dat$xf, dat$yf),]
as2 <- asreml(yield~gen, data=dat,
             residual = ~ar1(xf):ar1(yf),
             na.action=na.method(x="omit"))
preds$asrem2 <- predict(as2, data=dat, classify="gen")$pvals$predicted.value

lucid::vc(as2)
## effect component std.error z.ratio constr
## R!variance 48.7 7.155 6.8 pos
## R!xf.cor 0.6555 0.05638 12 unc
## R!yf.cor 0.4375 0.0806 5.4 unc
# Compare the estimates from the two asreml models.
# We see that Buckskin has correctly been shifted upward by the spatial model
plot(preds$as1, preds$as2, xlim=c(13,37), ylim=c(13,37),
     xlab="RCB", ylab="AR1xAR1", type='n')
title("stroup.nin: Comparison of predicted values")
text(preds$asreml1, preds$asreml2, preds$gen, cex=0.5)
abline(0,1)

# -----
# sommer
# Fixed gen, random row, col, 2D spline
libs(sommer)
dat <- transform(dat, yf = as.factor(row), xf = as.factor(col))
so1 <- mmer(yield ~ 0+gen,
             random = ~ vs(xf) + vs(yf) + vs(spl2D(row,col)),
             data=dat)
preds$so1 <- coef(so1)[,"Estimate"]
# spatPlot

# -----
# compare variety effects from different packages
lattice::splom(preds[,-1], main="stroup.nin")

## End(Not run)

---

**stroup.splitplot**

**Split-plot experiment of simulated data**

---

**Description**

A simulated dataset of a very simple split-plot experiment, used to illustrate the details of calculating predictable functions (broad space, narrow space, etc.).

For example, the density of narrow, intermediate and broad-space predictable function for factor level A1 is shown below (html help only)

**Format**

- **y** simulated response
- **rep** replicate, 4 levels
- **b** sub-plot, 2 levels
- **a** whole-plot, 3 levels

Used with permission of Walt Stroup.

**Source**

stroup.splitplot

References

Examples

```r
# Not run:
library(agridat)
data(stroup.splitplot)
dat <- stroup.splitplot

# ---- lme4 -----
# libs(lme4)
# m0 <- lmer(y~ -1 + a + b + a:b + (1|rep) + (1|a:rep), data=dat)
# No predict function

# ----- nlme -----  
# libs(nlme)
# m0 <- lme(y ~ -1 + a + b + a:b, data=dat, random = ~ 1|rep/a)

# ----- ASREML model -----  
libs(asreml)
m1 <- asreml(y~ -1 + a + b + a:b, random=~ rep + a:rep, data=dat)

libs(lucid)
# vc(m1) # Variance components match Stroup p. 41
# effect component std.error z.ratio bound
# rep 62.4 56.54 1.1 P 0
# a:rep 15.38 11.79 1.3 P 0
# units(R) 9.361 4.413 2.1 P 0

# Narrow space predictions
predict(m1, data=dat, classify="a", average=list(rep=NULL))
# a Predicted Std Err Status
# a1 32.88 1.082 Estimable
# a2 34.12 1.082 Estimable
# a3 25.75 1.082 Estimable

# Intermediate space predictions
predict(m1, data=dat, classify="a", ignore=list("a:rep"),
average=list(rep=NULL))
# a Predicted Std Err Status
# a1 32.88 2.24 Estimable
# a2 34.12 2.24 Estimable
# a3 25.75 2.24 Estimable

# Broad space predictions
predict(m1, data=dat, classify="a")
# a Predicted Std Err Status
# a1 32.88 4.54 Estimable
# a2 34.12 4.54 Estimable
```
# a3 25.75 4.54 Estimable

# ------ Mmcgllmm model ------
# Use the point estimates from REML with a prior distribution
libs(lattice,MCMCglmm)
prior2 = list(
  G = list(G1=list(V=62.40, nu=1), G2=list(V=15.38, nu=1)),
  R = list(V = 9.4, nu=1)
)
m2 <- MCMCglmm(y~ -1 + a + b + a:b,
  random=~ rep + a:rep, data=dat,
  pr=TRUE, # save random effects as columns of 'Sol'
nitt=23000, # double the default 13000
  prior=prior2, verbose=FALSE)

# Now create a matrix of coefficients for the prediction.
# Each column is for a different prediction. For example,
# the values in the column called 'a1a2n' are multiplied times
# the model coefficients (identified at the right side) to create
# the linear contrast for the the narrow-space predictions
# (also called adjusted mean) for the a1:a2 interaction.
# a1n  a1i  a1b  a1a2n  a1a2ib
cm <- matrix(c(1, 1, 1, 1, 1,  # a1
  0, 0, 0, -1, -1, # a2
  0, 0, 0, 0, 0,  # a3
  1/2, 1/2, 1/2, 0, 0,  # b2
  0, 0, 0, -1/2, -1/2, # a2:b2
  0, 0, 0, 0, 0,  # a3:b2
  1/4, 1/4, 0, 0, 0,  # r1
  1/4, 1/4, 0, 0, 0,  # r2
  1/4, 1/4, 0, 0, 0,  # r3
  1/4, 1/4, 0, 0, 0,  # r4
  1/4, 0, 0, 1/4, 0,  # a1r1
  0, 0, 0, -1/4, 0,  # a2r1
  0, 0, 0, 0, 0,  # a3r1
  1/4, 0, 0, 1/4, 0,  # a1r2
  0, 0, 0, -1/4, 0,  # a2r2
  0, 0, 0, 0, 0,  # a3r2
  1/4, 0, 0, 1/4, 0,  # a1r3
  0, 0, 0, -1/4, 0,  # a2r3
  0, 0, 0, 0, 0,  # a3r3
  1/4, 0, 0, 1/4, 0,  # a1r4
  0, 0, 0, -1/4, 0,  # a2r4
  0, 0, 0, 0, 0,  # a3r4
ncol=5, byrow=TRUE)
rownames(cm) <- c("a1", "a2", "a3", "b2", "a2:b2", "a3:b2",
  "r1", "r2", "r3", "r4",
  "a1r1", "a1r2", "a1r3", "a1r4", "a2r1", "a2r2",
  "a2r3", "a2r4", "a3r1", "a3r2", "a3r3", "a3r4")
print(cm)
# post2 <- as.mcmc(m2$Sol
post2 <- as.mcmc(crossprod(t(m2$Sol), cm))

# Following table has columns for A1 estimate (narrow, intermediate, broad)
# A1-A2 estimate (narrow and intermediate/broad).
# The REML estimates are from Stroup 1989.
est <- rbind("REML est"=c(32.88, 32.88, 32.88, -1.25, -1.25),
  "REML stderr"=c(1.08, 2.24, 4.54, 1.53, 3.17),
  "MCMC mode"=posterior.mode(post2),
  "MCMC stderr"=apply(post2, 2, sd))
round(est,2)

post22 <- lattice::make.groups(Narrow=post2[,1], Intermediate=post2[,2], Broad=post2[,3])
print(densityplot(~data|which, data=post22, groups=which, cex=.25, lty=1, layout=c(1,3),
xlab="MCMC model value of predictable function for A1"))

## End(Not run)

---

**student.barley**

**Multi-environment trial of barley**

**Description**

Yield for two varieties of barley grown at 51 locations in the years 1901 to 1906.

**Format**

A data frame with 102 observations on the following 7 variables.

- **year**: year, 1901-1906
- **farmer**: farmer name
- **place**: place (nearest town)
- **district**: district, geographical area
- **gen**: genotype, Archer and Goldthorpe
- **yield**: yield, 'stones' per acre (1 stone = 14 pounds)
- **income**: income per acre in shillings, based on yield and quality
Details

Experiments were conducted for six years by the Department of Agriculture in Ireland. A total of seven varieties were tested, but only Archer and Goldthorpe were tested in all six years (others were dropped after being found inferior, or were added later). Plots were two acres in size. The value of the grain depended on the yield and quality. Quality varied much from farm to farm, but not so much within the same farm.

The phrase "analysis of variance" first appears in the abstract (only) of a 1918 paper by Fisher. The 1923 paper by Student contained the first analysis of variance table (but not for this data).

One stone is 14 pounds. To convert lb/ac to tonnes/ha, multiply by 0.00112085116

Note: The analysis of Student cannot be reproduced exactly. For example, Student states that the maximum income of Goldthorpe is 230 shillings. A quick glance at Table I of Student shows that the maximum income for Goldthorpe is 220 shillings (11 pounds, 0 shillings) in 1901 at Thurles. Also, the results of Kempton could not be reproduced exactly, perhaps due to rounding or the conversion factor that was used.

Source


References


Examples

```r
## Not run:
library(agridat)
data(student.barley)
dat <- student.barley

libs(lattice)
bwplot(yield ~ gen|district, dat, main="student.barley - yield")

dat$year <- factor(dat$year)
dat$income <- NULL

# convert to tons/ha
dat <- transform(dat, yield=yield*14 * 0.00112085116)

# Define 'loc' the way that Kempton does
dat$loc <- rep("", nrow(dat))
dat[is.element(dat$farmer, c("Allardyce","Roche","Quinn")),"loc"] <- "1"
dat[is.element(dat$farmer, c("Luttrell","Dooley")), "loc"] <- "2"
dat[is.element(dat$year, c("1904","1905","1906")) & dat$farmer=="Kearney","loc"] <- "2"
dat[dat$farmer=="Mulhall","loc"] <- "3"

dat <- transform(dat, loc=factor(paste(place,loc,sep="")))
```
tai.potato

Multi-environment trial of potato

Description
Multi-environment trial of potato tuber yields

Usage
data("tai.potato")

Format
A data frame with 48 observations on the following 6 variables.
yield yield, kg/plot
gen genotype code
variety variety name
env environment code
loc location
year year
Details

Mean tuber yield of 8 genotypes in 3 locations over two years. Katahdin and Sebago are check varieties. Each location was planted as a 4-rep RCB design.

In Tai’s plot of the stability parameters, F5751 and Sebago were in the average stability area. The highest yielding genotype F6032 had an unstable performance.

Source


References


Examples

```r
library(agridat)
data(tai.potato)
dat <- tai.potato

libs(lattice)
dotplot(variety ~ yield|env, dat, main="tai.potato")

# fixme - need to add tai() example

# note, st4gi::tai assumes there are replications in the data
```

Description

Yield and 14 trait scores for each of 9 potato varieties at 12 locations in UK.

Usage

```r
data("talbot.potato.traits")
data("talbot.potato.yield")
```
### Format

The `talbot.potato.yield` dataframe has 126 observations on the following 3 variables.

- `gen` genotype/variety
- `trait` trait
- `score` trait score, 1-9

The `talbot.potato.yield` dataframe has 108 observations on the following 3 variables.

- `gen` genotype/variety
- `loc` location/center
- `yield` yield, t/ha

### Details

The `talbot.potato.yield` dataframe contains mean tuber yields (t/ha) of 9 varieties of potato at 12 centers in the United Kingdom over five years 1983-1987. The following abbreviations are used for the centers.

- BU  Bush
- CA  Cambridge
- CB  Conon Bridge
- CC  Crossacreve
- CP  Cockle Park
- CR  Craibstone
- GR  Greenmount
- HA  Harper Adams
- MO  Morley
- RO  Rosemaund
- SB  Sutton Bonnington
- TE  Terrington

Used with permission of Mike Talbot.

### Source


### Examples

```r
## Not run:
library(agridat)
library(pls, reshape2)
data(talbot.potato.traits)
```
datt <- talbot.potato.traits
data(talbot.potato.yield)
daty <- talbot.potato.yield
datt <- acast(datt, gen ~ trait, value.var='score')
daty <- acast(daty, gen ~ loc, value.var='yield')

# Transform columns to zero mean and unit variance
datt <- scale(datt)
daty <- scale(daty)

m1 <- plsr(daty ~ datt, ncomp=3)
summary(m1)

# Loadings factor 1
lo <- loadings(m1)[,1,drop=FALSE]
round(-1*lo[order(-1*lo),1,drop=FALSE],2)

biplot(m1, main="talbot.potato - biplot")

## End(Not run)

---

**theobald.barley**

*Multi-environment trial of barley, multiple years & fertilizer levels*

**Description**

Barley yields at multiple locs, years, fertilizer levels

**Usage**

```r
data("theobald.barley")
```

**Format**

A data frame with 105 observations on the following 5 variables.

- **yield**: yield, tonnes/ha
- **gen**: genotype
- **loc**: location, 5 levels
- **nitro**: nitrogen kg/ha
- **year**: year, 2 levels

**Details**

Theobald and Talbot used BUGS to fit a fully Bayesian model for yield response curves.

Locations of the experiment were in north-east Scotland.

Assumed nitrogen cost 400 pounds per tonne. Grain prices used were 100, 110, and 107.50 pounds per tonne for Georgie, Midas and Sundance.
Source


Data provided by Chris Theobald and Mike Talbot.

Examples

```r
library(agridat)

data(theobald.barley)
dat <- theobald.barley
dat <- transform(dat, env=paste(loc,year,sep="-"))
dat <- transform(dat, income=100*yield - 400*nitro/1000)

libs(lattice)
xyplot(income~nitro|env, dat, groups=gen, type="b",
auto.key=list(columns=3), main="theobald.barley")
```

theobald.covariate

Multi-environment trial of corn silage, Year * Loc * Variety with covariate

Description

Corn silage yields for maize in 5 years at 7 districts for 10 hybrids.

Format

A data frame with 256 observations on the following 5 variables.

- **year** year, 1990-1994
- **env** environment/district, 1-7
- **gen** genotype, 1-10
- **yield** dry-matter silage yield for corn
- **chu** corn heat units, thousand degrees Celsius

Used with permission of Chris Theobald.

Details

The trials were carried out in seven districts in the maritime provinces of Eastern Canada. Different fields were used in successive years. The covariate CHU (Corn Heat Units) is the accumulated average daily temperatures (thousands of degrees Celsius) during the growing season at each location.
Source


Examples

library(agridat)

data(theobald.covariate)
dat <- theobald.covariate
libs(lattice)
xyplot(yield ~ chu|gen, dat, type=c('p','smooth'),
xlab = "chu = corn heat units",
main="theobald.covariate - yield vs heat")

## Not run:
# REML estimates (Means) in table 3 of Theobald 2002
libs(lme4)
dat <- transform(dat, year=factor(year))
m0 <- lmer(yield ~ -1 + gen + (1|year/env) + (1|gen:year), data=dat)
round(fixef(m0),2)

## End(Not run)

## Not run:

# Use JAGS to fit Theobald (2002) model 3.2 with 'Expert' prior
libs(reshape2)
ymat <- acast(dat, year+env~gen, value.var='yield')
chu <- acast(dat, year+env~., mean, value.var='chu', na.rm=TRUE)
chu <- as.vector(chu - mean(chu)) # Center the covariate
dat$yr <- as.numeric(dat$year)
yridx <- as.vector(acast(dat, year+env~., mean, value.var='yr', na.rm=TRUE))
dat$loc <- as.numeric(dat$env)
locidx <- acast(dat, year+env~., mean, value.var='loc', na.rm=TRUE)
locidx <- as.vector(locidx)

jdat <- list(nVar = 10, nYear = 5, nLoc = 7, nYL = 29, yield = ymat,
chu = chu, year = yridx, loc = locidx)

libs(rjags)
m1 <- jags.model(file=system.file(package="agridat", "files/theobald.covariate.jag"),
data=jdat, n.chains=2)

# Table 3, Variety deviations from means (Expert prior)
c1 <- coda.samples(m1, variable.names=c('alpha'),
n.iter=10000, thin=10)
s1 <- summary(c1)
effs <- s1$statistics[, 'Mean']
Multi-environment trial of corn & soybean, 1930-1962, with temperature and precipitation

Description

Average yield of corn and soybeans in five U.S. states (IA, IL, IN, MO, OH) during the years 1930-1962. Pre-season precipitation and average temperature and precipitation during each month of the growing season is included.

Format

state  state
year  year, 1930-1962
rain0  pre-season precipitation in inches
temp5  may temperature, Fahrenheit
rain6  june rain, inches
temp6  june temp
rain7  july rain
temp7  july temp
rain8  august rain
temp8  august temp
corn  corn yield, bu/acre
soy  soybean yield, bu/acre

Details

Note: The Iowa corn data has sometimes been identified (in other sources) as the "Iowa wheat" data, but this is incorrect.

The ‘year’ variable affects yield through (1) improvements in plant genetics (2) changes in management techniques such as fertilizer, chemicals, tillage, planting date, and (3) climate, pest infestations, etc.

Double-cross corn hybrids were introduced in the 1920s. Single-cross hybrids became common around 1960.

During World War II, nitrogen was used in the production of TNT for bombs. After the war, these factories switched to producing ammonia for fertilizer. Nitrogen fertilizer use greatly increased after WWII and is a major reason for yield gains of corn. Soybeans gain little benefit from nitrogen
fertilizer. The other major reason for increasing yields in both crops is due to improved plant genetics.

Crops are often planted in May, and harvest begins in September.

Yields in 1936 were very low due to July being one of the hottest and driest on record.

Some relevant maps of yield, heat, and precipitation can be found in *Atlas of crop yield and summer weather patterns, 1931-1975*, https://www.isws.illinois.edu/pubdoc/C/ISWSC-150.pdf

The following notes pertain to the Iowa data.

The 1947 June precipitation of 10.33 inches was the wettest June on record (a new Iowa June record of 10.34 inches was set in 2010). As quoted in *Monthly Weather Review* (Dec 1957, p. 396) "The dependence of Iowa agriculture upon the vagaries of the weather was closely demonstrated during the 1947 season. A cool wet spring delayed crop planting activity and plant growth; then, in addition, a hard freeze on May 29th ... further set back the corn. The heavy rains and subsequent floods during June caused appreciable crop acreage to be abandoned ... followed by a hot dry weather regime that persisted from mid-July through the first week of September."

In 1949 soybean yields were average while corn yields were low. From the same source above, "The year 1949 saw the greatest infestation of corn borer in the history of corn in Iowa".

1955 yields were reduced due to dry weather in late July and August.

**Source**


**References**


**Examples**

```r
library(agridat)

data(thompson.cornsoy)
dat <- thompson.cornsoy

# The droughts of 1934/36 were severe in IA/MO. Less so in OH.
libs(lattice)
xyplot(corn+soy~year|state, dat,
    type=c('p','l','r'), auto.key=list(columns=2),
    main="thompson.cornsoy",
    layout=c(5,1),ylab='yield')

# In 1954, only Missouri suffered very hot, dry weather
## xyplot(corn~year, dat,
##     groups=state, type=c('p','l'),
##     main="thompson.cornsoy",
##     auto.key=list(columns=5), ylab='corn yield')
```
# Rain and temperature have negative correlation in each month.
# July is a critical month: temp and yield are negatively correlated,
# while rain and yield are positively correlated.
# splom(~dat[-1,-1], col=dat$state, cex=.5, main="thompson.cornsoy")

## Not run:
# Plots similar to those in Venables' Exegeses paper.

dat.ia <- subset(dat, state=="Iowa")

libs(splines)
m2 <- aov(corn ~ ns(rain0, 3) + ns(rain7, 3) +
ns(temp8, 3) + ns(year,3), dat.ia)
op <- par(mfrow=c(2,2))
termplot(m2, se=TRUE, rug=TRUE, partial=TRUE)
par(op)

# do NOT use gam package
libs(mgcv)
m1 <- gam(corn ~ s(year, k=5) + s(rain0, k=5) +
s(rain7, k=5) + s(temp8, k=5), data=dat.ia)
op <- par(mfrow=c(2,2))
plot.gam(m1, residuals=TRUE, se=TRUE, cex=2)
par(op)

## End(Not run)

---

### Description

Herbicide control of larkspur

### Usage

```r
data("turner.herbicide")
```

### Format

A data frame with 12 observations on the following 4 variables.

- **rep**: rep factor
- **rate**: rate of herbicide
- **live**: number of live plants before application
- **dead**: number of plants killed by herbicide
Details

Effectiveness of the herbicide Picloram on larkspur plants at 4 doses (0, 1.1, 2.2, 4.5) in 3 reps. Experiment was done in 1986 at Manti, Utah.

Source


References

Christopher Bilder, Thomas Loughin. Analysis of Categorical Data with R.

Examples

```r
## Not run:
library(agridat)
data(turner.herbicide)
dat <- turner.herbicide
dat <- transform(dat, prop=dead/live)
# xyplot(prop~rate,dat, pch=20, main="turner.herbicide", ylab="Proportion killed")

m1 <- glm(prop~rate, data=dat, weights=live, family=binomial)
coef(m1) # -3.46, 2.6567 Same as Turner eqn 3

# Make conf int on link scale and back-transform
p1 <- expand.grid(rate=seq(0,to=5,length=50))
p1 <- cbind(p1, predict(m1, newdata=p1, type='link', se.fit=TRUE))
p1 <- transform(p1, lo = plogis(fit - 2*se.fit),
              fit = plogis(fit),
              up = plogis(fit + 2*se.fit))

# Figure 2 of Turner
libs(latticeExtra)
foo1 <- xyplot(prop~rate,dat, cex=1.5,
            main="turner.herbicide (model with 2*S.E.)",
            xlab="Herbicide rate", ylab="Proportion killed")
foo2 <- xyplot(fit~rate, p1, type='l')
foo3 <- xyplot(lo+up~rate, p1, type='l', lty=1, col='gray')
print(foo1 + foo2 + foo3)

# What dose gives a LD90 percent kill rate?
# libs(MASS)
# dose.p(m1, p=.9)
## Dose           SE
## p = 0.9: 2.12939 0.128418

# Alternative method
# libs(car) # logit(.9) = 2.197225
```
What is a 95 percent confidence interval for LD90? Bilder & Loughin page 138

```r
root <- function(x, prob=.9, alpha=0.05){
  co <- coef(m1) # b0, b1
  covs <- vcov(m1) # b00, b11, b01
  # .95 = b0 + b1*x
  # (b0+b1*x) + Z(alpha/2) * sqrt(b00 + x^2*b11 + 2*x*b01) > .95
  # (b0+b1*x) - Z(alpha/2) * sqrt(b00 + x^2*b11 + 2*x*b01) < .95
  f <- abs(co[1] + co[2]*x - log(prob/(1-prob))) / #
    sqrt(covs[1,1] + x^2 * covs[2,2] + 2*x*covs[1,2])
  return( f - qnorm(1-alpha/2))
}
lower <- uniroot(f=root, c(0,2.13))
upper <- uniroot(f=root, c(2.12, 5))
c(lower$root, upper$root)
# 1.92 2.45

## End(Not run)
```

### urquhart.feedlot

**Weight gain calves in a feedlot**

**Description**

Weight gain calves in a feedlot, given three different diets.

**Usage**

```r
data("urquhart.feedlot")
```

**Format**

A data frame with 67 observations on the following 5 variables.

- **animal**: animal ID
- **herd**: herd ID
- **diet**: diet: Low, Medium, High
- **weight1**: initial weight
- **weight2**: slaughter weight

**Details**

Calves born in 1975 in 11 different herds entered a feedlot as yearlings. Each animal was fed one of three diets with low, medium, or high energy. The original sources explored the use of some contrasts for comparing breeds.
<table>
<thead>
<tr>
<th>Herd</th>
<th>Breed</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>New Mexico Herefords</td>
</tr>
<tr>
<td>16</td>
<td>New Mexico Herefords</td>
</tr>
<tr>
<td>3</td>
<td>Utah State University Herefords</td>
</tr>
<tr>
<td>32</td>
<td>Angus</td>
</tr>
<tr>
<td>24</td>
<td>Angus x Hereford (cross)</td>
</tr>
<tr>
<td>31</td>
<td>Charolais x Hereford</td>
</tr>
<tr>
<td>19</td>
<td>Charolais x Hereford</td>
</tr>
<tr>
<td>36</td>
<td>Charolais x Hereford</td>
</tr>
<tr>
<td>34</td>
<td>Brangus</td>
</tr>
<tr>
<td>35</td>
<td>Brangus</td>
</tr>
<tr>
<td>33</td>
<td>Southern Select</td>
</tr>
</tbody>
</table>

**Source**


**References**


Also available in the 'emmeans' package as the 'feedlot' data.

**Examples**

```r
## Not run:

library(agridat)
data(urquhart.feedlot)
dat <- urquhart.feedlot

library(reshape2)
d2 <- melt(dat, id.vars=c('animal', 'herd', 'diet'))

library(latticeExtra)
useOuterStrips(xyplot(value ~ variable|diet*herd, data=d2, group=animal,
                      type='l',
                      xlab="Initial & slaughter timepoint for each diet",
                      ylab="Weight for each herd",
                      main="urquhart.feedlot - weight gain by animal"))

# simple fixed-effects model
dat <- transform(dat, animal = factor(animal), herd=factor(herd))
m1 <- lm(weight2 ~ weight1 + herd*diet, data = dat)
coef(m1) # weight1 = 1.1373 match Urquhart table 5 common slope

# random-effects model might be better, for example
# libs(lme4)
```
usgs.herbicides

Concentrations of herbicides in streams in the United States

Description

Concentrations of selected herbicides and degradation products determined by laboratory method analysis code GCS for water samples collected from 51 streams in nine Midwestern States, 2002

Usage

data("usgs.herbicides")

Format

A data frame with 184 observations on the following 19 variables.

- mapnum  map number
- usgsid  USGS ID
- long    longitude
- lat     latitude
- site    site name
- city    city
- sampletype  sample type code
- date    date sample was collected
- hour    hour sample was collected
- acetochlor concentration as character
- alachlor concentration as character
- ametryn concentration as character
- atrazine concentration as character
- CIAT concentration as character
- CEAT concentration as character
- cyanazine concentration as character
- CAM concentration as character
- dimethenamid concentration as character
- flufenacet concentration as character
Details
Concentrations of selected herbicides and degradation products determined by laboratory method analysis code GCS for water samples collected from 51 streams in nine Midwestern States, 2002.
All concentrations are micrograms/liter, "<" means "less than". The data are in character format to allow for "<".
The original report contains data for more herbicides. This data is for illustrative purposes.
Sample types: CR = concurrent replicate sample, FB = field blank, LD = laboratory duplicate, S1 = sample from pre-emergence runoff, S2 = sample from post-emergence runoff, S3 = sample from harvest-season runoff.

Source

References
None.

Examples
### Not run:
```r
library(agridat)
data(usgs.herbicides)
dat <- usgs.herbicides

libs(NADA)
# create censored data for one trait
dat$y <- as.numeric(dat$atrazine)
dat$ycen <- is.na(dat$y)
dat$y[is.na(dat$y)] <- .05

# percent censored
with(dat, censummary(y, censored=ycen))
# median/mean
with(dat, cemle(y, ycen, dist="lognormal"))
# boxplot
with(dat, cenboxplot(obs=y, cen=ycen, log=FALSE))
# with(dat, boxplot(y))
pp <- with(dat, ros(obs=y, censored=ycen, forwardT="log")) # default lognormal
plot(pp)

plotfun <- function(vv){
dat$y <- as.numeric(dat[[vv]])
dat$ycen <- is.na(dat$y)
```
vaneeuwijk.drymatter

Multi-environment trial of maize, dry matter content

Description

Multi-environment trial of maize, dry matter content

Usage

data("vaneeuwijk.drymatter")

Format

A data frame with 168 observations on the following 5 variables.

- year
- site
- variety
- y

Details

Percent dry matter is given.

Site codes are soil type classifications: SS=Southern Sand, CS=Central Sand, NS=Northern Sand, RC=River Clay.

These data are a balanced subset of the data analyzed in van Eeuwijk, Keizer, and Bakker (1995b) and Kroonenberg, Basford, and Ebskamp (1995).

Used with permission of Fred van Eeuwijk.

Source

References


Examples

```r
## Not run:

library(agridat)
data(vaneeuwijk.drymatter)
dat <- vaneeuwijk.drymatter
dat <- transform(dat, year=factor(year))
dat <- transform(dat, env=factor(paste(year,site)))

libs(HH)
HH::interaction2wt(y ~ year+site+variety,dat,rot=c(90,0),
                   x.between=0, y.between=0,
                   main="vaneeuwijk.drymatter")

# anova model
m1 <- aov(y ~ variety+env+variety:env, data=dat)
anova(m1) # Similar to VanEeuwijk table 2
m2 <- aov(y ~ year*site*variety, data=dat)
anova(m2) # matches Sahai table 5.5

# variance components model
libs(lme4)
libs(lucid)
m3 <- lmer(y ~ (1|year) + (1|site) + (1|variety) +
           (1|year:site) + (1|year:variety) + (1|site:variety),
           data=dat)
vc(m3) # matches Sahai page 266
##  grp    var1    var2    vcov   sdcor
## year:variety (Intercept) <NA>  0.3187  0.5645
## year:site (Intercept) <NA>   7.735  2.781
## site:variety (Intercept) <NA>  0.83502 0.1871
## year (Intercept) <NA>    6.272  2.504
## variety (Intercept) <NA>   0.4867  0.6976
## site (Intercept) <NA>     6.594  2.55
## Residual <NA> <NA>  0.8885  0.9426

## End(Not run)
```
Description

Infection of wheat varieties by Fusarium strains from 1990 to 1993

Usage

data("vaneeuwijk.fusarium")

Format

A data frame with 560 observations on the following 4 variables.

- year: year, 1990-1993
- strain: strain of fusarium
- gen: genotype/variety
- y

Details

Data come from Hungary. There were 20 wheat varieties infected with 7 strains of Fusarium in the years 1990-1993. The measured value is a rating of the severity of disease due to Fusarium head blight, expressed as a number 1-100.

Three-way interactions for varieties 21 and 23 were the only ones in 1992 suffering from strain infections. This was due to incorrect storage of the inoculum (strain) which rendered it incapable of infecting most other varieties.

The data is a subset of the data analyzed by VanEeuwijk et al. 1995.

Used with permission of Fred van Eeuwijk.

Source


References

Examples

```r
## Not run:

library(agridat)
data(vaneeuwijk.fusarium)
dat <- vaneeuwijk.fusarium
dat <- transform(dat, year=factor(year))
dat <- transform(dat, logity=log((y/100)/(1-y/100)))

libs(HH)
position(dat$year) <- c(3,9,14,19)
position(dat$strain) <- c(2,5,8,11,14,17,20)
HH::interaction2wt(logity ~ gen*year*strain, dat, rot=c(90,0),
x.between=0, y.between=0,
main="vaneeuwijk.fusarium")

# anova on logit scale. Near match to VanEeuwijk table 6
m1 <- aov(logity ~ gen*strain*year, data=dat)
anova(m1)
## Response: logity
## Df Sum Sq Mean Sq  F value Pr(>F)
## gen  19 157.55   8.292  107.331  < 2e-16 ***
## strain  6  91.54  15.256  107.331  < 2e-16 ***
## year  3 321.99 107.331  107.331  < 2e-16 ***
## gen:strain 114  34.03   0.299    0.262
## gen:year  57 140.94   2.473    0.003 *
## strain:year 18 236.95  13.164    0.011 *
## gen:strain:year 342  93.15   0.272

## End(Not run)
```

vaneeuwijk.nematodes  
Number of cysts on 11 potato genotypes for 5 potato cyst nematode populations.

Description

The number of cysts on 11 potato genotypes for 5 potato cyst nematode populations.

Usage

```r
data("vaneeuwijk.nematodes")
```

Format

A data frame with 55 observations on the following 3 variables.

- `gen` potato genotype
vaneeuwijk.nematodes

pop  nematode population
y   number of cysts

Details

The number of cysts on 11 potato genotypes for 5 potato cyst nematode populations belonging to the species *Globodera pallida*. This is part of a larger table in . The numbers are the means over four or five replicates.

Van Eeuwijk used this data to illustrate fitting a generalized linear model.

Source


References


Examples

```r
library(agridat)
data(vaneeuwijk.nematodes)
dat <- vaneeuwijk.nematodes

# show non-normality
op <- par(mfrow=c(2,1), mar=c(5,4,3,2))
boxplot(y ~ pop, data=dat, las=2,
        ylab="number of cysts")
boxplot(y ~ gen, data=dat, las=2)
par(op)

## Not run:
# normal distribution
lm1 <- lm(y ~ gen + pop, data=dat)

# poisson distribution
glm1 <- glm(y ~ gen+pop, data=dat, family=quasipoisson(link=log))
anova(glm1)

libs(gnm)

# main-effects non-interaction model
gnm0 <- gnm(y ~ pop + gen, data=dat,
            family=quasipoisson(link=log))

# one interaction
gnm1 <- gnm(y ~ pop + gen + Mult(pop,gen,inst=1), data=dat,
```
family = quasipoisson(link = log))
# two interactions
gnm2 <- gnm(y ~ pop + gen + Mult(pop, gen, inst = 1) + Mult(pop, gen, inst = 2),
data = dat,
family = quasipoisson(link = log))

# anova(gnm0, gnm1, gnm2, test = "F")
# only 2, not 3 axes needed
# match vaneeuwijk table 2
# anova(gnm2)
## Df Deviance Resid. Df Resid. Dev
## NULL 54 8947.4
## pop 4 690.6 50 8256.8
## gen 10 7111.4 40 1145.4
## Mult(pop, gen, inst = 1) 13 716.0 27 429.4
## Mult(pop, gen, inst = 2) 11 351.1 16 78.3

# compare residual qq plots from models
op <- par(mfrow = c(2, 2))
plot(lm1, which = 2, main = "LM")
plot(glm1, which = 2, main = "GLM")
plot(gnm0, which = 2, main = "GNM, no interaction")
plot(gnm2, which = 2, main = "GNM, 2 interactions")
par(op)

# extract interaction-term coefficients, make a biplot
pops <- pickCoef(gnm2, ".[.pop")
gens <- pickCoef(gnm2, ".[.gen")
coefs <- coef(gnm2)
A <- matrix(coefs[pops], nc = 2)
B <- matrix(coefs[gens], nc = 2)
A2 = scale(A)
B2 = scale(B)
rownames(A2) <- levels(dat$pop)
rownames(B2) <- levels(dat$gen)
# near-match with vaneeuwijk figure 1
biplot(A2, B2, expand = 2.5, ylim = c(-2, 2), xlim = c(-2, 2),
main = "vaneeuwijk.nematodes - GAMMI biplot")

## End(Not run)
Usage

data("vargas.txe.covs")
data("vargas.txe.yield")

Format

The ‘vargas.txe.covs’ data has 10 years of measurements on 28 environmental covariates:

- `year`: Year
- `MTD`: Mean maximum temperature in December
- `MTJ`: Mean maximum temperature in January
- `MTF`: Mean maximum temperature in February
- `MTM`: Mean maximum temperature in March
- `MTA`: Mean maximum temperature in April
- `mTD`: Mean minimum temperature in December
- `mTJ`: Mean minimum temperature in January
- `mTF`: Mean minimum temperature in February
- `mTM`: Mean minimum temperature in March
- `mTA`: Mean minimum temperature in April
- `mTUD`: Mean minimum temperature in December
- `mTUJ`: Mean minimum temperature in January
- `mTUF`: Mean minimum temperature in February
- `mTUM`: Mean minimum temperature in March
- `mTUA`: Mean minimum temperature in April
- `PRD`: Total monthly precipitation in December
- `PRJ`: Total monthly precipitation in January
- `PRF`: Total monthly precipitation in February
- `PRM`: Total monthly precipitation in March
- `SHD`: Sun hours per day in December
- `SHJ`: Sun hours per day in January
- `SHF`: Sun hours per day in February
- `EVD`: Total monthly evaporation in December
- `EVJ`: Total monthly evaporation in January
- `EVF`: Total monthly evaporation in February
- `EVM`: Total monthly evaporation in March
- `EVA`: Total monthly evaporation in April

The ‘vargas.txe.yield’ dataframe contains 240 observations on three variables:

- `year`: Year
- `trt`: Treatment. See details section
- `yield`: Grain yield, kg/ha
Details

The treatment names indicate:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>deep knife</td>
</tr>
<tr>
<td>t</td>
<td>no deep knife</td>
</tr>
<tr>
<td>S</td>
<td>sesbania</td>
</tr>
<tr>
<td>s</td>
<td>soybean</td>
</tr>
<tr>
<td>M</td>
<td>chicken manure</td>
</tr>
<tr>
<td>m</td>
<td>no chicken manure</td>
</tr>
<tr>
<td>0</td>
<td>no nitrogen</td>
</tr>
<tr>
<td>n</td>
<td>100 kg/ha nitrogen</td>
</tr>
<tr>
<td>N</td>
<td>200 kg/ha nitrogen</td>
</tr>
</tbody>
</table>

Used with permission of Jose Crossa.

Source


Examples

```r
## Not run:
library(agridat)
data(vargas.txe.covs)
data(vargas.txe.yield)

libs(reshape2)
libs(lattice)
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
Z <- vargas.txe.yield
Z <- acast(Z, year ~ trt, value.var="yield")
levelplot(Z, col.regions=redblue,
          main="vargas.txe.yield", xlab="year", ylab="treatment",
          scales=list(x=list(rot=90)))

# Double-centered like AMMI
Z <- sweep(Z, 1, rowMeans(Z))
Z <- sweep(Z, 2, colMeans(Z))
# Vargas figure 1
biplot(prcomp(Z, scale.=FALSE), main="vargas.txe.yield")

# Now, PLS relating the two matrices
U <- vargas.txe.covs
U <- scale(U) # Standardized covariates
libs(pls)
m1 <- plsR(Z~U)
```
Vargas wheat1

Wheat yields in 7 years with genetic and environment covariates

Description

Yield of Durum wheat, 7 genotypes, 6 years, with 16 genotypic variates and 16 environment variates.

Usage

data("vargas.wheat1.covs")
data("vargas.wheat1.traits")

Format

The vargas.wheat1.covs dataframe has 6 observations on the following 17 variables.

- **year**: year, 1990-1995
- **MTD**: Mean daily max temperature December, deg C
- **MTJ**: January
- **MTF**: February
- **MTM**: March
- **mTD**: Mean daily minimum temperature December, deg C
- **mTJ**: January
- **mTF**: February
- **mTM**: March
- **PRD**: Monthly precipitation in December, mm
- **PRJ**: January
- **PRF**: February
- **PRM**: March
- **SHD**: a numeric vector
- **SHJ**: January
- **SHF**: February
- **SHM**: March

The vargas.wheat1.traits dataframe has 126 observations on the following 19 variables.

- **year**: year, 1990-1995
rep  replicate, 3 levels
gen  genotype, 7 levels
yield yield, kg/ha
ANT  anthesis, days after emergence
MAT  maturity, days after emergence
GFI  grainfill, MAT-ANT
PLH  plant height, cm
BIO  biomass above ground, kg/ha
HID  harvest index
STW  straw yield, kg/ha
NSM  spikes / m^2
NGM  grains / m^2
NGS  grains per spike
TKW  thousand kernel weight, g
WTI  weight per tiller, g
SGW  spike grain weight, g
VGR  vegetative growth rate, kg/ha/day, STW/ANT
KGR  kernel growth rate, mg/kernel/day

Details
Conducted in Ciudad Obregon, Mexico.

Source
Data provided by Jose Crossa.

Examples

```r
library(agridat)

## Not run:
data(vargas.wheat1.covs)
data(vargas.wheat1.traits)
libs(pls)
libs(reshape2)

# Yield as a function of non-yield traits
Y0 <- vargas.wheat1.traits[,c('gen','rep','year','yield')]
```
Y0 <- acast(Y0, gen ~ year, value.var='yield', fun=mean)
Y0 <- sweep(Y0, 1, rowMeans(Y0))
Y0 <- sweep(Y0, 2, colMeans(Y0)) # GxE residuals
Y0 <- scale(Y0) # scaled columns
X1 <- vargas.wheat1.traits[, -4] # omit yield
X1 <- aggregate(cbind(ANT,MAT,GFI,PLH,BIO,HID,STW,NSM,NGM,
                      NGS,TKW,WTI,SGW,VGR,KGR) ~ gen, data=X1, FUN=mean)
rownames(X1) <- X1$gen
X1$gen <- NULL
X1 <- scale(X1) # scaled columns
m1 <- plsr(Y1~X1)
loadings(m1)[,1,drop=FALSE] # X loadings in Table 1 of Vargas
biplot(m1, cex=.5, which="x", var.axes=TRUE,
       main="vargas.wheat1 - gen ~ trait") # Vargas figure 2a

# Yield as a function of environment covariates
Y2 <- t(Y0)
X2 <- vargas.wheat1.covs
rownames(X2) <- X2$year
X2$year <- NULL
Y2 <- scale(Y2)
X2 <- scale(X2)

m2 <- plsr(Y2~X2)
loadings(m2)[,1,drop=FALSE] # X loadings in Table 2 of Vargas

## End(Not run)

vargas.wheat2  Multi-environment trial of wheat with environmental covariates

Description

The yield of 8 wheat genotypes was measured in 21 low-humidity environments. Each environment had 13 covariates recorded.

Usage

data("vargas.wheat2.covs")
data("vargas.wheat2.yield")

Format

The 'vargas.wheat2.covs' data frame has 21 observations on the following 14 variables.

desc: environment
desc: length of growth cycle in days
mtc: mean daily minimum temperature in degrees Celsius
MTC  mean daily maximum temperature  
SHC  sun hours per day  
μTV  mean daily minimum temp during vegetative stage  
MTV  mean daily maximum temp during vegetative stage  
SHV  sun hours per day during vegetative stage  
μTS  mean daily minimum temp during spike growth stage  
MTS  mean daily maximum temp during spike growth stage  
SHS  sun hours per day during spike growth stage  
μTG  mean daily minimum temp during grainfill stage  
MTG  mean daily maximum temp during grainfill stage  
SHG  sun hours per day during grainfill stage  
The 'vargas.wheat2.yield' data frame has 168 observations on the following 3 variables.

env  environment  
gen  genotype  
yield  yield (kg/ha)

Details

Grain yields (kg/ha) for 8 wheat genotypes at 21 low-humidity environments grown during 1990-1994. The data is environment-centered and genotype-centered. The rows and columns of the GxE matrix have mean zero. The locations of the experiments were:

- OBD  Ciudad Obregon, Mexico, planted in December
- SUD  Wad Medani, Sudan
- TLD  Tlaltizapan, Mexico, planted in December
- TLF  Tlaltizapan, Mexico, planted in February
- IND  Dharwar, India
- SYR  Aleppo, Syria
- NIG  Kadawa, Nigeria

Source


Data provided by Jose Crossa.

Examples

```r
## Not run:
library(agridat)
libs(pls,reshape2)
```
data(vargas.wheat2.covs)
datc <- vargas.wheat2.covs

data(vargas.wheat2.yield)
daty <- vargas.wheat2.yield

# Cast to matrix
daty <- acast(daty, env ~ gen, value.var='yield')
rownames(datc) <- datc$env
datc$env <- NULL

# The pls package centers, but does not (by default) use scaled covariates
# Vargas says you should
# daty <- scale(daty)
datc <- scale(datc)
m2 <- plsr(daty ~ datc)

# Plot predicted vs observed for each genotype using all components
plot(m2)

# Loadings
# plot(m2, "loadings", xaxt='n')
# axis(1, at=1:ncol(datc), labels=colnames(datc), las=2)

# Biplots
biplot(m2, cex=.5, which="y", var.axes=TRUE, main="vargas.wheat2 ~ daty ~ datc") # Vargas figure 2a
biplot(m2, cex=.5, which="x", var.axes=TRUE) # Vectors form figure 2 b
# biplot(m2, cex=.5, which="scores", var.axes=TRUE)
# biplot(m2, cex=.5, which="loadings", var.axes=TRUE)

## End(Not run)

---

**verbyla.lupin**  
*Multi-environment trial of lupin, multiple varieties and densities*

## Description
Yield of 9 varieties of lupin at different planting densities across 2 years and multiple locations.

## Format
- **gen** genotype, 9 varieties
- **site** site, 11 levels
- **rep** rep, 2-3 levels
- **rate** seeding rate in plants/m^2
Details

Nine varieties of lupin were tested for yield response to plant density at 11 sites. The target density in 1991 was 10, 20, ..., 60 plants per m^2, and in 1992 was 20, 30, ..., 70 plants per m^2.

Plot dimensions are not given.

The variety Myallie was grown only in 1992.

Each site had 2 reps in 1991 and 3 reps in 1992. Each rep was laid out as a factorial RCB design; one randomization was used for all sites in 1991 and one (different) randomization was used for all sites in 1992. (This was confirmed with the principal investigator.)

In 1991 at the Mt. Barker location, the data for columns 5 and 6 was discarded due to problems with weeds.

Variety 'Myallie' was called '84L:439' in Verbyla 1997.

The year of release for the varieties is

<table>
<thead>
<tr>
<th>Variety</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unicrop</td>
<td>1973</td>
</tr>
<tr>
<td>Illyarrie</td>
<td>1979</td>
</tr>
<tr>
<td>Yandee</td>
<td>1980</td>
</tr>
<tr>
<td>Danja</td>
<td>1986</td>
</tr>
<tr>
<td>Gungurru</td>
<td>1988</td>
</tr>
<tr>
<td>Yorrel</td>
<td>1989</td>
</tr>
<tr>
<td>Warrah</td>
<td>1989</td>
</tr>
<tr>
<td>Merrit</td>
<td>1991</td>
</tr>
<tr>
<td>Myallie</td>
<td>1995</td>
</tr>
</tbody>
</table>

Data retrieved Oct 2010 from https://www.blackwellpublishers.co.uk/rss. (No longer available).

Used with permission of Blackwell Publishing.

Source


Examples

```r
## Not run:
library(agridat)
data(verbyla.lupin)
dat <- verbyla.lupin

# The same RCB randomization was used at all sites in each year
libs(desplot)
desplot(dat, gen~col+row|site,
out1=rep, num=rate,
# aspect unknown
main="verbyla.lupin - experiment design")

# Figure 3 of Verbyla
libs(lattice)
foo <- xyplot(yield ~ rate|loc*gen, data=dat, subset=year==92,
type=c('p','smooth'), cex=.5,
main="verbyla.lupin: 1992 yield response curves",
ylab="Yield (tons/ha)",
strip=strip.custom(par.strip.text=list(cex=.7)))

libs(latticeExtra) # for useOuterStrips
useOuterStrips(foo,
strip=strip.custom(par.strip.text=list(cex=.7)),
strip.left=strip.custom(par.strip.text=list(cex=.7)))

# ---------

libs(asreml,lucid)
# asreml 4

# We try to reproduce the analysis of Verbyla 1999.
# May not be exactly the same, but is pretty close.

# Check nlevels for size of random-coefficient structures
# length(with(dat, table(gen))) # 9 varieties for RC1
# length(with(dat, table(gen,site))) # 99 site:gen combinations for RC2

# Make row and col into factors
dat <- transform(dat, colf=factor(col), rowf=factor(row))
# sort for asreml
dat <- dat[order(dat$site, dat$rowf, dat$colf),]

# Make site names more useful for plots
```
# dat <- transform(dat, site=factor(paste0(year,".",substring(loc,1,4)))))

# Initial model from top of Verbyla table 9.
m0 <- asreml(yield ~ 1 + site + linrate + site:linrate, 
data = dat, 
random = ~ spl(rate) + dev(rate) + site:spl(rate) + site:dev(rate) + str(~gen+gen:linrate, ~us(2):id(9)) # RC1 
+ gen:spl(rate) + gen:dev(rate) + str(~site:gen+site:gen:linrate, ~us(2):id(99)) # RC2 
+ site:gen:spl(rate) + site:gen:dev(rate), 
residual = ~ dsum(~ ar1(rowf):ar1(colf)|site) # Spatial AR1 x AR1 
)
m0 <- update(m0) 
m0 <- update(m0) 
m0 <- update(m0) 
m0 <- update(m0) 
m0 <- update(m0) 

# Variograms match Verbyla 1999 figure 7 (scale slightly different)
plot(varioGram(m0), xlim=c(1:19), zlim=c(0,2), 
main="verbyla.lupin - variogram by site")

# Sequence of models in Verbyla 1999 table 10
m1 <- update(m0, fixed= ~ . + at(site, c(2,5,6,8,9,10)):lincol 
+ at(site, c(3,5,7,8)):linrow 
+ at(site, c(2,3,5,7,8,9,11)):serp, random = ~ . 
+ at(site, c(3,6,7,9)):rowf 
+ at(site, c(1,2,3,9,10)):colf 
+ at(site, c(5,7,8,10)):units)
m1 <- update(m1)

m2 <- update(m1, 
random = ~ ., 
- site:gen:spl(rate) - site:gen:dev(rate))

m3 <- update(m2, 
random = ~ ., 
- site:dev(rate) - gen:dev(rate))

m4 <- update(m3, 
random = ~ ., 
- dev(rate))
vold.longterm

535

m5 <- update(m4,
random = ~ .
- at(site, c(5,7,8,10)):units + at(site, c(5,7,8)):units)

# Variance components are a pretty good match to Verbyla 1997, table 15
libs(lucid)
vc(m5)
.001004/sqrt(.005446*.0003662) # .711 correlation for RC1
.00175/sqrt(.01881*.000167) # .987 correlation for RC2

# Matches Verbyla 1999 figure 5
plot(varioGram(m5),
    main="verbyla.lupin - final model variograms",
    xlim=c(1:19), ylim=c(0,1.5))

## End(Not run)

vold.longterm Long-term barley yields at different fertilizer levels

Description

Long-term barley yields at different fertilizer levels

Usage

data("vold.longterm")

Format

A data frame with 76 observations on the following 3 variables.

year
nitro nitrogen fertilizer, grams/m^2
yield yield, grams/m^2

Details

Trials conducted at Osaker, Norway. Nitrogen fertilizer amounts were increased by twenty percent in 1978.

Vold (1998) fit a Michaelis-Menten type equation with a different maximum in each year and a decreasing covariate for non-fertilizer nitrogen.

Miguez used a non-linear mixed effects model with asymptotic curve.

Source

vold.longterm

References


Examples

```r
## Not run:

library(agridat)
data(vold.longterm)
dat <- vold.longterm

libs(lattice)
foo1 <- xyplot(yield ~ nitro | factor(year), data = dat,
               as.table=TRUE, type = "o",
               main=list("vold.longterm", cex=1.5),
               xlab = list("N fertilizer",cex=1.5,font=4),
               ylab = list("Yield", cex=1.5))

# Long term trend shows decreasing yields
xyplot(yield ~ year , data = dat, group=nitro, type="o",
       main="vold.longterm - yield level by nitrogen",
       auto.key=list(columns=4))

if(0){
  # Global model
  m1.nls <- nls(yield ~ SSasymp(nitro, max, int, lograte), data=dat)
  summary(m1.nls)
  libs(MASS) # for 'confint'
  confint(m1.nls)

  # Raw data plus global model. Year variation not modeled.
  pdat <- data.frame(nitro=seq(0,14,0.5))
  pdat$pred <- predict(m1.nls, newdata=pdat)
  libs(latticeExtra) # for layers
  foo1 + xyplot(pred ~ nitro , data = pdat,
                as.table=TRUE, type='l', col='red', lwd=2)
}

# Separate fit for each year. Overfitting with 3x19=57 params.
libs(nlme)
m2.lis <- nlsList(yield ~ SSasymp(nitro,max,int,lograte) | year, data=dat)
plot(intervals(m2.lis),layout = c(3,1)) # lograte might be same for each year

# Fixed overall asymptotic model, plus random deviations for each year
# Simpler code, but less clear about what model is fit: m3.lme <- nlme(m2.lis)
libs(nlme)
m3.lme <- nlme(yield ~ SSasymp(nitro, max, int, lograte), data=dat,
               groups =~ year,
               fixed = list(max~1, int~1, lograte~1),
```
```
random= max + int + lograte ~ 1,
start= c(max=300, int=100, rate=-2))
## # Fixed effects are similar for the nls/lme models
## coef(m1.nls)
## fixef(m3.lme)
## # Random effects are normally distributed
## qqnorm(m3.lme, ~ ranef(.),col="black")
## # Note the trend in intercept effects over time
## plot(ranef(m3.lme),layout=c(3,1))
## # Correlation between int,lograte int,max may not be needed
## intervals(m3.lme,which="var-cov")
## pairs(m3.lme,pch=19,col="black")
## # Model with int uncorrelated with max,lograte. AIC is worse.
## # fit4.lm3 <- update(m3.lme, random=pdBlocked(list(max+lograte~1,int ~ 1)))
## # intervals(fit4.lm3, which="var-cov")
## # anova(m3.lme, fit4.lm3)

# Plot the random-effect model. Excellent fit with few parameters.
pdat2 <- expand.grid(year=1970:1988, nitro=seq(0,15,length=50))
pdat2$pred <- predict(m3.lme, new=pdat2)
pdat2$predf <- predict(m3.lme, new=pdat2, level=0)
foo1 <- update(foo1, type="p",
    key=simpleKey(c("Observed","Fixed","Random"),
    col=c("blue","red","darkgreen"),
    points=FALSE, columns=3))
libs(latticeExtra)
foo2 <- xyplot(pred~nitro|year, data=pdat2, type="l", col="darkgreen", lwd=2)
foo3 <- xyplot(predf~nitro|year, data=pdat2, type="l", col="red",lwd=1)
foo1 + foo2 + foo3

## # Income is maximized at about 15
## pdat2 <- transform(pdat2, income = predf*2 - 7*nitro)
## with(pdat2, xyplot(income~nitro))
## End(Not run)
```

---

**vsn.lupin3**

*Multi-environment trial of lupin, early generation trial*

---

**Description**

Early generation lupin trial with 3 sites, 330 test lines, 6 check lines.

**Format**

A data frame with 1236 observations on the following 5 variables.
Details

An early-stage multi-environment trial, with 6 check lines and 300 test lines. The 6 check lines were replicated in each environment.

Used with permission of Arthur Gilmour, Brian Cullis, Robin Thompson.

Source


Examples

## Not run:

```r
library(agridat)
data(vsn.lupin3)
dat <- vsn.lupin3

# Split gen into check/test, make factors
dat <- within(dat, {
  check <- ifelse(gen>336, 0, gen)
  check <- ifelse(check<7, check, 7)
  check <- factor(check)
  test <- factor(ifelse(gen>6 & gen<337, gen, 0))
  gen=factor(gen)
})

libs(desplot)
desplot(dat, yield~ col*row|site,
  # midpoint="midrange",
  # aspect unknown
  main="vsn.lupin3 - yield")

# Site 1 & 2 used same randomization
desplot(dat, check~ col*row|site,
  main="vsn.lupin3: check plot placement")

# asreml 4
libs(asreml,lucid)

# Single-site analyses suggested random row term for site 3,
# random column terms for all sites,
# AR1 was unnecessary for the col dimension of site 3
dat <- transform(dat, colf=factor(col), rowf=factor(row))
dat <- dat[order(dat$site, dat$colf, dat$rowf),] # Sort for asreml
```
m1 <- asreml(yield ~ site + check:site, data=dat, 
random = ~ at(site):colf + at(site,3):rowf + test, 
residual = ~ dsum( - ar1(colf):ar1(rowf) + 
   id(colf):ar1(rowf) | site, 
   levels=list(1:2, 3) ) 
)

m1$loglik
## [1] -314.2616

vc(m1)
## effect component std.error z.ratio constr
## at(site, S1):colf!colf.var 0.6228 0.4284 1.5 pos
## at(site, S2):colf!colf.var 0.159 0.1139 1.4 pos
## at(site, S3):colf!colf.var 0.04832 0.02618 1.8 pos
## at(site, S3):rowf!rowf.var 0.0235 0.008483 2.8 pos
## test!test.var 0.1031 0.01468 7 pos
## site_S1!variance 2.771 0.314 8.8 pos
## site_S1!colf.cor 0.1959 0.05375 3.6 uncon
## site_S1!rowf.cor 0.6503 0.03873 17 uncon
## site_S2!variance 0.9926 0.1079 9.2 pos
## site_S2!colf.cor 0.2868 0.05246 5.5 uncon
## site_S2!rowf.cor 0.5744 0.0421 14 uncon
## site_S3!variance 0.1205 0.01875 6.4 pos
## site_S3!rowf.cor 0.6394 0.06323 10 uncon

# Add site:test
m2 <- update(m1, random=~. + site:test)

m2$loglik
## [1] -310.8794

# CORUH structure on the site component of site:test
m3 <- asreml(yield ~ site + check:site, data=dat, 
random = ~ at(site):colf + at(site,3):rowf + corh(site):test, 
residual = ~ dsum( - ar1(colf):ar1(rowf) + 
   id(colf):ar1(rowf) | site, 
   levels=list(1:2, 3) )
)

m3$loglik
## [1] -288.4837

# Unstructured genetic variance matrix
m4 <- asreml(yield ~ site + check:site, data=dat, 
random = ~ at(site):colf + at(site,3):rowf + us(site):test, 
residual = ~ dsum( - ar1(colf):ar1(rowf) + 
   id(colf):ar1(rowf) | site, 
   levels=list(1:2, 3) )
)

m4$loglik
## [1] -286.8239

# Note that a 3x3 unstructured matrix can be written LL'+Psi with 1 factor L
# Explicitly fit the factor analytic model
m5 <- asreml(yield ~ site + check:site, data=dat, 
random = ~ at(site):colf + at(site,3):rowf + fa(site,1, init=c(.7,.1,.1,.5,.3,.2)):test,
\[
\text{residual = ~ dsum(} ~ \text{ar1(colf):ar1(rowf) + id(colf):ar1(rowf)} \mid \text{site, levels=list(1:2, 3))}
\]

\[
m5$loglik \# \text{Same as m4}
## [1] -286.8484

\# Model 4, Unstructured (symmetric) genetic variance matrix
un <- diag(3)
un[upper.tri(un,TRUE)] <- m4$vparameters[5:10]
round(un+t(un)-diag(diag(un)),3)
## [,1] [,2] [,3]
## [1,] 0.992 0.158 0.132
## [2,] 0.158 0.073 0.078
## [3,] 0.132 0.078 0.122

\# Model 5, FA matrix = LL' + Psi. Not quite the same as unstructured,
\# since the FA model fixes site 2 variance at 0.
psi <- diag(m5$vparameters[5:7])
lam <- matrix(m5$vparameters[8:10], ncol=1)
round(tcrossprod(lam,lam)+psi,3)
## [,1] [,2] [,3]
## [1,] 0.991 0.156 0.133
## [2,] 0.156 0.092 0.078
## [3,] 0.133 0.078 0.122

## End(Not run)

---

**wallace.iowaland**

*Iowa farmland values by county in 1925*

**Description**

Iowa farmland values by county in 1925

**Usage**

data("wallace.iowaland")

**Format**

A data frame with 99 observations on the following 10 variables.

- **county**: county factor, 99 levels
- **fips**: FIPS code (state+county)
- **lat**: latitude
- **long**: longitude
- **yield**: average corn yield per acre (bu)
corn  percent of land in corn
grain percent of land in small grains
untillable percent of land untillable
fedval land value (excluding buildings) per acre, 1925 federal census
stval land value (excluding buildings) per acre, 1925 state census

Details
None.

Source

References
Larry Winner. Spatial Data Analysis. https://www.stat.ufl.edu/~winner/data/iowaland.txt

Examples
```r
## Not run:
library(agridat)
data(wallace.iowaland)
dat <- wallace.iowaland

# Interesting trends involving latitude
libs(lattice)
splom(~dat[,c(1:2)], type=c('p','smooth'), lwd=2, main="wallace.iowaland")

# Means. Similar to Wallace table 1
apply(dat[, c('yield','corn','grain','untillable','fedval')], 2, mean)

# Correlations. Similar to Wallace table 2
round(cor(dat[, c('yield','corn','grain','untillable','fedval')]),2)

m1 <- lm(fedval ~ yield + corn + grain + untillable, dat)
summary(m1) # estimates similar to Wallace, top of p. 389

# Choropleth map
libs(maps)
data(county.fips)
dat <- transform(dat, polnm = paste0('iowa,',county)) # polnm example: iowa,adair

libs("latticeExtra") # for mapplot
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
mapplot(polnm=fedval , data=dat, colramp=redblue,
main="wallace.iowaland - Federal land values",
xislab="Land value, dollars per acre",
scales=list(draw=FALSE),
"wallace.iowaland")
```

```r
library(agridat)
data(wallace.iowaland)
dat <- wallace.iowaland

# Interesting trends involving latitude
libs(lattice)
splom(~dat[,c(1:2)], type=c('p','smooth'), lwd=2, main="wallace.iowaland")

# Means. Similar to Wallace table 1
apply(dat[, c('yield','corn','grain','untillable','fedval')], 2, mean)

# Correlations. Similar to Wallace table 2
round(cor(dat[, c('yield','corn','grain','untillable','fedval')]),2)

m1 <- lm(fedval ~ yield + corn + grain + untillable, dat)
summary(m1) # estimates similar to Wallace, top of p. 389

# Choropleth map
libs(maps)
data(county.fips)
dat <- transform(dat, polnm = paste0('iowa,',county)) # polnm example: iowa,adair

libs("latticeExtra") # for mapplot
redblue <- colorRampPalette(c("firebrick", "lightgray", "#375997"))
mapplot(polnm=fedval , data=dat, colramp=redblue,
main="wallace.iowaland - Federal land values",
xislab="Land value, dollars per acre",
scales=list(draw=FALSE),
"wallace.iowaland")
```
walsh.cottonprice

## End(Not run)

walsh.cottonprice  

### Description

Acres and price of cotton 1910-1943

### Format

A data frame with 34 observations on the following 9 variables.

- **year**: year, numeric 1910-1943
- **acres**: acres of cotton (1000s)
- **cotton**: price per pound (cents) in previous year
- **cottonseed**: price per ton (dollars) in previous year
- **combined**: cotton price/pound + 1.857 x cottonseed price/pound (cents)
- **index**: price index, 1911-1914=100
- **adjcotton**: adjusted cotton price per pound (cents) in previous year
- **adjcottonseed**: adjusted cottonseed price per ton (dollars) in previous year
- **adjcombined**: adjusted combined price/pound (cents)

### Details

The 'index' is a price index for all farm commodities.

### Source


### Examples

```r
library(agridat)
data(walsh.cottonprice)
dat <- walsh.cottonprice
dat <- transform(dat, acres=acres/1000) # convert to million acres
percentchg <- function(x){ # percent change from previous to current
```
ix <- 2:(nrow(dat))
c(NA, (x[ix]-x[ix-1])/x[ix-1])

# Compare percent change in acres with percent change in previous price
# using constant dollars
dat <- transform(dat, chga = percentchg(acres), chgp = percentchg(adjcombined))

with(dat, cor(chga, chgp, use='pair')) # .501 correlation
libs(lattice)
xyplot(chga~chgp, dat, type=c('p','r'),
       main="walsh.cottonprice",
       xlab="Percent change in previous price", ylab="Percent change in acres")

wassom.brome.uniformity

Uniformity trials of bromegrass

Description


Usage

data("wassom.brome.uniformity")

Format

A data frame with 1296 observations on the following 3 variables.

expt experiment
row row
col column
yield forage yield, pounds

Details

Experiments were conducted at Ames, Iowa. The response variable is forage yield in pounds of
green weight.

Optimum plot size was estimated to be about 3.5 x 7.5 feet.

Wassom and Kalton used two different methods to estimate optimum plot size. 1. Relative efficiency of different plot sizes. 2. Regression of the log variance of yield vs log plot size.

There are three Experiments:

Experiment E1 was broadcast seeded, harvested in 1950.
Experiment E2 was row planted, harvested in 1950.
Experiment E3 was broadcast seeded, harvested in 1951. This field contained a mixture of alfalfa and brome in about equal proportions.

Each plot was 3.5 ft x 4 ft, but the orientation of the plot is not clear.

Field width: 36 plots
Field length: 36 plots

Source


Examples

```r
## Not run:
library(agridat)
data(wassom.brome.uniformity)
dat <- wassom.brome.uniformity

libs(desplot)
desplot(dat, yield~col*row|expt,
     flip=TRUE, aspect=1, # approximate aspect
     main="wassom.brome.uniformity")

## End(Not run)
```

---

**waynick.soil**  
*Soil nitrogen and carbon in two fields*

**Description**

Soil nitrogen and carbon in two fields

**Format**

A data frame with 200 observations on the following 6 variables.

- **field** field name, 2 levels
- **sample** sample number
- **x** x ordinate
- **y** y ordinate
- **nitro** nitrogen content, percent
- **carbon** carbon content, percent
Details

Two fields were studied, one at University Farm in Davis, the other near Oakley. The Davis field is silty clay loam, the Oakley field is blow sand.

Source


Examples

```r
## Not run:

library(agridat)
data(waynick.soil)
dat <- waynick.soil

# Strong relationship between N,C
libs(lattice)
xyplot(nitro~carbon|field, data=dat, main="waynick.soil")

# Spatial plot
libs(sp, gstat)
d1 <- subset(dat, field=="Davis")
d2 <- subset(dat, field=="Oakley")
coordinates(d1) <- data.frame(x=d1$x, y=d1$y)
coordinates(d2) <- data.frame(x=d2$x, y=d2$y)
spplot(d1, zcol = "nitro", cuts=8, cex = 1.6,
       main = "waynick.soil - Davis field - nitrogen",
       col.regions = bpy.colors(8), key.space = "right")

# Variogram
v1 <- gstat::variogram(nitro~1, data=d1)
plot(v1, main="waynick.soil - Davis field - nitrogen") # Maybe hasn't reached sill

## End(Not run)
```

---

**wedderburn.barley**

*Multi-environment trial of barley, percent of leaves affected by leaf blotch*

Description

Percent of leaf area affected by leaf blotch on 10 varieties of barley at 9 sites.
Format

A data frame with 90 observations on the following 3 variables.

y  Percent of leaf area affected, 0-100.
site Site factor, 9 levels
gen Variety factor, 10 levels

Details

Incidence of *Rhynchosporium secalis* (leaf blotch) on the leaves of 10 varieties of barley grown at 9 sites in 1965.

Source


Wedderburn credits the original data to an unpublished thesis by J. F. Jenkyn.

References


Examples

```r
library(agridat)
data(wedderburn.barley)
dat <- wedderburn.barley
dat$y <- dat$y/100

lattice
dotplot(gen~y|site, dat, main="wedderburn.barley")

# Use the variance function mu(1-mu). McCullagh page 330
# Note, 'binomial' gives same results as 'quasibinomial', but also a warning
m1 <- glm(y ~ gen + site, data=dat, family="quasibinomial")
summary(m1)

# Same shape (different scale) as McCullagh fig 9.1a
plot(m1, which=1, main="wedderburn.barley")

# Compare data and model
dat$pbin <- predict(m1, type="response")
dotplot(gen~pbin+y|site, dat, main="wedderburn.barley: observed/predicted")

# Wedderburn suggested variance function: mu^2 * (1-mu)^2
```
# Millar shows how to do this explicitly.

```r
wedder <- list(varfun=function(mu) (mu*(1-mu))^2,
  validmu=function(mu) all(mu>=0) && all(mu<1),
  dev.resids=function(y, mu, wt) wt * ((y-mu)^2)/(mu*(1-mu))^2,
  initialize=expression({
    n <- rep.int(1, nobs)
    mustart <- pmax(0.001, pmin(0.99, y))
  }},
  name="(mu(1-mu))^2")

m2 <- glm(y ~ gen + site, data=dat, family=quasi(link="logit", variance=wedder))
#plot(m2)

## Not run:
# Alternatively, the 'gnm' package has the 'wedderburn' family.

libs(gnm)

m3 <- glm(y ~ gen + site, data=dat, family="wedderburn")
summary(m3)

## Similar to McCullagh fig 9.2
plot(m3, which=1)

## End(Not run)
```

---

**weiss.incblock**

*Soybean balanced incomplete block experiment*

**Description**

Soybean balanced incomplete block experiment

**Usage**

```r
data("weiss.incblock")
```

**Format**

A data frame with 186 observations on the following 5 variables.

- **block**: block factor
- **gen**: genotype (variety) factor
- **yield**: yield (bu/ac)
- **row**: row
- **col**: column
Details

Grown at Ames, Iowa in 1937. Each plot was 6 feet by 16 feet (2 rows, 3 feet apart). Including space between plots, the entire experiment was 252 ft x 96 feet (7 block * 6 plots * 6 feet = 252, 16*5 plots plus 4 gaps of 4 feet). Weiss shows a figure of the field (that was later doubled in size via using two rows per plot).

Note that only 30 varieties were tested. Varieties 7 and 14 are the same variety (Mukden). Although total yields of these varieties were not equal, the correction for blocks adjusted their means to identical values. Such accuracy is not, however, claimed to be a constant characteristic of the design.

Field width: 96 feet
Field length: 252 feet

Source


Examples

```r
## Not run:

library(agridat)
data(weiss.incblock)
dat <- weiss.incblock

# True aspect as shown in Weiss and Cox
libs(desplot)
desplot(dat, yield~col*row, 
  text=gen, shorten='none', cex=.6, out1=block, 
  aspect=252/96, # true aspect
  main="weiss.incblock")

# asreml
# Standard inc block analysis used by Weiss and Cox
libs(asreml) # asreml 4
m1 <- asreml(yield ~ gen + block , data=dat)
predict(m1, data=dat, classify="gen")$pvals

## gen pred.value std.error est.stat
## G01 24.59 0.8312 Estimable
## G02 26.92 0.8312 Estimable
## G03 32.62 0.8312 Estimable
## G04 26.97 0.8312 Estimable
## G05 26.02 0.8312 Estimable

## End(Not run)
```
Description

Lattice experiment in soybeans.

Usage

data("weiss.lattice")

Format

A data frame with 196 observations on the following 5 variables.

- **yield**: yield (bu/ac)
- **gen**: genotype factor, 49 levels
- **rep**: rep factor, 4 levels
- **col**: column
- **row**: row

Details

Yield test of 49 soybean varieties, grown at Ames, IA, in 1938. Plot dimensions were 3x16 feet. The varieties are compared to variety 26 (Mukden).

It is not clear how the reps were positioned in the field. On the one hand, the middle three columns of each rep/square are higher yielding, giving the appearance of the reps being stacked on top of each other. On the other hand, the analysis by Weiss uses 24 degrees of freedom 4*(7-1) to fit a separate effect for each column in each rep (instead of across reps).

Source


Examples

```r
## Not run:
library(agridat)
data(weiss.lattice)
dat <- weiss.lattice

libs(desplot)
desplot(dat, yield~col*row|rep,
       text=gen, shorten="none", cex=.8, aspect=3/16, # true aspect)
```
main="weiss.lattice (layout uncertain)", xlab="Soybean yields")

dat <- transform(dat, xf=factor(col), yf=factor(row))
m1 <- lm(terms(yield ~ rep + rep:xf + rep:yf + gen, keep.order=TRUE), data=dat)
anova(m1) # Matches Weiss table 7
## Response: yield
## Df Sum Sq Mean Sq F value Pr(>F)
## rep 3 91.57 30.525 4.7414 0.0039709 **
## rep:xf 24 2913.43 121.393 18.8557 < 2.2e-16 ***
## rep:yf 24 390.21 16.259 2.5254 0.0007734 ***
## gen 48 1029.87 21.456 3.3327 2.652e-07 ***
## Residuals 96 618.05 6.438

# ----------

libs(asreml) # asreml4
m2 <- asreml(yield ~ rep + rep:xf + rep:yf + gen, data=dat)
# Weiss table 6 means
wald(m2)
predict(m2, data=dat, classify="gen")$pvals

## gen pred.value std.error est.stat
## G01 27.74 1.461 Estimable
## G02 24.95 1.461 Estimable
## G03 24.38 1.461 Estimable
## G04 28.05 1.461 Estimable
## G05 19.6 1.461 Estimable
## G06 23.79 1.461 Estimable

## End(Not run)

welch.bermudagrass

Factorial experiment of bermuda grass, N, P, K fertilizers

Description

Factorial experiment of bermuda grass, N, P, K fertilizers

Format

A data frame with 64 observations on the following 4 variables.

n  nitrogen fertilizer, pounds/acre
p  phosphorus, lb/ac
k  potassium, lb/ac
yield  yield of grass, tons/ac
Details

The experiment was conducted 1955, 1956, and 1957. There were 3 blocks. The harvests were oven-dried. Each value is the mean for 3 years and 3 replications. In most cases, the yield increased with additions of the fertilizer nutrients.

Source


References


Examples

```r
## Not run:
library(agridat)
data(welch.bermudagrass)
dat <- welch.bermudagrass
# Welch uses 100-pound units of n,p,k.
dat <- transform(dat, n=n/100, p=p/100, k=k/100)

libs(latticeExtra)
useOuterStrips(xyplot(yield~n|factor(p)*factor(k), data=dat, type='b',
main="welch.bermudagrass: yield for each P*K",
xlab="Nitro for each Phosphorous level",
ylab="Yield for each Potassium level"))

# Fit a quadratic model
m1 <- lm(yield ~ n + p + k + I(n^2) + I(p^2) + I(k^2) + n:p + n:k + p:k + n:p:k, data=dat)
signif(coef(m1),4) # These match the 3-yr coefficients of Welch, Table 2
## (Intercept) n p k I(n^2) I(p^2) I(k^2) n:p n:k p:k n:p:k
## 1.94300 2.00700 1.47100 0.61880 -0.33150 -1.29500 -0.37430 0.20780 0.18740 0.23480 0.02789

# Welch Fig 4. Modeled response curves
d1 <- expand.grid(n=seq(0, 4, length=50), p=0, k=0)
d1$pred <- predict(m1, d1)
d2 <- expand.grid(n=0, p=0, k=seq(0, 1.68, length=50))
d2$pred <- predict(m1, d2)
d3 <- expand.grid(n=0, p=seq(0, .88, length=50), k=0)
d3$pred <- predict(m1, d3)
op <- par(mfrow=c(1,3), mar=c(5,3,4,1))
```
plot(pred~n, data=d1, type='l', ylim=c(0,6), xlab="N 100 lb/ac", ylab="")
plot(pred~k, data=d2, type='l', ylim=c(0,6), xlab="K 100 lb/ac", ylab="")
title("welch.bermudagrass - Predicted yield vs fertilizer", outer=TRUE, line=-3)
plot(pred~p, data=d3, type='l', ylim=c(0,6), xlab="P 100 lb/ac", ylab="")
par(op)

# Brute-force grid-search optimization of fertilizer quantities, using
# $25/ton for grass, $.12/lb for N, $.18/lb for P, $.07/lb for K
# Similar to Example 5 in Table 4 of Welch

d4 <- expand.grid(n=seq(3,4,length=20), p=seq(.5, 1.5, length=20), k=seq(.8, 1.8, length=20))
d4$pred <- predict(m1, newdata=d4)
d4 <- transform(d4, income = 25*pred - .12*n*100 + -.18*p*100 -.07*k*100)
d4[which.max(d4$income),] # Optimum at 300 lb N, 71 lb P, 148 lb K

# ----- JAGS -----
if(0){
  # Congdon (2007) p. 124, provides a Bayesian model based on a GLM
  # by McCullagh & Nelder. We use JAGS and simplify the code.
  # y ~ gamma with shape = nu, scale = nu * eps_i
  # 1/eps = b0 + b1/(N+a1) + b2/(P+a2) + b3/(K+a3)
  # N,P,K are added fertilizer amounts, a1,a2,a3 are background
  # nutrient levels and b1,b2,b3 are growth parameters.

  libs(rjags)
  mod.bug =
  "model {
    for(i in 1:nobs) {
      yield[i] ~ dgamma(nu, mu[i])
      mu[i] <- nu * eta[i]
      eta[i] <- b0 + b1 / (N[i]+a1) + b2 / (P[i]+a2) + b3 / (K[i]+a3)
      yhat[i] <- 1 / eta[i]
    }
  }

  # Hyperparameters
  nu ~ dgamma(0.01, 0.01) # Informative priors
  a1 ~ dnorm(40, 0.01)
  a2 ~ dnorm(22, 0.01)
  a3 ~ dnorm(32, 0.01)
  b0 ~ dnorm(0, 0.0001)
  b1 ~ dnorm(0, 0.0001) I(0,) # Keep b1 non-negative
  b2 ~ dnorm(0, 0.0001) I(0,)
  b3 ~ dnorm(0, 0.0001) I(0,)
  }
"
  jdat <- with(welch.bermudagrass,
    list(yield=yield, N=n, P=p, K=k, nobs=64))
  jinit = list(a1=40, a2=22, a3=32, b0=.1, b1=10, b2=1, b3=1)
  oo <- textConnection(mod.bug)
wheatley.carrot

Insecticide treatments for carrot fly larvae

Description

Insecticide treatments for carrot fly larvae. Two insecticides with five depths.

Usage

data("wheatley.carrot")

Format

A data frame with 36 observations on the following 6 variables.

treatment treatment factor, 11 levels
insecticide insecticide factor
depth depth
rep block
damaged number of damaged plants
total total number of plants
Details

In 1964 an experiment was conducted with microplots to evaluate the effectiveness of treatments against carrot fly larvae. The treatment factor is a combination of insecticide and depth.

Hardin & Hilbe used this data to fit a generalized binomial model.

Famoye (1995) used the same data to fit a generalized binomial regression model. Results for Famoye are not shown.

Source

G A Wheatley & H Freeman. (1982). A method of using the proportions of undamaged carrots or parsnips to estimate the relative population densities of carrot fly (Psila rosae) larvae, and its practical applications. Annals of Applied Biology, 100, 229-244. Table 2.

https://doi.org/10.1111/j.1744-7348.1982.tb01935.x

References


Examples

library(agridat)
data(wheatley.carrot)
dat <- wheatley.carrot

# Observed proportions of damage
data <- transform(dat, prop=damaged/total)
libs(lattice)
xyplot(prop~depth|insecticide, data=dat, subset=treatment!="T11",
cex=1.5, main="wheatley.carrot", ylab="proportion damaged")

# Model for Wheatley. Deviance for treatment matches Wheatley, but other
# deviances do not. Why?
# treatment:rep is the residual
m1 <- glm(cbind(damaged,total-damaged) ~ rep + treatment + treatment:rep,
data=dat, family=binomial("cloglog"))
anova(m1)

# GLM of Hardin & Hilbe p. 161. By default, R uses T01 as the base,
# but Hardin uses T11. Results match.
m2 <- glm(cbind(damaged,total-damaged) ~ rep + C(treatment, base=11),
data=dat, family=binomial("cloglog"))
summary(m2)
Uniformity trial of wheat at Aberdeen, Idaho, 1927.

A data frame with 1500 observations on the following 3 variables.

row row
col column (series)
yield yield in grams per plot

Yield trial conducted in 1927 near Aberdeen, Idaho. The crop was Federation wheat (C.I. no 4734). Plots were seeded on April 18 with a drill that sowed eight rows at a time. Individual rows were harvested in August and threshed with a small nursery thresher. Some authors recommend analyzing the square root of the yields.

Rows were 15 feet long, 1 foot apart.

Field width: 12 columns * 15 feet = 180 feet wide.

Field length: 125 rows * 12 in = 125 feet


Examples

```r
## Not run:

library(agridat)
data(wiebe.wheat.uniformity)
dat <- wiebe.wheat.uniformity

libs(desplot)
desplot(dat, yield~col+row,
   aspect=125/180, flip=TRUE, # true aspect
   main="wiebe.wheat.uniformity: yield") # row 1 is at south

# Preece (1981) found the last digits have an interesting distribution
# with 0 and 5 much more common than other digits.
dig <- substring(dat$yield, nchar(dat$yield))
dig <- as.numeric(dig)
hist(dig, breaks=0:10-.5, xlab="Last digit",
   main="wiebe.wheat.uniformity - histogram of last digit")
table(dat$col, dig) # Table 3 of Preece

# Wilkinson (1983, p. 152) noted that an 8-row planter was used which
# produced a recurring pattern of row effects on yield. This can be seen
# in the high autocorrelations of row means at lag 8 and lag 16
rowm <- tapply(dat$yield, dat$row, mean)
acf(rowm, main="wiebe.wheat.uniformity row means")

# Plot the row mean against the planter row unit 1-8
libs("lattice")
xyplot(rowm~rep(1:8, length=125),
   main="wiebe.wheat.uniformity",
   xlab="Planter row unit", ylab="Row mean yield")

# Wiebe (1937) and Yates (1939) show the effect of "guess rows"
# caused by the 8-row drill passing back and forth through
# the field.
# Yates gives the distance between strips (8 rows per strip) as:
# 10.2,12.4,11.7,13.4,10.6,14.2,11.8,13.8,12.2,13.1,11.2,14,11.3,12.9,12.4

# First give each row 12 inches of growing width between rows
tmp <- data.frame(row=1:125,area=12)
# Distance between rows 8,9 is 10.2 inches, so we give these two
# rows 6 inches (on the 'inside' of the strip) and 10.2/2=5.1 inches
# on the outside of the strip, total 11.1 inches

tmp$Area[8:9] <- 6 + 10.2/2
tmp$Area[16:17] <- 6 + 12.4/2
tmp$Area[24:25] <- 6 + 11.7/2
tmp$Area[32:33] <- 6 + 13.4/2
tmp$Area[40:41] <- 6 + 10.6/2
tmp$Area[48:49] <- 6 + 14.2/2
tmp$Area[56:57] <- 6 + 11.8/2
tmp$Area[64:65] <- 6 + 13.8/2
tmp$Area[72:73] <- 6 + 12.2/2
```
tmp$area[80:81] <- 6 + 13.1/2
tmp$area[88:89] <- 6 + 11.2/2
tmp$area[96:97] <- 6 + 14.0/2
tmp$area[104:105] <- 6 + 11.3/2
tmp$area[112:113] <- 6 + 12.9/2
tmp$area[120:121] <- 6 + 12.4/2
dat <- merge(dat, tmp)

# It's not clear if Wiebe used border rows...we delete them
dat <- subset(dat, row > 1 & row < 125)

# Wiebe (1937) calculated a moving average to adjust for fertility
# effects, then used only the OUTER rows of each 8-row drill strip
# and found 21.5 g / inch of space between rows. We used all the
# data without correcting for fertility and obtained 33.1 g / inch.
xypplot(yield ~ area, dat, type=c('p','r'),
       xlab="Average area per row", ylab="Yield")
coef(lm(yield ~ area, dat))[2]
# 33.1

## End(Not run)

---

**wiedemann.safflower.uniformity**

*Uniformity trial of safflower*

### Description

Uniformity trial of safflower at Farmington, Utah, 1960.

### Usage

```
data("wiedemann.safflower.uniformity")
```

### Format

A data frame with 1782 observations on the following 3 variables.

- **row** row
- **col** column
- **yield** yield, grams

### Details

This trial was planted at University Field Station, Farmington, Utah, in 1960, on a plot of land about one half acre in size. The soil was not too uniform...the northern third of the field was clay and the rest was gravelly. Rows were planted 22 inches apart, 62 rows total, each row running the length of the field. Before harvest, 4 rows were removed from each side, and 12 feet was removed from each
Each row was harvested in five-foot lengths, threshed, and the seed weighed to the nearest gram.
The northern third of the field had yields twice as high as the remaining part of the field because the soil had better moisture retention. The remaining part of the field had yields that were more uniform.

Wiedemann determined the optimum plot size to be about 8 basic plots. The shape of the plot was not very important. But, two-row plots were recommended for simplicity of harvest, so 3.33 feet by 20 feet.

Based on operational costs, K1=74 percent and K2=26 percent.

Field width: 33 plots/ranges * 5ft = 165 feet
Field length: 54 rows * 22 in/row = 99 feet

For this R package, the tables in Wiedemann were converted by OCR to digital format, and all values were checked by hand.

The original source document has columns labeled 33, 32, ... 1. Here the columns are labeled 1:33 so that plotting tools work normally. See Wiedemann figure 8.

Wiedemann notes the statistical analysis of the data required 100 hours of labor. Today the analysis takes only a second.

Source


References

None.

Examples

```r
## Not run:

library(agridat)

data(wiedemann.safflower.uniformity)
dat <- wiedemann.safflower.uniformity

# CV of entire field = 39
sd(dat$yield)/mean(dat$yield)

libs(desplot)
desplot(dat, yield~col*row,
       flip=TRUE, tick=TRUE, aspect =99/165, # true aspect
       main="wiedemann.safflower.uniformity (true shape)")

libs(agricolae)
libs(reshape2)
dmat <- acast(dat, row~col, value.var='yield')
```
Williams barley uniformity

Uniformity trial of barley

Description

Uniformity trial of barley at Narrabri, New South Wales, 1984.

Format

A data frame with 720 observations on the following 3 variables.

row row
col column
yield grain yield kg/ha divided by 10

Details

Grown at Roseworthy Agricultural College. Plots were 5 m long (4 m sown, 3.3 m harvested) by 0.75 m wide.

A three-plot seeder was used, planting in a serpentine fashion. Williams noted that it appears that the middle plot of each pass has a lower yield, possibly due to soil compaction from the tractor.

Field width: 48 plots * .75 m = 36 m
Field length: 15 plots * 5 m = 75 m

Source


References

Examples

```r
## Not run:

library(agridat)
data(williams.barley.uniformity)
dat <- williams.barley.uniformity

libs(desplot)
desplot(dat, yield ~ col*row,
   aspect= 75/36, # true aspect
   main="williams.barley.uniformity")

# Smoothed contour/persp plot like Williams Fig 1b, 2b
libs(lattice)
dat$fit <- fitted(loess(yield~col*row, dat, span=.1))
contourplot(fit~col*row, data=dat,
   aspect=75/36, region=TRUE, col.regions=RedGrayBlue,
   main="williams.barley.uniformity")
wireframe(fit~col*row, data=dat, zlim=c(100, 350),
   main="williams.barley.uniformity")

# Williams table 1
anova(aov(yield ~ factor(row) + factor(col), dat))

## End(Not run)
```

williams.cotton.uniformity

*Uniformity trial of cotton*

Description


Format

A data frame with 288 observations on the following 3 variables.

<table>
<thead>
<tr>
<th>row</th>
<th>row</th>
</tr>
</thead>
<tbody>
<tr>
<td>col</td>
<td>column</td>
</tr>
<tr>
<td>yield</td>
<td>lint yield, kg/ha divided by 10</td>
</tr>
</tbody>
</table>

Details

Cotton uniformity trial grown at Narrabri, New South Wales, 1984-1985. Plots were 12m long, 1m apart, 12 rows by 24 columns, with an irrigation furrow between columns.

Field width: 24 plots * 1 m = 24 m
Field length: 12 plots * 12 m = 144 m
williams.trees

Multi-environment trial of trees, height / survival of 37 species at 6 sites in Thailand

Description

Multi-environment trial of trees, height / survival of 37 species at 6 sites in Thailand

Format

A data frame with 222 observations on the following 4 variables.

- env  Environment factor, 6 levels
- gen  Genetic factor, 37 levels
- height  Height (cm)
- survival  Survival percentage
Details

Planted in 1985 at six sites in Thailand. RCB with 3 reps. The data here is the mean of the three reps. Plots were 5 meters square with spacing 2m x 2m. Measurements collected at 24 months. The gen column in the data is actually seedlot, as some tree species have multiple seed lots. The trees are mostly acacia and eucalyptus.

Used with permission of Emlyn Williams.

Source


References


Examples

library(agridat)

data(williams.trees)
dat <- williams.trees

libs(lattice)

xyplot(survival~height|env,dat, main="williams.trees", xlab="Height", ylab="Percent surviving")

---

woodman.pig

Weight gain in pigs for different treatments

Description

Weight gain in pigs for different treatments, with initial weight and feed eaten as covariates.

Usage

data("woodman.pig")

Format

A data frame with 30 observations on the following 7 variables.

pen pen
treatment diet
pig  pig number
sex  sex
weight1  initial weight in pounds, week 0
weight2  final weight in pounds, week 16
feed  feed eaten in pounds
w0  initial weight
g  average weekly gain
h  half rate of change in growth

Details
Six pigs in each of 5 pens were fed individually. From each litter there were 3 males and 3 females chosen for a pen. Three different diet treatments were used.

Note: Woodman gives the initial weights to the nearest 0.5 pounds.
The w0, g, h columns are from Wishart 1938. Wishart used the weekly weight measurements (not available) to fit quadratic growth curves for each pig and then reported the constants. These are the data that are widely used by many authors.

Source

References
Wishart (1950) Table 2, p 17.
Peter H Westfall, Randall Tobias, Russell D Wolfinger (2011). Multiple Comparisons and Multiple Tests using SAS. Sec 8.3.

Examples
## Not run:
library(agridat)
data(woodman.pig)
dat <- woodman.pig

# add day of year for each weighing
dat <- transform(dat, date1=36, date2=148)
plot(NA, xlim=c(31,153), ylim=c(28,214),
  xlab="day of year", ylab="weight")
segments(dat$dat1, dat$weight1, dat$date2, dat$weight2,
wyatt.multi.uniformity

Uniformity trial of oats and wheat on the same ground.

Description
Uniformity trial of oats and wheat on the same ground.

Usage
data("wyatt.multi.uniformity")

Format
A data frame with 258 observations on the following 5 variables.
col column
crop row
yield yield, bu/ac
year year
crop

Details
Experiments conducted at the Soils Experimental field at the University of Alberta, Canada. Oats were grown in 1925, The average yield was 88 bushels. Wheat was grown in 1926, with an average yield of 32.2 bu/acre. The data reported are relative yields within each year.

The plot size in rows 1 and 2 (Series A and B in the original paper) is 1/10th acre. The plot size in row 3 is 1/11 acre.
Field length: 3 plots (140 ft, 140 ft, 128 ft) + 2 roads * 16 feet = 440 feet.

Field width: 43 plots * 37 ft = 1591 feet.

Source


References

None

Examples

```
library(agridat)
data(wyatt.multi.uniformity)
dat <- wyatt.multi.uniformity

# range of yields. Wyatt has 48.6 bu/ac for oats, 10.4 for wheat
# diff(range(na.omit(subset(dat, crop=="oats")$yield)/100*88)) # 48.4
# diff(range(na.omit(subset(dat, crop=="wheat")$yield)/100*32.8)) # 10.5

# std dev. Wyatt has 9.18 bu/ac for oats, 2.06 for wheat, 2.06 for wheat
# sd(na.omit(subset(dat, crop=="oats")$yield)/100*88) # 9.11
# sd(na.omit(subset(dat, crop=="wheat")$yield)/100*32.8) # 2.14

# correlation across years. Wyatt has .08
# cor(reshape2::acast(dat, row+col ~ crop, value.var="yield"), use="pair")

# Fig 3
libs(lattice)
xyplot(yield ~ col|factor(row), dat, group=crop,
       main="wyatt.multi.uniformity",
       type="l", layout=c(1,3), auto.key=TRUE )

## Not run:
libs(desplot)
desplot(dat, yield ~ col*row, subset=crop=="oats",
        tick=TRUE,
        aspect=(440)/(1591), # true aspect
        main="wyatt.multi.uniformity - 1925 oats")
desplot(dat, yield ~ col*row, subset=crop=="wheat",
        aspect=(440)/(1591), # true aspect
        main="wyatt.multi.uniformity - 1926 wheat")

## End(Not run)
```
yan.winterwheat  

Multi-environment trial of winter wheat in Ontario

Description
Yield of 18 varieties of winter wheat grown at 9 environments in Ontario in 1993.

Format
A data frame with 162 observations on the following 3 variables.

- gen  genotype
- env  environment
- yield  yield in metric tons per hectare

Used with permission of Weikai Yan.

Details
The yield is the mean of several reps, measured in metric tons per hectare.
This data has often been used to illustrate GGE biplots.

Source
Weikai Yan and Nicholas A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Table 1.

References

Examples
```r
## Not run:
library(agridat)

data(yan.winterwheat)
dat <- yan.winterwheat

libs(gge)
ml <- gge(dat, yield ~ gen*env)
biplot(ml, flip=c(1,1), hull=TRUE,
    main="yan.winterwheat - GGE biplot")
```
## Description

Yield of 6 barley varieties at 18 locations in Alberta.

## Usage

data("yang.barley")

## Format

A data frame with 108 observations on the following 3 variables.

- `site` site factor, 18 levels
- `gen` genotype factor, 6 levels
- `yield` yield, Mg/ha

## Details

From an experiment in 2003. Yang (2013) uses this data to illustrate a procedure for bootstrapping biplots.

<table>
<thead>
<tr>
<th>site</th>
<th>long</th>
<th>lat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beaverlodge</td>
<td>119.43</td>
<td>55.21</td>
</tr>
<tr>
<td>BigLakes</td>
<td>113.70</td>
<td>53.61</td>
</tr>
<tr>
<td>Calmar</td>
<td>113.85</td>
<td>53.26</td>
</tr>
<tr>
<td>CdcNorth</td>
<td>113.33</td>
<td>53.63</td>
</tr>
<tr>
<td>DawsonCreek</td>
<td>120.23</td>
<td>55.76</td>
</tr>
<tr>
<td>FtKent</td>
<td>110.61</td>
<td>54.31</td>
</tr>
<tr>
<td>FtStJohn</td>
<td>120.85</td>
<td>56.25</td>
</tr>
<tr>
<td>Irricana</td>
<td>113.60</td>
<td>51.32</td>
</tr>
<tr>
<td>Killam</td>
<td>111.85</td>
<td>52.78</td>
</tr>
<tr>
<td>Lacombe</td>
<td>113.73</td>
<td>52.46</td>
</tr>
<tr>
<td>LethbridgeDry</td>
<td>112.81</td>
<td>49.70</td>
</tr>
<tr>
<td>LethbridgeIrr</td>
<td>112.81</td>
<td>49.70</td>
</tr>
<tr>
<td>Lomond</td>
<td>112.65</td>
<td>50.35</td>
</tr>
<tr>
<td>Neapolis</td>
<td>113.86</td>
<td>51.65</td>
</tr>
<tr>
<td>NorthernSunrise</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Olds</td>
<td>114.09</td>
<td>51.78</td>
</tr>
<tr>
<td>StPaul</td>
<td>111.28</td>
<td>53.98</td>
</tr>
<tr>
<td>Stettler</td>
<td>112.71</td>
<td>52.31</td>
</tr>
</tbody>
</table>
yates.missing

Factorial experiment with missing values

Description

Potato factorial experiment with missing values
**Format**

A data frame with 80 observations on the following 3 variables.

- **trt**: treatment factor with levels 0 K N P NK KP NP NKP
- **block**: block, 10 levels
- **y**: infection intensity

**Details**

The response variable $y$ is the intensity of infection of potato tubers inoculated with * Phytophthora Erythroseptica.*

Yates (1933) presents an iterative algorithm to estimate missing values in a matrix, using this data as an example.

**Source**

F. Yates, 1933. The analysis of replicated experiments when the field results are incomplete. *Emp. J. Exp. Agric.*, 1, 129–142.

**References**


**Examples**

```r
## Not run:
library(agridat)
data(yates.missing)
dat <- yates.missing

libs(lattice)
bwplot(y ~ trt, data=dat,
  xlab="Treatment", ylab="Infection intensity",
  main="yates.missing")

libs(reshape2)
mat0 <- acast(dat[, c('trt','block','y')], trt~block,
  id.var=c('trt','block'), value.var='y')

# Use lm to estimate missing values. The estimated missing values
# are the same as in Yates (1933)
m1 <- lm(y~trt+block, dat)
dat$pred <- predict(m1, new=dat[, c('trt','block')])
dat$filled <- ifelse(is.na(dat$y), dat$pred, dat$y)
mat1 <- acast(dat[, c('trt','block','pred')], trt~block,
  id.var=c('trt','block'), value.var='pred')

# Another method to estimate missing values via PCA
```
\begin{verbatim}
libs("nipals")
m2 <- nipals(mat0, center=FALSE, ncomp=3, fitted=TRUE)
# mat2 <- m2$scores
mat2 <- m2$fitted

# Compare
ord <- c("0","N","K","P","NK","NP","KP","NKP")
print(mat0[ord,,], na.print=".")
round(mat1[ord,,] ,2)
round(mat2[ord,,] ,2)

# SVD with 3 components recovers original data better
sum((mat0-mat1)^2, na.rm=TRUE)
sum((mat0-mat2)^2, na.rm=TRUE) # Smaller SS => better fit

## End(Not run)
\end{verbatim}

\textbf{yates.oats}  \hspace{1cm} \textit{Split-plot experiment of oats}

\section*{Description}

The yield of oats from a split-plot field trial conducted at Rothamsted in 1931. Varieties were applied to the main plots. Manurial (nitrogen) treatments were applied to the sub-plots. Each plot is \(\frac{1}{80}\) acre = 28.4 links * 44 links. Field width: 4 plots * 44 links = 176 links. Field length: 18 rows * 28.4 links = 511 links

The 'block' numbers in this data are as given in the Rothamsted Report. The 'grain' and 'straw' values are the actual pounds per sub-plot as shown in the Rothamsted Report. Each sub-plot is \(\frac{1}{80}\) acre, and a 'hundredweight (cwt)' is 112 pounds, so converting from sub-plot weight to hundredweight/acre needs a conversion factor of 80/112.

The 'yield' values are the values as they appeared in the paper by Yates, who used 1/4-pounds as the units (i.e. he multiplied the original weight by 4) for simpler calculations.

\section*{Format}

\begin{verbatim}
row row
col column
yield yield in 1/4 pounds per sub-plot, each 1/80 acre
nitro nitrogen treatment in hundredweight per acre
gen genotype, 3 levels
block block, 6 levels
grain grain weight in pounds per sub-plot
straw straw weight in pounds per sub-plot
\end{verbatim}
Source

References

Examples
```r
## Not run:
library(agridat)
data(yates.oats)
dat <- yates.oats

## # Means match Rothamsted report p. 144
## libs(dplyr)
## dat
## ## summarize(grain=mean(grain)*80/112,
## ## straw=mean(straw)*80/112)

libs(desplot)
# Experiment design & yield heatmap
desplot(dat, block ~ col*row, col.regions=c("black","yellow"),
        out1=block, num=nitro, col=gen,
        cex=1, aspect=511/176, # true aspect
        main="yates.oats")

# Roughly linear gradient across the field. The right-half of each
# block has lower yield. The blocking is inadequate!
libs("lattice")
xyplot(yield ~ col|factor(nitro), dat,
       type = c('p','r'), xlab='col', as.table = TRUE,
       main="yates.oats")

libs(lme4)
# Typical split-plot analysis. Non-significant gen differences
m3 <- lmer(yield ~ factor(nitro) * gen + (1|block/gen), data=dat)
# Residuals still show structure
xyplot(resid(m3) ~ dat$col, xlab='col', type=c('p','smooth'),
       main="yates.oats")

# Add a linear trend for column
m4 <- lmer(yield ~ col + factor(nitro) * gen + (1|block/gen), data=dat)
# xyplot(resid(m4) ~ dat$col, type=c('p','smooth'), xlab='col')

## Compare fits
AIC(m3,m4)
```
## df AIC
## m3  9 581.2372
## m4 10 557.9424 # Substantially better

# --------

# Marginal predictions from emmeans package and asreml::predict

# --- nlme ---
libs(nlme)
libs(emmeans)
# create unbalance
dat2 <- yates.oats[-c(1,2,3,5,8,13,21,34,55),]
m5l <- lme(yield ~ factor(nitro) + gen, random = ~1 | block/gen,
          data = dat2)

# asreml r 4 has a bug with asreml(factor(nitro))
dat2$nitrof <- factor(dat2$nitro)

# --- asreml4 ---
libs(asreml)
m5a <- asreml(yield ~ nitrof + gen, random = ~ block + block:gen, data=dat2)
libs(lucid)
vc(m5l)
vc(m5a)

emmeans::emmeans(m5l, "gen")
predict(m5a, data=dat2, classify="gen")$pvals

# --------

if(0){
  # Demonstrate use of regress package, compare to lme
  libs(regress)
m6 <- regress(yield ~ nitrof + gen, ~block + I(block:gen), identity=TRUE,
           verbose=1, data=dat)
  summary(m6)
  ## Variance Coefficients:
  ##          Estimate Std. Error
  ## block  214.468    168.794
  ## I(block:gen) 109.700     67.741
  ## In    162.558     32.189
  # ordinal causes clash with VarCorr
  if(is.element("package:ordinal", search())) detach(package:ordinal)
m7 <- lme(yield ~ nitrof + gen, random = ~ 1|block/gen, data=dat)
lme4::VarCorr(m7)
  ##       Variance StdDev

### zuidhof.broiler

Daily weight, feed, egg measurements for a broiler chicken

**Description**

Daily weight, feed, egg measurements for a broiler chicken

**Format**

A data frame with 59 observations on the following 6 variables.

- **bw** Body weight, grams
- **targetbw** Target body weight, grams
- **adfi** Average daily feed intake, grams
- **adg** Average daily gain, grams
- **eggwt** Egg weight, grams
- **age** Age, days

**Details**

Using graphs like the one in the examples section, the authors discovered that a drop in body weight commonly occurs around the time of first egg production.

Used with permission of Martin Zuidhof.

**Source**

Examples

```r
library(agridat)

data(zuidhof.broiler)
dat <- zuidhof.broiler

dat <- transform(dat, age=age/7) # Change days into weeks

# Reproducing figure 1 of Zuidhof et al.

# Plot using left axis
op <- par(mar=c(5,4,4,4))
plot(bw~age, dat, xlab="Age (weeks)", ylab="Bodyweight (g)",
     main="zuidhof.broiler",
     xlim=c(20,32), ylim=c(0,4000), pch=20)
lines(targetbw~age, subset(dat, !is.na(targetbw)), col="black")

# Now plot using the right axis
par(new=TRUE)
plot(adfi~age, subset(dat, !is.na(adfi)),
     xlab="", ylab="", xlim=c(20,32), xaxt="n", yaxt="n",
     ylim=c(-50,175), type="s", lty=2)
axis(4, at=c(-50,-25,0,25,50,75,100,125,150,175), col="red", col.axis="red")
mtext("Weight (g)", side=4, line=2, col="red")
lines(adg~age, subset(dat, !is.na(adg)), col="red", type="s", lty=1, lwd=2)
abline(h=c(0,52), col="red")
with(dat, segments(age, 0, age, eggwt, col="red"))

legend(20, -40, c("Body weight", "Target BW", "Feed/day", "Gain/day", "Egg wt"),
       bty="n", cex=.5, ncol=5,
       col=c("black","black","red","red","red"),
       lty=c(-1,1,2,1,1), lwd=c(1,1,1,2,1), pch=c(20,-1,-1,-1,-1))
par(op)
```
Index

* datasets
  acorsi.grayleafspot, 11
  adugna.sorghum, 14
  alwan.lamb, 25
  ansari.wheat.uniformity, 28
  arankacami.groundnut.uniformity, 29
  archbold.apple, 30
  ars.earlywhitecorn96, 32
  baker.barley.uniformity, 35
  baker.strawberry.uniformity, 37
  baker.wheat.uniformity, 38
  bancroft.peanut.uniformity, 39
  barrero.maize, 41
  battese.survey, 45
  beall.webworms, 48
  beaven.barley, 50
  becker.chicken, 51
  besag.bayesian, 52
  besag.beans, 54
  besag.checks, 56
  besag.endive, 59
  besag.triticale, 64
  blackman.wheat, 66
  bliss.borers, 67
  bond.diallel, 69
  bose.multi.uniformity, 71
  box.cork, 72
  bradley.multi.uniformity, 74
  brandle.rape, 76
  brandt.switchback, 78
  broadbalk.wheat, 81
  bryan.corn.uniformity, 83
  buntaran.wheat, 84
  burgueno.alpha, 85
  burgueno.unreplicated, 90
  butron.maize, 91
  byers.apple, 94
  caribbean.maize, 95
  carlson.germination, 97
  carmer.density, 98
  cate.potassium, 100
  chinloy.fractionalfactorial, 101
  christidis.competition, 103
  christidis.cotton.uniformity, 105
  christidis.wheat.uniformity, 106
  cleveland.soil, 107
  cochran.beets, 109
  cochran.crd, 111
  cochran.eelworms, 112
  cochran.factorial, 114
  cochran.latin, 115
  cochran.lattice, 116
  cochran.wireworms, 117
  connolly.potato, 119
  cramer.cucumber, 126
  crampton.pig, 129
  crossa.wheat, 130
  cullis.earlygen, 137
  damesa.maize, 140
  darwin.maize, 141
  dasilva.maize, 143
  dasilva.soybean.uniformity, 145
  davidian.soybean, 146
  davies.pasture.uniformity, 148
  day.wheat.uniformity, 149
  denis.ryegrass, 152
  depalluel.sheep, 153
  devries.pine, 154
  digby.jointregression, 156
  diggle.cow, 158
  draper.safflower.uniformity, 160
  eden.nonnormal, 167
  eden.potato, 170
  eden.tea.uniformity, 172
  edwards.oats, 173
  engelstad.nitro, 175
  fan.stability, 177
federer.diagcheck, 178
fisher.barley, 183
fisher.latin, 185
foulley.calving, 187
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202
fokkelman.mud.chestnut, 205
fouler.mud.chestnut, 205
fox.wheat, 189
garber.multi.uniformity, 190
gartner.corn, 191
garber.multi.uniformity, 190
gift.yield, 202

arankacami.groundnut.uniformity, 29
archbold.apple, 19, 30
ars.earlywhitecorn96, 19, 32
australia.soybean, 19, 33
baker.barley.uniformity, 16, 35
baker.strawberry.uniformity, 16, 37
baker.wheat.uniformity, 16, 38
bancroft.peanut.uniformity, 16, 39
barrero.maize, 19, 23, 41
batchelor.apple.uniformity, (batchelor.uniformity), 43
batchelor.apple.uniformity (batchelor.uniformity), 43
batchelor.lemon.uniformity, 16
batchelor.lemon.uniformity (batchelor.uniformity), 43
batchelor.navel1.uniformity, 16
batchelor.navel1.uniformity (batchelor.uniformity), 43
batchelor.navel2.uniformity, 16
batchelor.navel2.uniformity (batchelor.uniformity), 43
batchelor.valencia.uniformity, 16
batchelor.valencia.uniformity (batchelor.uniformity), 43
batchelor.walnut.uniformity, 16
batchelor.walnut.uniformity (batchelor.uniformity), 43
battese.survey, 19, 45
beall.webworms, 19, 48
beaven.barley, 19, 50
becker.chicken, 18, 51
besag.bayesian, 19, 52
besag.beans, 19, 54
besag.checks, 19, 56
besag.elbatan, 19, 57
besag.endive, 19, 59
besag.met, 19, 60
besag.triticale, 19, 64
blackman.wheat, 19, 66
bliss.borers, 19, 67
bond.diallel, 19, 69
bose.multi.uniformity, 16, 71
box.cork, 18, 72
bradley.multi.uniformity, 16, 74
brandle.rape, 19, 76
brandt.switchback, 18, 78
bridges.cucumber, 19, 79
broadbalk.wheat, 23, 81
bryan.corn.uniformity, 16, 83
buntaran.wheat, 19, 84
burgueno.alpha, 19, 85
burgueno.rowcol, 19, 88
burgueno.unreplicated, 19, 90
byers.apple, 23, 94
caribbean.maize, 19, 95
carlsone.germination, 19, 97
carmer.density, 19, 98
cate.potassium, 23, 100
chinloy.fractionalfactorial, 19, 101
christidis.competition, 19, 103
christidis.cotton.uniformity, 16, 105
christidis.wheat.uniformity, 16, 106
cleveland.soil, 23, 107
cochran.beets, 19, 109
cochran.bib, 19, 110
cochran.crd, 19, 111
cochran.eelworms, 19, 112
cochran.factorial, 19, 114
cochran.latin, 19, 115
cochran.lattice, 19, 116
cochran.wireworms, 19, 117
connolly.potato, 19, 119
cornelius.maize, 19, 121
corsten.interaction, 19, 123
cox.strip.split, 19, 124
cramer.cucumber, 19, 126
crampton.pig, 18, 129
crossa.wheat, 19, 130
crowder.seeds, 19, 132
cullis.earlygen, 19, 137
damesa.maize, 20, 140
darwin.maize, 20, 141
dasilva.maize, 20, 143
dasilva.soybean.uniformity, 16, 145
davidian.soybean, 20, 146
davies.pasture.uniformity, 16, 148
day.wheat.uniformity, 16, 149
denis.missing, 20, 151
denis.ryegrass, 20, 152
depalluel.sheep, 18, 153
devries.pine, 18, 154
digby.jointregression, 20, 156
diggle.cow, 18, 158
INDEX

draper.safflower.uniformity, 16, 160
durban.competition, 20, 162
durban.rowcol, 20, 163
durban.splitplot, 20, 165

eden.nonnormal, 20, 167
eden.potato, 20, 170
eden.tea.uniformity, 16, 172
edwards.oats, 20, 23, 173
engelstad.nitro, 20, 175

fan.stability, 20, 177
federer.diagcheck, 20, 178
federer.tobacco, 20, 181
fisher.barley, 20, 183
fisher.latin, 20, 185
foulley.calving, 18, 187
fox.wheat, 20, 189

garber.multi.uniformity, 16, 190
gartner.corn, 18, 191
gathmann.bt, 20, 194
gauch.soy, 20, 23, 196
george.wheat, 20, 23, 197
giles.wheat, 20, 199
gilmour.serentine, 20, 201
gilmour.slatehall, 20, 203
gomez.fractionalfactorial, 20, 204
gomez.groupsplit, 20, 206
gomez.heterogeneity, 207
gomez.heteroskedastic, 20, 208
gomez.multilocsplitplot, 20, 209
gomez.nitrogen, 20, 211
gomez.nonnormal1, 20, 212
gomez.nonnormal2, 20, 214
gomez.nonnormal3, 20, 215
gomez.rice.uniformity, 16, 217
gomez.seedrate, 20, 218
gomez.splitplot.subsample, 20, 219
gomez.splitsplit, 20, 221
gomez.striplot, 20, 222
gomez.stripssplitplot, 20, 225
gomez.wetdry, 20, 226
gotway.hessianfly, 20, 227
goulden.barley.uniformity, 16, 229
goulden.eggs, 18, 230
goulden.latin, 20, 231
goulden.splitsplit, 20, 232

graybill.heteroskedastic, 20, 234
gregory.cotton, 20, 236
grover.diallel, 20, 238
grover.rcb.subsample, 20, 239
gumpertz.pepper, 20, 240

hadasch.lettuce, 21, 242
hanks.sprinkler, 21, 244
hanover.whitepine, 18, 247
harris.multi.uniformity, 16, 248
harris.wateruse, 18, 251
harrison.priors, 23, 255
hartman.tomato.uniformity, 16, 257
harvey.lsmeans, 18, 259
harville.lamb, 18, 260
hayman.tobacco, 21, 262
hazell.vegetables, 21, 265
heady.fertilizer, 21, 267
heath.raddish.uniformity, 16, 269
henderson.milkfat, 18, 271
hernandez.nitrogen, 21, 272
hessling.argentina, 23, 274
hildebrand.systems, 21, 276
holland.arthropods, 18, 278
holshouser.splitstrip, 21, 280
holtsmark.timothy.uniformity, 16, 281
huehn.wheat, 21, 283
hughes.grapes, 21, 284
hunter.corn, 21, 286

igue.sugarcane.uniformity, 16, 287
ilri.sheep, 18, 288
immer.sugarbeet.uniformity, 16, 291
ivins.herbs, 21, 292
iyer.wheat.uniformity, 16, 294

jansen.apple, 21, 295
jansen.carrot, 21, 297
jansen.strawberry, 21, 298
jayaraman.bamboo, 21, 299
jenkyns.mildew, 21, 301
john.alpha, 21, 302
johnson.blight, 21, 305
johnson.douglasfir, 18, 307
jones.corn.uniformity, 308

kadim.millet.uniformity, 16, 310
kalamkar.potato.uniformity, 16, 311
kalamkar.wheat.uniformity, 16, 312
kang.maize, 21, 314
kang.peanut, 21, 315
karcher.turfgrass, 21, 317
kayad.alfalfa, 18, 318
keen.potatodamage, 21, 320
kempton.barley.uniformity, 16, 321
kempton.competition, 21, 323
kempton.rowcol, 21, 324
kempton.slatehall, 21, 326
kenward.cattle, 18, 328
kerr.sugarcane.uniformity, 16, 330
khin.rice.uniformity, 16, 332
kiesselbach.oats.uniformity, 16, 333
kreusler.maize, 23, 334
kristensen.barley.uniformity, 16, 337
kulkarni.sorghum.uniformity, 16, 338

lambert.soilttemp, 23, 340
lander.multi.uniformity, 16, 341
lasrosas.corn, 18, 343
lavoranti.eucalyptus, 18, 345
laycock.tea.uniformity, 16, 346
lee.potatoblight, 21, 348
lehmann.millet.uniformity, 16, 350
lehner.soybeanmold, 21, 352
lessman.sorghum.uniformity, 16, 354
li.millet.uniformity, 16, 356
libs, 357
lillemo.wheat, 21, 358
lin.superiority, 27, 360
lin.unbalanced, 21, 362
linder.wheat, 21, 364
little.splitblock, 21, 365
loesell.bean.uniformity, 16, 366
lonnquist.maize, 21, 368
lord.rice.uniformity, 16, 370
love.cotton.uniformity, 17, 371
lu.stability, 21, 373
lucas.switchback, 18, 374
lyon.potato.uniformity, 17, 376
lyons.wheat, 21, 377

magistad.pineapple.uniformity, 17, 378
masood.rice.uniformity, 17, 380
mcclelland.corn.uniformity, 17, 381
mcconway.turnip, 21, 383
mcleod.barley, 21, 384
mead.cauliflower, 21, 386
mead.cowpea.maize, 21, 387
mead.germination, 21, 389
mead.lamb, 18, 390
mead.strawberry, 21, 393
mead.turnip, 21, 394
mercer.mangold.uniformity, 17, 395
mercer.wheat.uniformity, 17, 397
miguez.biomass, 21, 398
minnesota.barley.weather, 21, 400
minnesota.barley.yield, 21, 403
montgomery.wheat.uniformity, 17, 405
moore.bushbean.uniformity, 17
moore.bushbean.uniformity
(moore.uniformity), 407
moore.carrot.uniformity, 17
moore.carrot.uniformity
(moore.uniformity), 407
moore.fallcauliflower.uniformity, 17
moore.fallcauliflower.uniformity
(moore.uniformity), 407
moore.polebean.uniformity, 17
moore.polebean.uniformity
(moore.uniformity), 407
moore.springcauliflower.uniformity, 17
moore.springcauliflower.uniformity
(moore.uniformity), 407
moore.sweetcorn.uniformity, 17
moore.sweetcorn.uniformity
(moore.uniformity), 407

nagai.strawberry.uniformity, 17, 409
nair.turmeric.uniformity, 17, 411
narain.sorghum.uniformity, 17, 412
nass.barley, 23
nass.barley(nass.corn), 413
nass.corn, 23, 413
nass.cotton, 23
nass.cotton(nass.corn), 413
nass.hay, 23
nass.hay(nass.corn), 413
nass.rice, 23
nass.rice(nass.corn), 413
nass.sorghum, 23
nass.sorghum(nass.corn), 413
nass.soybean, 23
nass.soybean(nass.corn), 413
nass.wheat, 23
nass.wheat(nass.corn), 413
nebraska.farmincome, 23, 415
nonnecke.peas.uniformity, 17, 416
INDEX

nonnecke.sweetcorn.uniformity, 17, 418
obsi.potato.uniformity, 17, 419
odland.soybean.uniformity, 17, 420
odland.soyhay.uniformity, 17
odland.soyhay.uniformity (odland.soybean.uniformity), 420
omer.sorghum, 21, 422
onofri.winterwheat, 21, 423
ortiz.tomato, 22, 424
pacheco.soybean, 22, 427
paez.coffee.uniformity, 17, 428
panse.cotton.uniformity, 17, 429
parker.orange.uniformity, 17, 431
patterson.switchback, 18, 433
pearce.apple, 18, 434
pearl.kernels, 23, 436
pederson.lettuce.repeated, 22, 437
perry.springwheat, 22, 439
petersen.sorghum.cowpea, 22, 441
piepho.barley.uniformity, 17, 442
piepho.cocksfoot, 22, 444
petersen.sorghum.uniformity, 17, 445
petersen.safflower.uniformity, 17, 447
reid.grasses, 22, 448
ridout.appleshoots, 22, 449
robinson.peanut.uniformity, 17, 452
rothamsted.brussels, 22, 453
rothamsted.oats, 22, 454
ryder.groundnut, 22, 456
salmon.bunt, 22, 458
sawyer.multi.uniformity, 17, 459
sayer.sugarcane.uniformity, 17, 462
senshu.rice, 22, 463
shafi.tomato.uniformity, 17, 465
shafi.rapeseed, 22, 23, 466
shaw.oats, 22, 23, 468
silva.cotton, 22, 469
sinclair.clover, 22, 472
smith.beans.uniformity, 17, 474
smith.corn.uniformity, 17, 476
sneadecor.asparagus, 22, 478
snijders.fusarium, 22, 481
stephens.sorghum.uniformity, 17, 483
steptoe.morex.genotype
(steptoe.morex.pheno), 484
steptoe.morex.pheno, 22, 484
stickler.sorghum.uniformity, 17, 487
stirret.borers, 23, 489
streibig.competition, 22, 490
strickland.apple.uniformity, 17, 492
strickland.grape.uniformity, 17, 493
strickland.peach.uniformity, 17, 494
strickland.tomato.uniformity, 17, 495
stroup.nin, 22, 496
stroup.splitplot, 22, 500
student.barley, 22, 503
tai.potato, 22, 505
talbot.potato, 22, 506
theobald.barley, 22, 508
theobald.covariate, 22, 509
thompson.cornsoy, 22, 511
turner.herbicide, 23, 513
urquhart.feedlot, 18, 515
usgs.herbicides, 23, 517
vaneewijk.drymatter, 22, 519
vaneewijk.fusarium, 22, 521
vaneewijk.nematodes, 22, 522
vargas.txe, 22, 524
vargas.wheat1, 22, 527
vargas.wheat2, 22, 529
verbyla.lupin, 22, 23, 531
vold.longterm, 22, 535
vsn.lupin3, 22, 537
wallace.iowaland, 23, 540
walsh.cottonprice, 23, 542
wassom.brome.uniformity, 17, 543
waynick.soil, 23, 544
wedderburn.barley, 22, 545
weiss.incblock, 22, 547
weiss.lattice, 22, 549
welch.bermudagrass, 22, 550
wheat2, 498
wheatley.carrot, 22, 553
wiebe.wheat.uniformity, 17, 555
wiedemann.safflower.uniformity, 17, 557
williams.barley.uniformity, 17, 559
williams.cotton.uniformity, 17, 560
williams.trees, 18, 561
woodman.pig, 18, 562
wyatt.multi.uniformity, 17, 564

yan.winterwheat, 22, 566
yang.barley, 22, 567
yates.missing, 22, 568
yates.oats, 23, 570

zuidhof.broiler, 18, 573