Package 'spatstat.data'

October 21, 2022

Version 3.0-0

Date 2022-10-20

Title Datasets for 'spatstat' Family

Maintainer Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>

Depends R (>= 3.5.0)

Imports spatstat.utils (>= 3.0-0), Matrix

Suggests spatstat.geom, spatstat.random, spatstat.explore,
     spatstat.model, spatstat.linnet

Additional_repositories <https://spatstat.r-universe.dev>

Description Contains all the datasets for the 'spatstat' family of packages.

License GPL (>= 2)

URL <http://spatstat.org/>

LazyData true

NeedsCompilation no

ByteCompile true

BugReports <https://github.com/spatstat/spatstat.data/issues>

Author Adrian Baddeley [aut, cre] (<<https://orcid.org/0000-0001-9499-8382>>),
     Rolf Turner [aut] (<<https://orcid.org/0000-0001-5521-5218>>),
     Ege Rubak [aut] (<<https://orcid.org/0000-0002-6675-533X>>),
     W Aherne [ctb],
     Freda Alexander [ctb],
     Qi Wei Ang [ctb],
     Sourav Banerjee [ctb],
     Mark Berman [ctb],
     R Bernhardt [ctb],
     Thomas Berndtsen [ctb],
     Andrew Bevan [ctb],
     Jeffrey Betts [ctb],
     Ray Cartwright [ctb],
     Lucia Cobo Sanchez [ctb],

1

Richard Condit [ctb],
Francis Crick [ctb],
Marcelino de la Cruz Rot [ctb],
Jack Cuzick [ctb],
Tilman Davies [ctb],
Peter Diggle [ctb],
Michael Drinkwater [ctb],
Stephen Eglen [ctb],
Robert Edwards [ctb],
AE Esler [ctb],
Gregory Evans [ctb],
Bernard Fingleton [ctb],
Olivier Flores [ctb],
David Ford [ctb],
Robin Foster [ctb],
Janet Franklin [ctb],
Neba Funwi-Gabga [ctb],
DJ Gerrard [ctb],
Andy Green [ctb],
Tim Griffin [ctb],
Ute Hahn [ctb],
RD Harkness [ctb],
Arthur Hickman [ctb],
Stephen Hubbell [ctb],
Austin Hughes [ctb],
Jonathan Huntington [ctb],
MJ Hutchings [ctb],
Jackie Inwald [ctb],
Valerie Isham [ctb],
Aruna Jammalamadaka [ctb],
Carl Knox-Robinson [ctb],
Mahdieh Khanmohammadi [ctb],
Tero Kokkila [ctb],
Bas Kooijman [ctb],
Kenneth Kosik [ctb],
Peter Kovesi [ctb],
Lily Kozmian-Ledward [ctb],
Robert Lamb [ctb],
NA Laskurain [ctb],
George Leser [ctb],
Marie-Colette van Lieshout [ctb],
AF Mark [ctb],
Jorge Mateu [ctb],
Annikki Makela [ctb],
Enrique Miranda [ctb],
Nicoletta Nava [ctb],
M Numata [ctb],
Matti Nummelin [ctb],

Jens Randel Nyengaard [ctb],
Yosihiko Ogata [ctb],
Si Palmer [ctb],
Antti Penttinen [ctb],
Sandra Pereira [ctb],
Nicolas Picard [ctb],
William Platt [ctb],
Stephen Rathbun [ctb],
Brian Ripley [ctb],
Roger Sainsbury [ctb],
Dietrich Stoyan [ctb],
David Strauss [ctb],
L Strand [ctb],
Masaharu Tanemura [ctb],
Graham Upton [ctb],
Bill Venables [ctb],
Sasha Voss [ctb],
Rasmus Waagepetersen [ctb],
Keith Watkins [ctb],
H Wendrock [ctb]

R **topics documented:**

---

spatstat.data-package   *The spatstat.data Package*

Description

The **spatstat.data** package contains the datasets for the **spatstat** family of packages.

Details

The **spatstat.data** package contains the datasets for the **spatstat** family of packages.

These are spatial datasets; they are objects belonging to classes of spatial data defined in other sub-packages of the **spatstat** family. In order to handle these datasets correctly, we recommend loading the **spatstat** package.

Licence

This library and its documentation are usable under the terms of the "GNU General Public License", a copy of which is distributed with R.

Author(s)

Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>, Rolf Turner <r.turner@auckland.ac.nz> and Ege Rubak <rubak@math.aau.dk>.

---

amacrine   *Hughes' Amacrine Cell Data*

Description

Austin Hughes' data: a point pattern of displaced amacrine cells in the retina of a rabbit. A marked point pattern.

Usage

```
data(amacrine)
```

Format

An object of class ″ppp″ representing the point pattern of cell locations. Entries include

x	Cartesian x-coordinate of cell
y	Cartesian y-coordinate of cell
marks	factor with levels off and on
	indicating "off" and "on" cells

See `ppp.object` for details of the format.

Notes

Austin Hughes' data: a point pattern of displaced amacrine cells in the retina of a rabbit. 152 "on" cells and 142 "off" cells in a rectangular sampling frame.

The true dimensions of the rectangle are 1060 by 662 microns. The coordinates here are scaled to a rectangle of height 1 and width $1060/662 = 1.601$ so the unit of measurement is approximately 662 microns.

The data were analysed by Diggle (1986).

Source

Peter Diggle, personal communication

References

Diggle, P. J. (1986). Displaced amacrine cells in the retina of a rabbit: analysis of a bivariate spatial point pattern. *J. Neurosci. Meth.* **18**, 115–125.

Examples

```
  if(require(spatstat.geom)) {
amacrine
(Ama <- rescale(amacrine))
  }
```

---

anemones   *Beadlet Anemones Data*

Description

These data give the spatial locations and diameters of sea anemones (beadlet anemone *Actinia equina*) in a sample plot on the north face of a boulder, well above low tide level, at Quiberon (Bretagne, France) in May 1976.

The data were originally described and discussed by Kooijman (1979a). Kooijman (1979b) shows a hand-drawn plot of the original data. The data are discussed by Upton and Fingleton (1985) as Example 1.8 on pages 64–67.

The anemones dataset is taken directly from Table 1.11 of Upton and Fingleton (1985). The co-ordinates and diameters are integer multiples of an idiosyncratic unit of length. The boundary is a rectangle 280 by 180 units.

Usage

```
data(anemones)
```

Format

anemones is an object of class ″ppp″ representing the point pattern of anemone locations. It is a marked point pattern with numeric marks representing anemone diameter. See ppp.object for details of the format.

Units

There is some confusion about the correct physical scale for these data. According to Upton and Fingleton (1985), one unit in the dataset is approximately 0.475 cm. According to Kooijman (1979a, 1979b) and also quoted by Upton and Fingleton (1985), the physical size of the sample plot was 14.5 by 9.75 decimetres (145 by 97.5 centimetres). However if the data are plotted at this scale, they are too small for a rectangle of this size, and the appearance of the plot does not match the original hand-drawn plot in Kooijman (1979b). To avoid confusion, we have not assigned a unit scale to this dataset.

Source

Table 1.11 on pages 62–63 of Upton and Fingleton (1985), who acknowledge Kooijman (1979a) as the source.

References

Kooijman, S.A.L.M. (1979a) The description of point patterns. In *Spatial and temporal analysis in ecology* (ed. R.M. Cormack and J.K. Ord), International Cooperative Publishing House, Fairland, Maryland, USA. Pages 305–332.

Kooijman, S.A.L.M. (1979b) Inference about dispersal patterns. *Acta Biotheoretica* **28**, 149–189.

Upton, G.J.G. and Fingleton, B. (1985) *Spatial data analysis by example*. Volume 1: Point pattern and quantitative data. John Wiley and Sons, Chichester.

Examples

```
data(anemones)
if(require(spatstat.geom)) {
# plot diameters on same scale as x, y
plot(anemones, markscale=1)
}
```

---

ants   *Harkness-Isham ants' nests data*

Description

These data give the spatial locations of nests of two species of ants, *Messor wasmanni* and *Cataglyphis bicolor*, recorded by Professor R.D. Harkness at a site in northern Greece, and described in Harkness & Isham (1983). The full dataset (supplied here) has an irregular polygonal boundary, while most analyses have been confined to two rectangular subsets of the pattern (also supplied here).

The harvester ant *M. wasmanni* collects seeds for food and builds a nest composed mainly of seed husks. *C. bicolor* is a heat-tolerant desert foraging ant which eats dead insects and other arthropods. Interest focuses on whether there is evidence in the data for intra-species competition between *Messor* nests (i.e. competition for resources) and for preferential placement of *Cataglyphis* nests in the vicinity of *Messor* nests.

The full dataset is displayed in Figure 1 of Harkness & Isham (1983). See **Usage** below to produce a comparable plot. It comprises 97 nests (68 Messor and 29 Cataglyphis) inside an irregular convex polygonal boundary, together with annotations showing a foot track through the region, the boundary between field and scrub areas inside the region, and indicating the two rectangular subregions A and B used in their analysis.

Rectangular subsets of the data were analysed by Harkness & Isham (1983), Isham (1984), Takacs & Fiksel (1986), S\"arkk\"a (1993, section 5.3), H\"ogmander and S\"arkk\"a (1999) and Baddeley & Turner (2000). The full dataset (inside its irregular boundary) was first analysed by Baddeley & Turner (2005b).

The dataset `ants` is the full point pattern enclosed by the irregular polygonal boundary. The x and y coordinates are eastings (E-W) and northings (N-S) scaled so that 1 unit equals 0.5 feet. This is a multitype point pattern object, each point carrying a mark indicating the ant species (with levels `Cataglyphis` and `Messor`).

The dataset `ants.extra` is a list of auxiliary information:

A **and** B The subsets of the pattern within the rectangles A and B demarcated in Figure 1 of Harkness & Isham (1983). These are multitype point pattern objects.

`trackNE` **and** `trackSW` coordinates of two straight lines bounding the foot track.

`fieldscrub` The endpoints of a straight line separating the regions of 'field' and 'scrub': scrub to the North and field to the South.

`side` A `function(x,y)` that determines whether the location `(x,y)` is in the scrub or the field. The function can be applied to numeric vectors x and y, and returns a factor with levels `"scrub"` and `"field"`. This function is useful as a spatial covariate.

`plotit` A function which produces a plot of the full dataset.

Usage

```
data(ants)
```

Format

`ants` is an object of class `"ppp"` representing the full point pattern of ants' nests. See [ppp.object](#) for details of the format. The coordinates are scaled so that 1 unit equals 0.5 feet. The points are marked by species (with levels `Cataglyphis` and `Messor`).

`ants.extra` is a list with entries

A point pattern of class `"ppp"`

B point pattern of class `"ppp"`

trackNE data in format `list(x=numeric(2),y=numeric(2))` giving the two endpoints of line markings

trackSW data in format `list(x=numeric(2),y=numeric(2))` giving the two endpoints of line markings

fieldscrub data in format `list(x=numeric(2),y=numeric(2))` giving the two endpoints of line markings

side Function with arguments `x,y`

plotit Function

Source

Harkness and Isham (1983). Nest coordinates kindly provided by Prof Valerie Isham. Polygon coordinates digitised by Adrian Baddeley from a reprint of Harkness & Isham (1983).

References

Baddeley, A. and Turner, R. (2000) Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics* **42**, 283–322.

Baddeley, A. and Turner, R. (2005a) Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**:6, 1–42. URL: `www.jstatsoft.org`, ISSN: 1548-7660.

Baddeley, A. and Turner, R. (2005b) Modelling spatial point patterns in R. In: A. Baddeley, P. Gregori, J. Mateu, R. Stoica, and D. Stoyan, editors, *Case Studies in Spatial Point Pattern Modelling*, Lecture Notes in Statistics number 185. Pages 23–74. Springer-Verlag, New York, 2006. ISBN: 0-387-28311-0.

Harkness, R.D. and Isham, V. (1983) A bivariate spatial point pattern of ants' nests. *Applied Statistics* **32**, 293–303.

Högmander, H. and Särkkä, A. (1999) Multitype spatial point patterns with hierarchical interactions. *Biometrics* **55**, 1051–1058.

Isham, V.S. (1984) Multitype Markov point processes: some approximations. *Proceedings of the Royal Society of London, Series A*, **391**, 39–53.

Takacs, R. and Fiksel, T. (1986) Interaction pair-potentials for a system of ants' nests. *Biometrical Journal* **28**, 1007–1013.

Särkkä, A. (1993) *Pseudo-likelihood approach for pair potential estimation of Gibbs processes*. Number 22 in Jyväskylä Studies in Computer Science, Economics and Statistics. University of Jyväskylä, Finland.

Examples

```
if(require(spatstat.geom)) {

# Equivalent to Figure 1 of Harkness and Isham (1983)

data(ants)
ants.extra$plotit()

# Data in subrectangle A, rotated
# Approximate data used by Sarkka (1993)

angle <- atan(diff(ants.extra$fieldscrub$y)/diff(ants.extra$fieldscrub$x))
```

```
plot(rotate(ants.extra$A, -angle))

# Approximate window used by Takacs and Fiksel (1986)

tfwindow <- boundingbox(Window(ants))
antsTF <- ppp(ants$x, ants$y, window=tfwindow)
plot(antsTF)
}
```

---

austates   *Australian States and Mainland Territories*

Description

The states and large mainland territories of Australia are represented as polygonal regions forming
a tessellation.

Usage

```
data(austates)
```

Format

Object of class "tess".

Details

Western Australia, South Australia, Queensland, New South Wales, Victoria and Tasmania (which
are states of Australia) and the Northern Territory (which is a 'territory' of Australia) are represented
as polygonal regions.

Offshore territories, and smaller mainland territories, are not shown.

The dataset austates is a tessellation object (class "tess") whose tiles are the states and territories.

The coordinates are latitude and longitude in degrees, so the space is effectively a Mercator projec-
tion of the earth.

Source

Obtained from the **oz** package and reformatted.

Examples

```
data(austates)
if(require(spatstat.geom)) {
plot(austates)
}
```

---

bdspots   *Breakdown Spots in Microelectronic Materials*

Description

A list of three point patterns, each giving the locations of electrical breakdown spots on a circular electrode in a microelectronic capacitor.

Usage

```
data(bdspots)
```

Format

A list (of class `"listof"`) of three spatial point patterns, each representing the spatial locations of breakdown spots on an electrode. The three electrodes are circular discs, of radii 169, 282 and 423 microns respectively. Spatial coordinates are given in microns.

Details

The application of successive voltage sweeps to the metal gate electrode of a microelectronic capacitor generates multiple breakdown spots on the electrode. The spatial distribution of these breakdown spots in MIM (metal-insulator-metal) and MIS (metal-insulator-semiconductor) structures was observed and analysed by Miranda et al (2010, 2013) and Saura et al (2013a, 2013b, 2014).

The data given here are the breakdown spot patterns for three circular electrodes of different radii, 169, 282 and 423 microns respectively, in MIM structures analysed in Saura et al (2013a).

Source

Professor Enrique Miranda, Departament d'Enginyeria Electrònica, Escola d'Enginyeria, Universitat Autònoma de Barcelona, Barcelona, Spain.

References

Miranda, E. and O'Connor, E. and Hurley, P.K. (2010) Simulation of the breakdown spots spatial distribution in high-*K* dielectrics and model validation using the **spatstat** package for *R* language. *ECS Transactions* **33** (3) 557–562.

Miranda, E., Jiménez, D., Suñé, J., O'Connor, E., Monaghan, S., Povey, I., Cherkaoui, K. and Hurley, P. K. (2013) Nonhomogeneous spatial distribution of filamentary leakage current paths in circular area Pt/HfO2/Pt capacitors. *J. Vac. Sci. Technol. B* **31**, 01A107.

Saura, X., Suñé, J., Monaghan, S., Hurley, P.K. and Miranda, E. (2013a) Analysis of the breakdown spot spatial distribution in Pt/HfO2/Pt capacitors using nearest neighbor statistics. *J. Appl. Phys.* **114**, 154112.

Saura, X., Moix, D., Suñé, J., Hurley, P.K. and Miranda, E. (2013b) Direct observation of the generation of breakdown spots in MIM structures under constant voltage stress. *Microelectronics Reliability* **53**, 1257–1260.

Saura, X., Monaghan, S., Hurley, P.K., Suñé, J. and Miranda, E. (2014) Failure analysis of MIM and MIS structures using point-to-event distance and angular probability distributions. *IEEE Transactions on Devices and Materials Reliability* **14** (4) 1080–1090.

Examples

```
data(bdspots)
  if(require(spatstat.geom)) {
plot(bdspots, equal.scales=TRUE)
  }
```

---

bei   *Tropical rain forest trees*

Description

A point pattern giving the locations of 3605 trees in a tropical rain forest. Accompanied by covariate data giving the elevation (altitude) and slope of elevation in the study region.

Usage

```
data(bei)
```

Format

`bei` is an object of class ″ppp″ representing the point pattern of tree locations. See `ppp.object` for details of the format.

`bei.extra` is a list containing two pixel images, `elev` (elevation in metres) and `grad` (norm of elevation gradient). These pixel images are objects of class ″im″, see `im.object`.

Notes

The dataset `bei` gives the positions of 3605 trees of the species *Beilschmiedia pendula* (Lauraceae) in a 1000 by 500 metre rectangular sampling region in the tropical rainforest of Barro Colorado Island.

The accompanying dataset `bei.extra` gives information about the altitude (elevation) in the study region. It is a list containing two pixel images, `elev` (elevation in metres) and `grad` (norm of elevation gradient).

All spatial coordinates are given in metres.

These data are part of a much larger dataset containing the positions of hundreds of thousands of trees belong to thousands of species; see Hubbell and Foster (1983), Condit, Hubbell and Foster (1996) and Condit (1998).

The present data were analysed by Møller and Waagepetersen (2007).

Source

Hubbell and Foster (1983), Condit, Hubbell and Foster (1996) and Condit (1998). Data files kindly supplied by Rasmus Waagepetersen. The data were collected in the forest dynamics plot of Barro Colorado Island. The study was made possible through the generous support of the U.S. National Science Foundation, the John D. and Catherine T. MacArthur Foundation, and the Smithsonian Tropical Research Institute.

References

Condit, R. (1998) *Tropical Forest Census Plots*. Springer-Verlag, Berlin and R.G. Landes Company, Georgetown, Texas.

Condit, R., Hubbell, S.P and Foster, R.B. (1996) Changes in tree species abundance in a neotropical forest: impact of climate change. *Journal of Tropical Ecology* **12**, 231–256.

Hubbell, S.P and Foster, R.B. (1983) Diversity of canopy trees in a neotropical forest and implications for conservation. In: *Tropical Rain Forest: Ecology and Management* (eds. S.L. Sutton, T.C. Whitmore and A.C. Chadwick), Blackwell Scientific Publications, Oxford, 25–41.

Møller, J. and Waagepetersen, R.P. (2007) Modern spatial point process modelling and inference (with discussion). *Scandinavian Journal of Statistics* **34**, 643–711.

---

betacells	*Beta Ganglion Cells in Cat Retina*

Description

Point pattern of cells in the retina, each cell classified as 'on' or 'off' and labelled with the cell profile area.

Usage

```
data(betacells)
```

Format

betacells is an object of class "ppp" representing the point pattern of cell locations. Entries include

x	Cartesian x-coordinate of cell
y	Cartesian y-coordinate of cell
marks	data frame of marks

Cartesian coordinates are given in microns.

The data frame of marks has two columns:

type	factor with levels off and on indicating "off" and "on" cells
area	numeric vector giving the

areas of cell profiles (in square microns)

See `ppp.object` for details of the format.

Notes

This is a new, corrected version of the old dataset `ganglia`. See below.

These data represent a pattern of beta-type ganglion cells in the retina of a cat recorded by W\"assle et al. (1981). Beta cells are associated with the resolution of fine detail in the cat's visual system. They can be classified anatomically as "on" or "off".

Statistical independence of the arrangement of the "on"- and "off"-components would strengthen the evidence for Hering's (1878) 'opponent theory' that there are two separate channels for sensing "brightness" and "darkness". See W\"assle et al (1981). There is considerable current interest in the arrangement of cell mosaics in the retina, see Rockhill et al (2000).

The dataset is a marked point pattern giving the locations, types ("on" or "off"), and profile areas of beta cells observed in a rectangle of dimensions 750×990 microns. Coordinates are given in microns (thousandths of a millimetre) and areas are given in square microns.

The original source is Figure 6 of W\"assle et al (1981), which is a manual drawing of the beta mosaic observed in a microscope field-of-view of a whole mount of the retina. Thus, all beta cells in the retina were effectively projected onto the same two-dimensional plane.

The data were scanned in 2004 by Stephen Eglen from Figure 6(a) of W\"assle et al (1981). Image analysis software was used to identify the soma (cell body). The x, y location of each cell was taken to be the centroid of the soma. The type of each cell ("on" or 'off") was identified by referring to Figures 6(b) and 6(d). The area of each soma (in square microns) was also computed.

Note that this is a corrected version of the `ganglia` dataset provided in earlier versions of **spatstat**. The earlier data `ganglia` were not faithful to the scale in the original paper and contain some scanning errors.

Source

W\"assle et al (1981), Figure 6(a), scanned and processed by Stephen Eglen <S.J.Eglen@damtp.cam.ac.uk>.

References

Hering, E. (1878) Zur Lehre von Lichtsinn. Vienna.

Van Lieshout, M.N.M. and Baddeley, A.J. (1999) Indices of dependence between types in multivariate point patterns. *Scandinavian Journal of Statistics* **26**, 511–532.

Rockhill, R.L., Euler, T. and Masland, R.H. (2000) Spatial order within but not between types of retinal neurons. *Proc. Nat. Acad. Sci. USA* **97**(5), 2303–2307.

W\"assle, H., Boycott, B. B. & Illing, R.-B. (1981). Morphology and mosaic of on- and off-beta cells in the cat retina and some functional considerations. *Proc. Roy. Soc. London Ser. B* **212**, 177–195.

Examples

```
  plot(betacells)
 if(require(spatstat.geom)) {
 area <- marks(betacells)$area
 plot(betacells %mark% sqrt(area/pi), markscale=1)
  }
```

---

bramblecanes   *Hutchings' Bramble Canes data*

Description

Data giving the locations and ages of bramble canes in a field. A marked point pattern.

Usage

```
data(bramblecanes)
```

Format

An object of class ″ppp″ representing the point pattern of plant locations. Entries include

x	Cartesian x-coordinate of plant
y	Cartesian y-coordinate of plant
marks	factor with levels 0,1, 2 indicating age

See `ppp.object` for details of the format.

Notes

These data record the (x, y) locations and ages of bramble canes in a field 9 metres square, rescaled to the unit square. The canes were classified according to age as either newly emergent, one or two years old. These are encoded as marks 0, 1 and 2 respectively in the dataset.

The data were recorded and analysed by Hutchings (1979) and further analysed by Diggle (1981a, 1981b, 1983), Diggle and Milne (1983), and Van Lieshout and Baddeley (1999). All analyses found that the pattern of newly emergent canes exhibits clustering, which Hutchings attributes to "vigorous vegetative reproduction".

Source

Hutchings (1979), data published in Diggle (1983)

References

Diggle, P. J. (1981a) Some graphical methods in the analysis of spatial point patterns. In *Interpreting multivariate data*, V. Barnett (Ed.) John Wiley and Sons.

Diggle, P. J. (1981b). Statistical analysis of spatial point patterns. *N.Z. Statist.* **16**, 22–41.

Diggle, P.J. (1983) *Statistical analysis of spatial point patterns*. Academic Press.

Diggle, P. J. and Milne, R. K. (1983) Bivariate Cox processes: some models for bivariate spatial point patterns. *Journal of the Royal Statistical Soc. Series B* **45**, 11–21.

Hutchings, M. J. (1979) Standing crop and pattern in pure stands of Mercurialis perennis and Rubus fruticosus in mixed deciduous woodland. *Oikos* **31**, 351–357.

Van Lieshout, M.N.M. and Baddeley, A.J. (1999) Indices of dependence between types in multivariate point patterns. *Scandinavian Journal of Statistics* **26**, 511–532.

Examples

```
if(require(spatstat.geom)) {
 bramblecanes
 # convert coordinates to metres
 (Bram <- rescale(bramblecanes))
 }
```

---

bronzefilter   *Bronze gradient filter data*

Description

These data represent a spatially inhomogeneous pattern of circular section profiles of particles, observed in a longitudinal plane section through a gradient sinter filter made from bronze powder, prepared by Ricardo Bernhardt, Dresden.

The material was produced by sedimentation of bronze powder with varying grain diameter and subsequent sintering, as described in Bernhardt et al. (1997).

The data are supplied as a marked point pattern of circle centres marked by circle radii. The co-ordinates of the centres and the radii are recorded in mm. The field of view is an 18×7 mm rectangle.

The data were first analysed by Hahn et al. (1999).

Usage

```
data(bronzefilter)
```

Format

An object of class ″ppp″ representing the point pattern of cell locations. Entries include

x	Cartesian x-coordinate of bronze grain profile centre
y	Cartesian y-coordinate of bronze grain profile centre
marks	radius of bronze grain profile

See `ppp.object` for details of the format. All coordinates are recorded in mm.

Source

R.\ Bernhardt (section image), H.\ Wendrock (coordinate measurement). Adjusted, formatted and communicated by U.\ Hahn.

References

Bernhardt, R., Meyer-Oblersleben, F. and Kieback, B. (1997) Fundamental investigation on the preparation of gradient structures by sedimentation of different powder fractions under gravity. *Proc. of the 4th Int. Conf. On Composite Engineering, July 6–12 1997, ICCE/4*, Hawaii, Ed. David Hui, 147–148.

Hahn U., Micheletti, A., Pohlink, R., Stoyan D. and Wendrock, H.(1999) Stereological analysis and modelling of gradient structures. *Journal of Microscopy*, **195**, 113–124.

Examples

```
data(bronzefilter)
if(require(spatstat.geom)) {
plot(bronzefilter, markscale=2)
}
```

---

btb   *Bovine Tuberculosis Data*

Description

Geospatial data of 873 farm locations with detected bovine tuberculosis in Cornwall, UK, over the years 1989-2002. This data-set was first analysed in Diggle, Zheng and Durr (2005).

Usage

```
data(btb)
```

Format

Loading this dataset supplies the point pattern btb and the additional object btb.extra.

btb is a marked point pattern (see [`ppp.object`](#)) containing 873 points. Its spatial coordinates are Eastings and Northings in kilometres giving the farm locations. It has two columns of marks:

year	Year of detection: a factor with levels 1989 to 2002
spoligotype	Spoligotype of tuberculosis: a factor with four levels "9", "12", "15", "20"

Loading the dataset btb will also load the object btb.extra containing additional data. This is a list (of class "solist") containing two elements,

standard	The standard version of the BTB dataset used in many publications. This is a marked point pattern, identical to bt
full	A more extensive dataset compiled from files supplied by Professor Diggle. This is a marked point pattern, identic

Source

Professor Peter Diggle.

Roger Sainsbury of the UK's State Veterinary Service helped to collect the data-set. Jackie Inwald and Si Palmer of the Department of Bacterial Diseases, Veterinary Laboratories Agency, Weybridge, UK carried out the spoligotyping.

Peter Diggle supplied the point coordinates, spoligotype data and year data, and the coordinates of the window used in `btb.extra`.

Tilman Davies drew the finer window used in `btb`.

References

Diggle, P.J., Zheng, P. and Durr, P. (2005) Nonparametric estimation of spatial segregation in a multivariate point process: bovine tuberculosis in Cornwall, UK. *Applied Statistics*, **54**, 645–658.

Examples

```
if(require(spatstat.geom)) {
  summary(btb)
  plot(subset(btb, select=spoligotype), cols=2:5)
}
```

---

cells   *Biological Cells Point Pattern*

Description

The data record the locations of the centres of 42 biological cells observed under optical microscopy in a histological section. The microscope field-of-view has been rescaled to the unit square.

The data were recorded by F.H.C. Crick and B.D. Ripley, and analysed in Ripley (1977, 1981) and Diggle (1983). They are often used as a canonical example of an 'ordered' point pattern.

Usage

```
data(cells)
```

Format

An object of class "ppp" representing the point pattern of cell centres. See `ppp.object` for details of the format.

Source

Crick and Ripley, see Ripley (1977)

References

Diggle, P.J. (1983) *Statistical analysis of spatial point patterns*. Academic Press.

Ripley, B.D. (1977) Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B* **39**, 172–212.

Ripley, B.D. (1981) *Spatial statistics*. John Wiley and Sons.

---

cetaceans	*Point patterns of whale and dolphin sightings.*

Description

Nine (independent replicate) point patterns of whale and dolphin sightings obtained from aircraft flying along eight parallel transects in the region of Great Barrier Island, the Hauraki Gulf and the Coromandel Peninsula (New Zealand). Most of the transects are interrupted by portions of land mass. Observations were recorded within narrow rectangles of total width 840 metres (420 metres on each side of the transect).

Usage

```
data(cetaceans)
```

Format

The object `cetaceans` is a *hyperframe* (see [hyperframe](#)()) with 9 rows and 4 columns. Each row of this hyperframe represents a replicate survey. The columns are `whales`, `dolphins`, `fish` and `plankton`.

Each entry in the hyperframe is a point pattern. The `dolphins` column consists of marked patterns (with marks having levels dd and tt) while the other columns contain unmarked point patterns.

The object `cetaceans.extra` is a list containing auxiliary data. It currently contains only one entry, `patterns`, which contains the same information as `cetaceans` in another form. This is a list, of class solist ("spatial object list"; see [solist](#)(), [as.solist](#)()). It is a list of length 9, in which each entry is a marked point pattern, representing the result of one survey. Each pattern was obtained by superimposing the `whales`, `dolphins`, `fish` and `plankton` patterns from the corresponding row of `cetaceans`. The marks of these patterns have levels be, dd, fi, tt and zo.

Details

The data were obtained from nine aerial surveys, conducted from 02/12/2013 to 22/04/2014. Each survey was conducted over the course of a single day. The gap between successive surveys ranged from two to six weeks (making it "not unreasonable" to treat the patterns obtained as being independent). The marks of the patterns referred to above may be interpreted as follows:

- be: whales — Bryde's whale (Balaenoptera edeni)
- dd: dolphins — Common dolphin (Delphinus delphis)
- fi: fish — Any species that forms schools

- `tt`: dolphins — Bottlenose dolphin (Tursiops truncatus)
- `zo`: plankton — Zooplankton

The window for the point patterns in these data sets is of type `polygonal` and consists of a number of thin rectangular strips. These are arranged along eight parallel transects.

The units in which the patterns are presented are kilometres.

These data are rather "sparse". For example there are a total of only eight whale observations in the entire data set (all nine surveys). Thus conclusions drawn from these data should be treated with even more than the usual amount of circumspection.

Source

These data were kindly supplied by Lily Kozmian-Ledward, who studied them in the course of writing her Master's thesis at the University of Auckland, under the joint supervision of Dr. Rochelle Constantine, University of Auckland and Dr Leigh Torres, Oregon State University.

References

Kozmian-Ledward, L. (2014). *Spatial ecology of cetaceans in the Hauraki Gulf, New Zealand.* Unpublished MSc thesis, University of Auckland, New Zealand.

Examples

```
if(require(spatstat.model)) {
   cet <- cetaceans
   cet$dMplank <- with(cet, distfun(plankton, undef=20))
   cet$dMfish <- with(cet, distfun(fish, undef=20))
   fit.whales <- mppm(whales ~ dMplank + dMfish,data=cet)
   anova(fit.whales,test="Chi")
   # Note that inference is *conditional* on the fish and
   # plankton patterns.
   cetPats <- cetaceans.extra$patterns
   plot(Window(cetPats[[1]]),main="The window")
   plot(cetPats,nrows=3,main="All data")
}
```

---

chicago   *Chicago Crime Data*

Description

This dataset is a record of spatial locations of crimes reported in the period 25 April to 8 May 2002, in an area of Chicago (Illinois, USA) close to the University of Chicago. The original crime map was published in the Chicago Weekly News in 2002.

The data give the spatial location (street address) of each crime report, and the type of crime. The type labels are interpreted as follows:

assault   battery/assault

burglary	burglary
cartheft	motor vehicle theft
damage	criminal damage
robbery	robbery
theft	theft
trespass	criminal trespass

All crimes occurred on or near a street. The data give the coordinates of all streets in the survey area, and their connectivity.

Spatial coordinates are expressed in feet (one foot is 0.3048 metres).

The dataset `chicago` is an object of class ″lpp″ representing a point pattern on a linear network. See [lpp](#) for further information on the format.

These data were published and analysed in Ang, Baddeley and Nair (2012).

Usage

```
data(chicago)
```

Format

Object of class ″lpp″. See [lpp](#).

Source

Chicago Weekly News, 2002. Manually digitised by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>.

References

Ang, Q.W. (2010) *Statistical methodology for events on a network.* Master's thesis, School of Mathematics and Statistics, University of Western Australia.

Ang, Q.W., Baddeley, A. and Nair, G. (2012) Geometrically corrected second-order analysis of events on a linear network, with applications to ecology and criminology. *Scandinavian Journal of Statistics* **39**, 591–617.

Chicago Weekly News website: <http://www.chicagoweeklynews.com>

Examples

```
data(chicago)
  if(require(spatstat.linnet)) {
plot(chicago)
plot(as.linnet(chicago), main="Chicago Street Crimes",col="green")
plot(as.ppp(chicago), add=TRUE, col="red", chars=c(16,2,22,17,24,15,6))
  }
```

---

chorley   *Chorley-Ribble Cancer Data*

Description

Spatial locations of cases of cancer of the larynx and cancer of the lung, and the location of a disused industrial incinerator. A marked point pattern.

Usage

```
data(chorley)
```

Format

The dataset chorley is an object of class "ppp" representing a marked point pattern. Entries include

x	Cartesian x-coordinate of home address
y	Cartesian y-coordinate of home address
marks	factor with levels larynx and lung
	indicating whether this is a case of cancer of the larynx
	or cancer of the lung.

See [ppp.object](#) for details of the format.

The dataset chorley.extra is a list with two components. The first component plotit is a function which will plot the data in a sensible fashion. The second component incin is a list with entries x and y giving the location of the industrial incinerator.

Coordinates are given in kilometres, and the resolution is 100 metres (0.1 km)

Notes

The data give the precise domicile addresses of new cases of cancer of the larynx (58 cases) and cancer of the lung (978 cases), recorded in the Chorley and South Ribble Health Authority of Lancashire (England) between 1974 and 1983. The supplementary data give the location of a disused industrial incinerator.

The data were first presented and analysed by Diggle (1990). They have subsequently been analysed by Diggle and Rowlingson (1994) and Baddeley et al. (2005).

The aim is to assess evidence for an increase in the incidence of cancer of the larynx in the vicinity of the now-disused industrial incinerator. The lung cancer cases serve as a surrogate for the spatially-varying density of the susceptible population.

The data are represented as a marked point pattern, with the points giving the spatial location of each individual's home address and the marks identifying whether each point is a case of laryngeal cancer or lung cancer.

Coordinates are in kilometres, and the resolution is 100 metres (0.1 km).

The dataset chorley has a polygonal window with 132 edges which closely approximates the boundary of the Chorley and South Ribble Health Authority.

Note that, due to the rounding of spatial coordinates, the data contain duplicated points (two points at the same location). To determine which points are duplicates, use `duplicated.ppp`. To remove the duplication, use `unique.ppp`.

Source

Coordinates of cases were provided by the Chorley and South Ribble Health Authority, and were kindly supplied by Professor Peter Diggle. Region boundary was digitised by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>, 2005, from a photograph of an Ordnance Survey map.

References

Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005) Residual analysis for spatial point processes. *Journal of the Royal Statistical Society, Series B* **67**, 617–666.

Diggle, P. (1990) A point process modelling approach to raised incidence of a rare phenomenon in the vicinity of a prespecified point. *Journal of the Royal Statistical Soc. Series A* **153**, 349-362.

Diggle, P. and Rowlingson, B. (1994) A conditional approach to point process modelling of elevated risk. *Journal of the Royal Statistical Soc. Series A* **157**, 433-440.

Examples

```
    chorley
  if(require(spatstat.geom)) {
    summary(chorley)
    chorley.extra$plotit()
  }
```

---

clmfires   *Castilla-La Mancha Forest Fires*

Description

This dataset is a record of forest fires in the Castilla-La Mancha region of Spain between 1998 and 2007. This region is approximately 400 by 400 kilometres. The coordinates are recorded in kilometres.

The dataset `clmfires` is a point pattern (object of class `"ppp"`) containing the spatial coordinates of each fire, with marks containing information about each fire. There are 4 columns of marks:

cause	cause of fire (see below)
burnt.area	total area burned, in hectares
date	the date of fire, as a value of class Date
julian.date	number of days elapsed since 1 January 1998

The cause of the fire is a factor with the levels `lightning`, `accident` (for accidents or negligence), `intentional` (for intentionally started fires) and `other` (for other causes including unknown cause).

The format of date is "Year-month-day", e.g. "2005-07-14" means 14 July, 2005.

The accompanying dataset clmfires.extra is a list of two items clmcov100 and clmcov200 containing covariate information for the entire Castilla-La Mancha region. Each of these two elements is a list of four images (objects of class "im") named elevation, orientation, slope and landuse. The landuse image is factor-valued with the factor having levels urban, farm (for farms or orchards), meadow, denseforest (for dense forest), conifer (for conifer forest or plantation), mixedforest, grassland, bush, scrub and artifgreen for artificial greens such as golf courses.

These images (effectively) provide values for the four covariates at every location in the study area. The images in clmcov100 are 100 by 100 pixels in size, while those in clmcov200 are 200 by 200 pixels. For easy handling, clmcov100 and clmcov200 also belong to the class "listof" so that they can be plotted and printed immediately.

Usage

```
data(clmfires)
```

Format

clmfires is a marked point pattern (object of class "ppp"). See ppp.object.

clmfires.extra is a list with two components, named clmcov100 and clmcov200, which are lists of pixel images (objects of class "im").

Remark

The precision with which the coordinates of the locations of the fires changed between 2003 and 2004. From 1998 to 2003 many of the locations were recorded as the centroid of the corresponding "district unit"; the rest were recorded as exact UTM coordinates of the centroids of the fires. In 2004 the system changed and the exact UTM coordinates of the centroids of the fires were used for *all* fires. There is thus a strongly apparent "gridlike" quality to the fire locations for the years 1998 to 2003.

There is however no actual duplication of points in the 1998 to 2003 patterns due to "jittering" having been applied in order to avoid such duplication. It is not clear just *how* the fire locations were jittered. It seems unlikely that the jittering was done using the jitter() function from R or the **spatstat** function rjitter.

Of course there are many sets of points which are *virtually* identical, being separated by distances induced by the jittering. Typically these distances are of the order of 40 metres which is unlikely to be meaningful on the scale at which forest fires are observed.

Caution should therefore be exercised in any analyses of the patterns for the years 1998 to 2003.

Source

Professor Jorge Mateu.

Examples

```
  if(require(spatstat.geom)) {
plot(clmfires, which.marks="cause", cols=2:5, cex=0.25)
plot(clmfires.extra$clmcov100)
```

```
# Split the clmfires pattern by year and plot the first and last years:
yr  <- factor(format(marks(clmfires)$date,format="%Y"))
X   <- split(clmfires,f=yr)
fAl <- c("1998","2007")
plot(X[fAl],use.marks=FALSE,main.panel=fAl,main="")
  }
```

---

concrete	*Air Bubbles in Concrete*

Description

Prof. Shin-ichi Igarashi's data: a point pattern of the locations, in a cross-section of a concrete body, of the centroids of air bubbles in the cement paste matrix surrounding particles of aggregate.

Usage

```
data("concrete")
```

Format

An object of class "ppp" representing the point pattern of air bubble centroid locations. Spatial coordinates are expressed in microns.

Details

The window of the point pattern is a binary mask (window of type "mask"; see owin and as.mask for more information about this type of window). This window in effect consists of the cement paste matrix, or equivalently of the complement (in the observed cross-section) of the aggregate.

Major scientific interest is focussed on analysing the distribution of the location of the air bubbles in the cement paste matrix. These bubbles are important in assuring frost resistance of the concrete. Each air bubble protects a region around it to a certain distance. To protect an entire concrete object against severe frost attack, it is necessary to cover the whole of the cement paste matrix with subsets of protected regions formed around the air bubbles. It is believed that the protected regions are related to the Dirichlet tessellation of the centroids of the bubbles, and the statistical properties of the protected regions can be determined from those of the Dirichlet tessellation. In this regard, the areas of the tiles are particularly important.

Source

Prof. Shin-ichi Igarashi, of the School of Geoscience and Civil Engineering, Kanazawa University, personal communication.

References

Natesaiyer, K., Hover, K.C. and Snyder, K.A. (1992). Protected-paste volume of air-entrained cement paste: part 1. *Journal of Materials in Civil Engineering* **4** No.2, 166 – 184.

Murotani, T., Igarashi, S. and Koto, H. (2019). Distribution analysis and modeling of air voids in concrete as spatial point processes. *Cement and Concrete Research* **115** 124 – 132.

Examples

```
if(require(spatstat.geom)) {
   plot(concrete,chars="+",cols="blue",col="yellow")
   # The aggregate is in yellow; the cement paste matrix is in white.

   # Unit of length: use \mu symbol for micron
   unitname(concrete) <- "\u00B5m"

   if(interactive()) {
     # Compute the Dirichlet tessellation
     dtc <- dirichlet(concrete)
     plot(dtc,ribbon=FALSE, col=sample(rainbow(dtc$n)))
     # Study Dirichlet tile areas
     areas <- tile.areas(dtc)
     aa <- areas/1000 # Divide by 1000 to avoid numerical instability
     # Fit a gamma distribution by the method of moments
     mm <- mean(aa)
     vv <- var(aa)
     shape <- mm^2/vv
     rate <- mm/vv
     rate <- rate/1000 # Adjust for rescaling
     hist(areas,probability=TRUE,ylim=c(0,7.5e-6),
        main="Histogram and density estimates for areas",ylab="",xlab="area")
     lines(density(areas),col="red")
     curve(dgamma(x,shape=shape,rate=rate),add=TRUE,col="blue")
     legend("topright",lty=1,col=c("red","blue"),
            legend=c("non-parametric","gamma fit"),bty="n")
   }
}
```

---

copper   *Berman-Huntington points and lines data*

Description

These data come from an intensive geological survey of a 70 x 158 km region in central Queensland, Australia. They consist of 67 points representing copper ore deposits, and 146 line segments representing geological 'lineaments'. Lineaments are linear features, visible on a satellite image, that are believed to consist largely of geological faults (Berman, 1986, p. 55). It would be of great interest to predict the occurrence of copper deposits from the lineament pattern, since the latter can easily be observed on satellite images.

These data were introduced and analysed by Berman (1986). They have also been studied by Berman and Diggle (1989), Berman and Turner (1992), Baddeley and Turner (2000, 2005), Foxall and Baddeley (2002) and Baddeley et al (2005).

Many analyses have been performed on the southern half of the data only. This subset is also provided.

Usage

```
data(copper)
```

Format

copper is a list with the following entries:

Points a point pattern (object of class ″ppp″) representing the full point pattern of copper deposits. See `ppp.object` for details of the format.

Lines a line segment pattern (object of class ″psp″) representing the lineaments in the full dataset. See `psp.object` for details of the format.

SouthWindow the window delineating the southern half of the study region. An object of class ″owin″.

SouthPoints the point pattern of copper deposits in the southern half of the study region. An object of class ″ppp″.

SouthLines the line segment pattern of the lineaments in the southern half of the study region. An object of class ″psp″.

All spatial coordinates are expressed in kilometres.

Source

Dr Jonathan Huntington, CSIRO Earth Science and Resource Engineering, Sydney, Australia. Co-ordinates kindly provided by Dr. Mark Berman and Dr. Andy Green, CSIRO, Sydney, Australia.

References

Baddeley, A. and Turner, R. (2000) Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics* **42**, 283–322.

Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005) Residual analysis for spatial point processes. *Journal of the Royal Statistical Society, Series B* **67**, 617–666.

Baddeley, A. and Turner, R. (2005) Modelling spatial point patterns in R. In: A. Baddeley, P. Gregori, J. Mateu, R. Stoica, and D. Stoyan, editors, *Case Studies in Spatial Point Pattern Modelling*, Lecture Notes in Statistics number 185. Pages 23–74. Springer-Verlag, New York, 2006. ISBN: 0-387-28311-0.

Berman, M. (1986). Testing for spatial association between a point process and another stochastic process. *Applied Statistics* **35**, 54–62.

Berman, M. and Diggle, P.J. (1989) Estimating Weighted Integrals of the Second-order Intensity of a Spatial Point Process. *Journal of the Royal Statistical Society, series B* **51**, 81–92.

Berman, M. and Turner, T.R. (1992) Approximating point process likelihoods with GLIM. *Applied Statistics* **41**, 31–38.

Foxall, R. and Baddeley, A. (2002) Nonparametric measures of association between a spatial point process and a random set, with geological applications. *Applied Statistics* **51**, 165–182.

Examples

```
data(copper)

if(require(spatstat.model)) {

# Plot full dataset

plot(copper$Points)
plot(copper$Lines, add=TRUE)

# Plot southern half of data
plot(copper$SouthPoints)
plot(copper$SouthLines, add=TRUE)

if(interactive()) {
  Z <- distmap(copper$SouthLines)
  plot(Z)
  X <- copper$SouthPoints
  ppm(X, ~D, covariates=list(D=Z))
}
}
```

---

copyExampleFiles   *Copy Data Files for Example*

Description

This command copies several data files to a folder (directory) chosen by the user, so that they can be used for a practice example.

Usage

```
copyExampleFiles(which, folder = getwd())
```

Arguments

which   Character string name (partially matched) of one of the datasets installed in
             spatstat for which the original data files are provided. If which is missing,
             a list of available options is printed.

folder   Character string path name of a folder (directory) in which the files will be
             placed. Defaults to the current working directory.

Details

The original text files containing data for the selected dataset are copied to the chosen folder.

This is part of an exercise described in Chapter 3 of Baddeley, Rubak and Turner (2015).

Author(s)

Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>, Rolf Turner <r.turner@auckland.ac.nz> and Ege Rubak <rubak@math.aau.dk>.

References

Baddeley, A., Rubak, E. and Turner, R. (2015) *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press.

See Also

[vesicles](), [finpines]()

Examples

```
copyExampleFiles()
```

---

demohyper   *Demonstration Example of Hyperframe of Spatial Data*

Description

This is an artificially constructed example of a hyperframe of spatial data. The data could have been obtained from an experiment in which there are two groups of experimental units, the response from each unit is a point pattern Points, and for each unit there is explanatory data in the form of a pixel image Image.

Usage

```
data(demohyper)
```

Format

A [hyperframe]() with 3 rows and 3 columns:

Points List of spatial point patterns (objects of class "ppp") serving as the responses in an experiment.

Image List of images (objects of class "im") serving as explanatory variables.

Group Factor with two levels a and b serving as an explanatory variable.

Source

Artificially generated by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>.

Examples

```
 if(require(spatstat.model)) {
plot(demohyper, quote({ plot(Image, main=""); plot(Points, add=TRUE) }),
     parargs=list(mar=rep(1,4)))
mppm(Points ~ Group/Image, data=demohyper)
 }
```

---

demopat   *Artificial Data Point Pattern*

Description

This is an artificial dataset, for use in testing and demonstrating the capabilities of the spatstat package.

It is a multitype point pattern in an irregular polygonal window. There are two types of points. The window contains a polygonal hole. Spatial coordinates are expressed in *furlongs* (one furlong equals 660 feet).

Usage

```
data(demopat)
```

Format

An object of class "ppp" representing the point pattern.

See [ppp.object](#) for details of the format of a point pattern object.

Source

Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>

---

dendrite   *Dendritic Spines Data*

Description

Dendrites are branching filaments which extend from the main body of a neuron (nerve cell) to propagate electrochemical signals. Spines are small protrusions on the dendrites.

This dataset gives the locations of 566 spines observed on one branch of the dendritic tree of a rat neuron. The spines are classified according to their shape into three types: mushroom, stubby or thin.

The data have been analysed in Jammalamadaka et al (2013) and Baddeley et al (2014). Please cite these papers and acknowledge the Kosik Lab, UC Santa Barbara, in any use of the data.

Usage

```
data("dendrite")
```

Format

Object of class `"lpp"`. See `lpp`.

Spatial coordinates are expressed in microns.

Source

Kosik Lab, UC Santa Barbara (Dr Kenneth Kosik, Dr Sourav Banerjee). Formatted for `spatstat` by Dr Aruna Jammalamadaka.

References

Baddeley, A, Jammalamadaka, A. and Nair, G. (2014) Multitype point process analysis of spines on the dendrite network of a neuron. *Applied Statistics* (Journal of the Royal Statistical Society, Series C), **63**, 673–694.

Jammalamadaka, A., Banerjee, S., Manjunath, B.S. and Kosik, K. (2013) Statistical Analysis of Dendritic Spine Distributions in Rat Hippocampal Cultures. *BMC Bioinformatics* **14**, 287.

Examples

```
  if(require(spatstat.linnet)) {
plot(dendrite,leg.side="bottom", main="", cex=0.75, cols=2:4)
  }
```

---

finpines   *Pine saplings in Finland.*

Description

The data record the locations of 126 pine saplings in a Finnish forest, their heights and their diameters.

The dataset `finpines` is a marked point pattern containing the locations of the saplings marked by their heights and their diameters.

Sapling locations are given in metres (to six significant digits); heights are in metres (rounded to the nearest 0.1 metre, except in one case to the nearest 0.05 metres); diameters are in centimetres (rounded to the nearest centimetre).

The data were recorded by Professor Seppo Kellomaki, Faculty of Forestry, University of Joensuu, Finland, and subsequently massaged by Professor Antti Penttinen, Department of Statistics, University of Jyv\"askyl\"a, Finland.

Originally the point locations were observed in polar coordinates with rather poor angular precision. Hence the coordinates are imprecise for large radius because of rounding errors: indeed the alignments can be observed by eye.

The data were manipulated by Prof Penttinen by making small angular perturbations at random. After this transformation, the original data (in a circular plot) were clipped to a square window, for convenience.

Professor Penttinen emphasises that the data were intended only for initial experimentation. They have some strange features. For example, if the height is less than 1.3 metres then the diameter can be uncertain. Also there are some very close pairs of points. Some pairs of trees (namely (58,59), (78,79), (96,97) and (102,103)) violate the requirement that the interpoint distance should be greater than half the sum of their diameters.

These data have subsequently been analysed by Van Lieshout (2004).

Usage

```
data(finpines)
```

Format

Object of class ″ppp″ representing the point pattern of sapling locations marked by their heights and diameters. See `ppp.object` for details of the format.

Source

Prof Antti Penttinen

References

Van Lieshout, M.N.M. (2004) A J-function for marked point patterns. Research Report PNA-R0404, June 2004. Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 2004.

Examples

```
   data(finpines)
 if(require(spatstat.geom)) {
   plot(unmark(finpines), main="Finnish pines: locations")
   plot(finpines, which.marks="height", main="heights")
   plot(finpines, which.marks="diameter", main="diameters")
   plot(finpines, which.marks="diameter",
           main="diameters to scale", markscale=1/200)
 }
```

---

flu   *Influenza Virus Proteins*

Description

Replicated spatial point patterns giving the locations of two different virus proteins on the membranes of cells infected with influenza virus.

Usage

```
data(flu)
```

Format

A `hyperframe` with 41 rows and four columns:

pattern List of spatial point patterns (objects of class ″ppp″) with points of two types, identifying the locations of two different proteins on a membrane sheet. Coordinates are expressed in nanometres (*nm*) and the window of observation is a square of side length 3331 *nm*.

virustype Factor identifying whether the infecting virus was the wild type (`wt`) or mutant (`mut1`).

stain Factor identifying whether the membrane sheet was stained for the proteins *M2* and *M1* (`stain="M2-M1"`) or stained for the proteins *M2* and *HA* (`stain="M2-HA"`).

frameid Integer. Serial number of the microscope frame in the original experiment. Frame identifier is not unique across different values of `virustype` and `stain`.

The row names of the hyperframe can be used as succinct labels in plots.

Details

The data consist of 41 spatial point patterns, each giving the locations of two different virus proteins on the membranes of cells infected with influenza virus.

Chen et al (2008) conducted the experiment and used spatial analysis to establish evidence for an interaction between the influenza virus proteins M1 and M2 that is important for the study of viral replication.

Canine kidney cells were infected with human influenza, Udorn strain, either the wild type or a mutant which encodes a defective M2 protein. At twelve hours post-infection, membrane sheets were prepared and stained for viral proteins, using two antibodies conjugated to gold particles of two sizes (6 nanometre and 12 nanometre diameter) enabling localisation of two different proteins on each sheet. The 6 nm particles were stained for M2 (ion channel protein), while the 12 nm particles were stained either for M1 (matrix protein) or for HA (hemagglutinin). Membrane sheets were visualised in electron microscopy.

Experimental technique and spatial analysis of the membranes stained for M2 and M1 is reported in Chen et al (2008). Analysis of the membranes stained for M2 and HA is reported in Rossman et al (2010). The M2-HA data shows a stronger association between the two proteins which has also been observed biochemically and functionally (Rossman et al, 2010).

The dataset `flu` is a `hyperframe` with one row for each membrane sheet. The column named `pattern` contains the spatial point patterns of gold particle locations, with two types of points (either M1 and M2 or HA and M2). The column named `virustype` is a factor identifying the virus: either wild type `wt` or mutant `mut1`. The column named `stain` is a factor identifying whether the membrane was stained for M1 and M2 (`stain="M2-M1"`) or stained for HA and M2 (`stain="M2-HA"`). The row names of the hyperframe are a succinct summary of the experimental conditions and can be used as labels in plots. See the Examples.

Source

Data generously provided by Dr G.P. Leser and Dr R.A. Lamb. Please cite Chen et al (2008) in any use of these data.

References

Chen, B.J., Leser, G.P., Jackson, D. and Lamb, R.A. (2008) The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding. *Journal of Virology* **82**, 10059–10070.

Rossman, J.S., Jing, X.H., Leser, G.P. and Lamb, R.A. (2010) Influenza virus M2 protein mediates ESCRT-independent membrane scission *Cell* **142**, 902–913.

Examples

```
  if(require(spatstat.geom)) {
flu
Y <- flu$pattern[10]
Y <- flu[10, 1, drop=TRUE]
wildM1 <- with(flu, virustype == "wt" & stain == "M2-M1")
plot(flu[wildM1, 1, drop=TRUE],
     main=c("flu data", "wild type virus, M2-M1 stain"),
     pch=c(3,16), cex=0.4, cols=2:3)
  }
```

---

ganglia	*Beta Ganglion Cells in Cat Retina, Old Version*

Description

Point pattern of retinal ganglion cells identified as 'on' or 'off'. A marked point pattern.

Usage

```
data(ganglia)
```

Format

An object of class ″ppp″ representing the point pattern of cell locations. Entries include

x	Cartesian x-coordinate of cell
y	Cartesian y-coordinate of cell
marks	factor with levels off and on
	indicating "off" and "on" cells

See `ppp.object` for details of the format.

Notes

Important: these data are INCORRECT. See below.

The data represent a pattern of beta-type ganglion cells in the retina of a cat recorded in Figure 6(a) of W\"assle et al. (1981).

The pattern was first analysed by W\"assle et al (1981) using nearest neighbour distances. The data used in their analysis are not available.

The present dataset `ganglia` was scanned from Figure 6(a) of W\"assle et al (1981) in the early 1990's, but we have no further information. This dataset is the one analysed by Van Lieshout and Baddeley (1999) using multitype J functions, and by Stoyan (1995) using second order methods (pair correlation and mark correlation).

It has now been discovered that these data are **incorrect**. They are not faithful to the scale in Figure 6 of W\"assle et al (1981), and they contain some scanning errors. Hence they should not be used to address the original scientific question. They have been retained only for comparison with other analyses in the statistical literature.

A new, corrected dataset, scanned from the original microscope image, has been provided under the name `betacells`. Use that dataset for any further study.

Warnings

These data are incorrect. Use the new corrected dataset `betacells`.

Source

W\"assle et al (1981), data supplied by Marie-Colette van Lieshout and attributed to Peter Diggle

References

Stoyan, D. (1995) Personal communication.

Van Lieshout, M.N.M. and Baddeley, A.J. (1999) Indices of dependence between types in multi-variate point patterns. *Scandinavian Journal of Statistics* **26**, 511–532.

W\"assle, H., Boycott, B. B. & Illing, R.-B. (1981). Morphology and mosaic of on- and off-beta cells in the cat retina and some functional considerations. *Proc. Roy. Soc. London Ser. B* **212**, 177–195.

---

gordon   *People in Gordon Square*

Description

This dataset records the location of people sitting on a grass patch in Gordon Square, London, at 3pm on a sunny afternoon.

The dataset gordon is a point pattern (object of class ″ppp″) containing the spatial coordinates of each person.

The grass patch is an irregular polygon with two holes.

Coordinates are given in metres.

Usage

```
data(gordon)
```

Source

Andrew Bevan, University College London.

References

Baddeley, A., Turner, R., Mateu, J. and Bevan, A. (2013) Hybrids of Gibbs point process models and their implementation. *Journal of Statistical Software* **55**:11, 1–43. DOI: `10.18637/jss.v055.i11`

Examples

```
data(gordon)
  if(require(spatstat.geom)) {
plot(gordon)
  }
```

---

gorillas   *Gorilla Nesting Sites*

Description

Locations of nesting sites of gorillas, and associated covariates, in a National Park in Cameroon.

Usage

```
data(gorillas)
```

Format

gorillas is a marked point pattern (object of class ″ppp″) representing nest site locations.

gorillas.extra is a named list of 7 pixel images (objects of class ″im″) containing spatial covariates. It also belongs to the class ″listof″.

All spatial coordinates are in metres. The coordinate reference system is WGS_84_UTM_Zone_32N.

Details

These data come from a study of gorillas in the Kagwene Gorilla Sanctuary, Cameroon, by the Wildlife Conservation Society Takamanda-Mone Landscape Project (WCS-TMLP). A detailed description and analysis of the data is reported in Funwi-Gabga and Mateu (2012).

The dataset gorillas is a marked point pattern (object of class ″ppp″) giving the spatial locations of 647 nesting sites of gorilla groups observed in the sanctuary over time. Locations are given as UTM (Zone 32N) coordinates in metres. The observation window is the boundary of the sanctuary, represented as a polygon. Marks attached to the points are:

group Identifier of the gorilla group that constructed the nest site: a categorical variable with values major or minor.

season Season in which data were collected: categorical, either rainy or dry.

date Day of observation. A value of class `"Date"`.

Note that the data contain duplicated points (two points at the same location). To determine which points are duplicates, use `duplicated.ppp`. To remove the duplication, use `unique.ppp`.

The accompanying dataset `gorillas.extra` contains spatial covariate information. It is a named list containing seven pixel images (objects of class `"im"`) giving the values of seven covariates over the study region. It also belongs to the class `"listof"` so that it can be plotted. The component images are:

aspect Compass direction of the terrain slope. Categorical, with levels `N`, `NE`, `E`, `SE`, `S`, `SW`, `W` and `NW`.

elevation Digital elevation of terrain, in metres.

heat Heat Load Index at each point on the surface (Beer's aspect), discretised. Categorical with values `Warmest` (Beer's aspect between 0 and 0.999), `Moderate` (Beer's aspect between 1 and 1.999), `Coolest` (Beer's aspect equals 2).

slopeangle Terrain slope, in degrees.

slopetype Type of slope. Categorical, with values `Valley`, `Toe` (toe slope), `Flat`, `Midslope`, `Upper` and `Ridge`.

vegetation Vegetation or cover type. Categorical, with values `Disturbed` (highly disturbed forest), `Colonising` (colonising forest), `Grassland` (savannah), `Primary` (primary forest), `Secondary` (secondary forest), and `Transition` (transitional vegetation).

waterdist Euclidean distance from nearest water body, in metres.

For further information see Funwi-Gabga and Mateu (2012).

Raw Data

For demonstration and training purposes, the raw data file for the `vegetation` covariate is also provided in the **spatstat.data** package installation, as the file `vegetation.asc` in the folder `rawdata/gorillas`. Use `system.file` to obtain the file path: `system.file("rawdata/gorillas/vegetation.asc", package="spatstat.data")`. This is a text file in the simple ASCII file format of the geospatial library GDAL. The file can be read by the function `readGDAL` in the **rgdal** package, or alternatively read directly using `scan`.

Source

Field data collector: Wildlife Conservation Society Takamanda-Mone Landscape Project (WCS-TMLP). *Please acknowledge WCS-TMLP in any use of these data.*

Data kindly provided by Funwi-Gabga Neba, Data Coordinator of A.P.E.S. Database Project, Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

The collaboration of Prof Jorge Mateu, Universitat Jaume I, Castellon, Spain is gratefully acknowledged.

References

Funwi-Gabga, N. (2008) *A pastoralist survey and fire impact assessment in the Kagwene Gorilla Sanctuary, Cameroon*. M.Sc. thesis, Geology and Environmental Science, University of Buea, Cameroon.

Funwi-Gabga, N. and Mateu, J. (2012) Understanding the nesting spatial behaviour of gorillas in the Kagwene Sanctuary, Cameroon. *Stochastic Environmental Research and Risk Assessment* **26** (6), 793–811.

Examples

```
if(require(spatstat.geom)) {
summary(gorillas)
plot(gorillas)
plot(gorillas.extra)
}
```

---

hamster   *Aherne's hamster tumour data*

Description

Point pattern of cell nuclei in hamster kidney, each nucleus classified as either 'dividing' or 'pyknotic'. A multitype point pattern.

Usage

```
data(hamster)
```

Format

An object of class ″ppp″ representing the point pattern of cell locations. Entries include

x	Cartesian x-coordinate of cell
y	Cartesian y-coordinate of cell
marks	factor with levels ″dividing″ and ″pyknotic″.

See `ppp.object` for details of the format.

Notes

These data were presented and analysed by Diggle (1983, section 7.3).

The data give the positions of the centres of the nuclei of certain cells in a histological section of tissue from a laboratory-induced metastasising lymphoma in the kidney of a hamster.

The nuclei are classified as either "pyknotic" (corresponding to dying cells) or "dividing" (corresponding to cells arrested in metaphase, i.e. in the act of dividing). The background void is occupied by unrecorded, interphase cells in relatively large numbers.

The sampling window is a square, originally about 0.25 mm square in real units, which has been rescaled to the unit square.

Source

Dr W. A. Aherne, Department of Pathology, University of Newcastle-upon-Tyne, UK. Data supplied by Prof. Peter Diggle

References

Diggle, P.J. (1983) *Statistical analysis of spatial point patterns*. Academic Press.

Examples

```
if(require(spatstat.geom)) {
hamster
## rescale to microns
(Ham <- rescale(hamster))
}
```

---

heather *Diggle's Heather Data*

Description

The spatial mosaic of vegetation of the heather plant (*Calluna vulgaris*) recorded in a 10 by 20 metre sampling plot in Sweden.

Usage

```
data(heather)
```

Format

A list with three entries, representing the same data at different spatial resolutions:

coarse	original heather data, 100 by 200 pixels
medium	current heather data, 256 by 512 pixels
fine	finest resolution data, 778 by 1570 pixels

Each of these entries is an object of class `"owin"` containing a binary pixel mask. All spatial coordinates are given in metres.

Notes on data

These data record the spatial mosaic of vegetation of the heather plant (*Calluna vulgaris*) in a 10 by 20 metre sampling plot near Jädraås, Sweden. They were recorded and first analysed by Diggle(1981).

The dataset `heather` contains three different versions of the data that have been analysed by different writers over the decades.

coarse: Data as originally digitised by Diggle in 1983 at 100 by 200 pixels resolution (i.e. 10 pixels = 1 metre).

    These data were entered by hand in the form of a run-length encoding (original file no longer available) and translated by a program into a 100 by 200 pixel binary image.

    There are known to be some errors in the image which arise from errors in counting the run-length so that occasionally there will be an unexpected 'spike' on one single column.

fine: A fine scale digitisation of the original map, prepared by CWI (Centre for Computer Science, Amsterdam, Netherlands) in 1994.

    The original hand-drawn map was scanned by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>, and processed by Chris Jonker, Henk Heijmans and Adrian Baddeley <Adrian.Baddeley@curtin.edu.au> to yield a clean binary image of 778 by 1570 pixels resolution.

medium: The version of the heather data currently supplied on Professor Diggle's website. This is a 256 by 512 pixel image. The method used to create this image is not stated.

History of analysis of data

The data were recorded, presented and analysed by Diggle (1983). He proposed a Boolean model consisting of discs of random size with centres generated by of a Poisson point process.

Renshaw and Ford (1983) reported that spectral analysis of the data suggested the presence of strong row and column effects. However, this may have been attributable to errors in the run-length encoding of the original data.

Hall (1985) and Hall (1988, pp 301-318) took a bootstrap approach.

Ripley (1988, pp. 121-122, 131-135] used opening and closing functions to argue that a Boolean model of discs is inappropriate.

Cressie (1991, pp. 763-770) tried a more general Boolean model.

Source

Peter Diggle

References

Cressie, N.A.C. (1991) *Statistics for Spatial Data*. John Wiley and Sons, New York.

Diggle, P.J. (1981) Binary mosaics and the spatial pattern of heather. *Biometrics* **37**, 531-539.

Hall, P. (1985) Resampling a coverage pattern. *Stochastic Processes and their Applications* **20** 231-246.

Hall, P. (1988) *An introduction to the theory of coverage processes*. John Wiley and Sons, New York.

Renshaw, E. and Ford, E.D. (1983) The interpretation of process from pattern using two-dimensional spectral analysis: Methods and problems of interpretation. *Applied Statistics* **32** 51-63.

Ripley, B.D. (1988) *Statistical Inference for Spatial Processes*. Cambridge University Press.

---

humberside	*Humberside Data on Childhood Leukaemia and Lymphoma*

Description

Spatial locations of cases of childhood leukaemia and lymphoma, and randomly-selected controls, in North Humberside. A marked point pattern.

Usage

```
data(humberside)
```

Format

The dataset humberside is an object of class ″ppp″ representing a marked point pattern. Entries include

x	Cartesian x-coordinate of home address
y	Cartesian y-coordinate of home address
marks	factor with levels case and control
	indicating whether this is a disease case
	or a control.

See [ppp.object](#) for details of the format.

Spatial coordinates are expressed as multiples of 100 metres.

The dataset humberside.convex is an object of the same format, representing the same point pattern data, but contained in a larger, 5-sided convex polygon.

Notes

Cuzick and Edwards (1990) first presented and analysed these data.

The data record 62 cases of childhood leukaemia and lymphoma diagnosed in the North Humberside region of England between 1974 and 1986, together with 141 controls selected at random from the birth register for the same period.

The data are represented as a marked point pattern, with the points giving the spatial location of each individual's home address (actually, the centroid for the postal code) and the marks identifying cases and controls.

Coordinates are expressed in units of 100 metres, and the resolution is 100 metres. At this resolution, there are some duplicated points. To determine which points are duplicates, use [duplicated.ppp](#). To remove the duplication, use [unique.ppp](#).

Two versions of the dataset are supplied, both containing the same point coordinates, but using different windows. The dataset humberside has a polygonal window with 102 edges which closely approximates the Humberside region, while humberside.convex has a convex 5-sided polygonal window originally used by Diggle and Chetwynd (1991) and shown in Figure 1 of that paper. (This pentagon has been modified slightly from the original data, by shifting two vertices horizontally by 1 unit, so that the pentagon contains all the data points.)

Source

Dr Ray Cartwright and Dr Freda Alexander. Published and analysed in Cuzick and Edwards (1990), see Table 1. Pentagonal boundary from Diggle and Chetwynd (1991), Figure 1. Point coordinates and pentagonal boundary supplied by Andrew Lawson. Detailed region boundary was digitised by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>, 2005, from a reprint of Cuzick and Edwards (1990).

References

J. Cuzick and R. Edwards (1990) Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society, series B*, **52** (1990) 73-104.

P.J. Diggle and A.G. Chetwynd (1991) Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics* 47 (1991) 1155-1163.

Examples

```
if(require(spatstat.geom)) {
 humberside
 summary(humberside)
 plot(humberside)
 plot(Window(humberside.convex), add=TRUE, lty=2)
 ## convert to metres
 (Hum <- rescale(humberside))
 ## convert to kilometres
 (HumK <- rescale(humberside, 10, "km"))
}
```

---

hyytiala   *Scots pines and other trees at Hyytiala*

Description

This dataset is a spatial point pattern of trees recorded at Hyytiälä, Finland. The majority of the trees are Scots pines. See Kokkila et al (2002).

The dataset hyytiala is a point pattern (object of class "ppp") containing the spatial coordinates of each tree, marked by species (a factor with levels aspen, birch, pine and rowan). The survey region is a 20 by 20 metre square. Coordinates are given in metres.

Usage

```
data(hyytiala)
```

Source

Nicolas Picard

References

Kokkila, T., Mäkelä, A. and Nikinmaa E. (2002) A method for generating stand structures using Gibbs marked point process. *Silva Fennica* **36** 265–277.

Picard, N, Bar-Hen, A., Mortier, F. and Chadoeuf, J. (2009) The multi-scale marked area-interaction point process: a model for the spatial pattern of trees. *Scandinavian Journal of Statistics* **36** 23–41

Examples

```
data(hyytiala)
  if(require(spatstat.geom)) {
plot(hyytiala, cols=2:5)
  }
```

---

japanesepines   *Japanese Pines Point Pattern*

Description

The data give the locations of saplings of Japanese black pine (*Pinus thunbergii*) in a square sampling region in a natural forest. The observations were originally collected by Numata (1961).

These data are used as a standard example in the textbook of Diggle (2003); see pages 1, 14, 19, 22, 24, 56–57 and 61.

Usage

```
data(japanesepines)
```

Format

An object of class ″ppp″ representing the point pattern of 65 tree sapling locations in a 5.7 x 5.7 metre square, rescaled to the unit square and rounded to two decimal places.

See `ppp.object` for details of the format of a point pattern object.

Source

Diggle (2003), obtained from Numata (1961)

References

Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns.* Arnold Publishers.

Numata, M. (1961) Forest vegetation in the vicinity of Choshi. Coastal flora and vegetation at Choshi, Chiba Prefecture. IV. *Bulletin of Choshi Marine Laboratory, Chiba University* **3**, 28–48 (in Japanese).

Examples

```
if(require(spatstat.geom)) {
 japanesepines
 summary(japanesepines)
 ## rescale to metres
 (Jpines <- rescale(japanesepines))
}
```

Kovesi   *Colour Sequences with Uniform Perceptual Contrast*

Description

A collection of 41 different sequences of colours, each sequence having a uniform perceptual contrast over its whole range. These sequences make very good colour maps which avoid introducing artefacts when displaying image data.

Usage

```
data(Kovesi)
```

Format

A [hyperframe](#) with the following columns:

linear	Logical: whether the sequence is linear.
diverging	Logical: whether the sequence is diverging.
rainbow	Logical: whether the sequence is a rainbow.
cyclic	Logical: whether the sequence is cyclic.
isoluminant	Logical: whether the sequence is isoluminant.
ternary	Logical: whether the sequence is ternary.
colsig	Character: colour signature (see Details)
l1, l2	Numeric: lightness parameters
chro	Numeric: average chroma (percent)
n	Numeric: length of colour sequence
cycsh	Numeric: cyclic shift (percent)
values	: Character: the colour values.

Details

Kovesi (2014, 2015) presented a collection of colour sequences that have uniform perceptual contrast over their whole range.

The dataset `Kovesi` provides these data. It is a `hyperframe` with 41 rows, in which each row provides information about one colour sequence.

Additional information in each row specifies whether the colour sequence is 'linear', 'diverging', 'rainbow', 'cyclic', 'isoluminant' and/or 'ternary' as defined by Kovesi (2014, 2015).

The 'colour signature' is a string composed of letters representing the successive hues, using the following code:

r	red
g	green
b	blue
c	cyan
m	magenta
y	yellow
o	orange
v	violet
k	black
w	white
j	grey (j rhymes with grey)

For example `kryw` is the sequence from black to red to yellow to white.

The column `values` contains the colour data themselves. The `i`th colour sequence is `Kovesi$values[[i]]`, a character vector of length 256.

Source

Dr Peter Kovesi, Centre for Exploration Targeting, University of Western Australia.

References

Kovesi, P. (2014) Website *CET Uniform Perceptual Contrast Colour Maps* https://www.peterkovesi.com/projects/colourmaps/

Kovesi, P. (2015) Good Colour Maps: How to Design Them. `arXiv:1509.03700 [cs.GR]`

Examples

```
Kovesi
LinearBMW <- Kovesi$values[[28]]
if(require(spatstat.geom)) {
plot(colourmap(LinearBMW, range=c(0,1)))

## The following would be suitable for spatstat.options(image.colfun)
BMWfun <- function(n) { interp.colours(LinearBMW, n) }
}
```

---

lansing   *Lansing Woods Point Pattern*

Description

Locations and botanical classification of trees in Lansing Woods.

The data come from an investigation of a 924 ft x 924 ft (19.6 acre) plot in Lansing Woods, Clinton County, Michigan USA by D.J. Gerrard. The data give the locations of 2251 trees and their botanical classification (into hickories, maples, red oaks, white oaks, black oaks and miscellaneous trees). The original plot size (924 x 924 feet) has been rescaled to the unit square.

Note that the data contain duplicated points (two points at the same location). To determine which points are duplicates, use `duplicated.ppp`. To remove the duplication, use `unique.ppp`.

Usage

```
data(lansing)
```

Format

An object of class `"ppp"` representing the point pattern of tree locations. Entries include

x	Cartesian x-coordinate of tree
y	Cartesian y-coordinate of tree
marks	factor with levels indicating species of each tree

The levels of `marks` are `blackoak`, `hickory`, `maple`, `misc`, `redoak` and `whiteoak`. See `ppp.object` for details of the format of a point pattern object.

References

Besag, J. (1978) Some methods of statistical analysis for spatial data. *Bull. Internat. Statist. Inst.* **44**, 77–92.

Cox, T.F. (1976) The robust estimation of the density of a forest stand using a new conditioned distance method. *Biometrika* **63**, 493–500.

Cox, T.F. (1979) A method for mapping the dense and sparse regions of a forest stand. *Applied Statistics* **28**, 14–19.

Cox, T.F. and Lewis, T. (1976) A conditioned distance ratio method for analysing spatial patterns. *Biometrika* **63**, 483–492.

Diggle, P.J. (1979a) The detection of random heterogeneity in plant populations. *Biometrics* **33**, 390–394.

Diggle, P.J. (1979b) Statistical methods for spatial point patterns in ecology. *Spatial and temporal analysis in ecology*. R.M. Cormack and J.K. Ord (eds.) Fairland: International Co-operative Publishing House. pages 95–150.

Diggle, P.J. (1981) Some graphical methods in the analysis of spatial point patterns. In *Interpreting Multivariate Data*. V. Barnett (eds.) John Wiley and Sons. Pages 55–73.

Diggle, P.J. (1983) *Statistical analysis of spatial point patterns*. Academic Press.

Gerrard, D.J. (1969) Competition quotient: a new measure of the competition affecting individual forest trees. Research Bulletin 20, Agricultural Experiment Station, Michigan State University.

Lotwick, H.W. (1981) *Spatial stochastic point processes*. PhD thesis, University of Bath, UK.

Ord, J.K. (1978) How many trees in a forest? *Mathematical Scientist* **3**, 23–33.

Examples

```
    data(lansing)
  if(require(spatstat.geom)) {
    plot(lansing)
    summary(lansing)
    plot(split(lansing))
    plot(split(lansing)$maple)
    ##  rescale to feet
    (Lan <- rescale(lansing))
  }
```

---

letterR   *Window in Shape of Letter R*

Description

A window in the shape of the capital letter R, for use in demonstrations.

Usage

```
  data(letterR)
```

Format

An object of class ″owin″ representing the capital letter R, in the same font as the R package logo. See `owin.object` for details of the format.

Source

Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>

---

longleaf   *Longleaf Pines Point Pattern*

Description

Locations and sizes of Longleaf pine trees. A marked point pattern.

The data record the locations and diameters of 584 Longleaf pine (*Pinus palustris*) trees in a 200 x 200 metre region in southern Georgia (USA). They were collected and analysed by Platt, Evans and Rathbun (1988).

This is a marked point pattern; the mark associated with a tree is its diameter at breast height (dbh), a convenient measure of its size. Several analyses have considered only the "adult" trees which are conventionally defined as those trees with dbh greater than or equal to 30 cm.

The pattern is regarded as spatially inhomogeneous.

Usage

```
data(longleaf)
```

Format

An object of class ″ppp″ representing the point pattern of tree locations. Entries include

x	Cartesian x-coordinate of tree in metres
y	Cartesian y-coordinate of tree in metres
marks	diameter at breast height, in centimetres.

See `ppp.object` for details of the format of a point pattern object.

Source

Platt, Evans and Rathbun (1988)

References

Platt, W. J., Evans, G. W. and Rathbun, S. L. (1988) The population dynamics of a long-lived Conifer (Pinus palustris). *The American Naturalist* **131**, 491–525.

Rathbun, S. L. and Cressie, N. (1994) A space-time survival point process for a longleaf pine forest in southern Georgia. *Journal of the American Statistical Association* **89**, 1164–1173.

Examples

```
   data(longleaf)
 if(require(spatstat.geom)) {
   plot(longleaf)
   plot(cut(longleaf, breaks=c(0,30,Inf), labels=c("Sapling","Adult")))
 }
```

---

mucosa   *Cells in Gastric Mucosa*

Description

A bivariate inhomogeneous point pattern, giving the locations of the centres of two types of cells in a cross-section of the gastric mucosa of a rat.

Usage

```
data(mucosa)
```

Format

An object of class ″ppp″, see `ppp.object`. This is a multitype point pattern with two types of points, ECL and other.

Details

This point pattern dataset gives the locations of cell centres in a cross-section of the gastric mucosa (mucous membrane of the stomach) of a rat. The rectangular observation window has been scaled to unit width. The lower edge of the window is closest to the outside of the stomach.

The cells are classified into two types: *ECL cells* (enterochromaffin-like cells) and *other* cells. There are 86 ECL cells and 807 other cells in the dataset. ECL cells are a type of neuroendocrine cell which synthesize and secrete histamine. One hypothesis of interest is whether the spatially-varying intensities of ECL cells and other cells are proportional.

The data were originally collected by Dr Thomas Berntsen. The data were discussed and analysed in Møller and Waagepetersen (2004, pp. 2, 169).

The associated object `mucosa.subwin` is the smaller window to which the data were restricted for analysis by Møller and Waagepetersen.

The scale of spatial coordinates is unknown (R. Waagepetersen, personal communication).

Source

Dr Thomas Berntsen and Prof Rasmus Waagepetersen.

References

Møller, J. and Waagepetersen, R. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC.

Examples

```
if(require(spatstat.geom)) {
plot(mucosa, chars=c(1,3), cols=c("red", "green"))
plot(mucosa.subwin, add=TRUE, lty=3)
}
```

---

murchison   *Murchison gold deposits*

Description

Data recording the spatial locations of gold deposits and associated geological features in the Murchison area of Western Australia. Extracted from a large scale (1:500,000) study of the Murchison area by the Geological Survey of Western Australia (Watkins and Hickman, 1990). The features recorded are

- the locations of gold deposits;
- the locations of geological faults;
- the region that contains greenstone bedrock.

The study region is contained in a 330×400 kilometre rectangle. At this scale, gold deposits are points, i.e. their spatial extent is negligible. Gold deposits in this region occur only in greenstone bedrock. Geological faults can be observed reliably only within the same region. However, some faults have been extrapolated (by geological "interpretation") outside the greenstone boundary from information observed in the greenstone region.

Deposit locations were extracted from the Minedex database (Geological Survey of Western Australia, n.d.) and include deposits of all sizes. The fault geometry and greenstone boundaries were mapped and collated by Watkins and Hickman (1990).

These data were analysed by Foxall and Baddeley (2002) and Brown et al (2002); see also Groves et al (2000), Knox-Robinson and Groves (1997), Baddeley, Rubak and Turner (2015) and Baddeley (2019). The main aim is to predict the intensity of the point pattern of gold deposits from the more easily observable fault pattern.

Usage

```
data(murchison)
```

Format

murchison is a list with the following entries:

gold a point pattern (object of class `"ppp"`) representing the point pattern of gold deposits. See `ppp.object` for details of the format.

faults a line segment pattern (object of class `"psp"`) representing the geological faults. See `psp.object` for details of the format.

greenstone the greenstone bedrock region. An object of class `"owin"`. Consists of multiple irregular polygons with holes.

All coordinates are given in **metres**.

Source

Data were kindly provided by Dr Carl Knox-Robinson of the Department of Geology and Geophysics, University of Western Australia. Permission to use the data is granted by Dr Tim Griffin, Geological Survey of Western Australia and by Dr Knox-Robinson. *Please make appropriate acknowledgement* to Watkins and Hickman (1990) and the Geological Survey of Western Australia.

References

Baddeley, A. (2018) A statistical commentary on mineral prospectivity analysis. In Daya Sagar, B.S., Cheng, Q. and Agterberg, F.P. (eds.) *Handbook of Mathematical Geosciences: Fifty Years of IAMG*. International Association for Mathematical Geosciences. Chapter 2, pages 25–65.

Baddeley, A., Rubak, E. and Turner, R. (2015) *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press.

Brown, W.M., Gedeon, T.D., Baddeley, A.J. and Groves, D.I. (2002) Bivariate J-function and other graphical statistical methods help select the best predictor variables as inputs for a neural network method of mineral prospectivity mapping. In U. Bayer, H. Burger and W. Skala (eds.) *IAMG 2002:*

8th Annual Conference of the International Association for Mathematical Geology, Volume 1, 2002. International Association of Mathematical Geology. Pages 257–268.

Foxall, R. and Baddeley, A. (2002) Nonparametric measures of association between a spatial point process and a random set, with geological applications. *Applied Statistics* **51**, 165–182.

Geological Survey of Western Australia (n.d.) MINEDEX database of Mines and Mineral Deposits. <http://dmp.wa.gov.au/Mines-and-mineral-deposits-1502.aspx>.

Groves, D.I., Goldfarb, R.J., Knox-Robinson, C.M., Ojala, J., Gardoll, S, Yun, G.Y. and Holyland, P. (2000) Late-kinematic timing of orogenic gold deposits and significance for computer-based exploration techniques with emphasis on the Yilgarn Block, Western Australia. *Ore Geology Reviews*, **17**, 1–38.

Knox-Robinson, C.M. and Groves, D.I. (1997) Gold prospectivity mapping using a geographic information system (GIS), with examples from the Yilgarn Block of Western Australia. *Chronique de la Recherche Miniere* **529**, 127–138.

Watkins, K.P. and Hickman, A.H. (1990) *Geological evolution and mineralization of the Murchison Province, Western Australia*. Bulletin 137, Geological Survey of Western Australia. 267 pages. Published by Department of Mines, Western Australia, 1990. Available online from Department of Industry and Resources, State Government of Western Australia, www.doir.wa.gov.au

Examples

```
if(require(spatstat.geom)) {
if(interactive()) {
data(murchison)
plot(murchison$greenstone, main="Murchison data", col="lightgreen")
plot(murchison$gold, add=TRUE, pch="+",col="blue")
plot(murchison$faults, add=TRUE, col="red")
}
## rescale to kilometres
Mur <- solapply(murchison, rescale, s=1000, unitname="km")
}
```

---

nbfires   *Point Patterns of New Brunswick Forest Fires*

Description

Point patterns created from yearly records, provided by the New Brunswick Department of Natural Resources, of all fires falling under their jurisdiction for the years 1987 to 2003 inclusive (with the year 1988 omitted until further notice).

Usage

```
data(nbfires)
```

Format

Executing `data(nbfires)` gives access to four objects: `nbfires`, `nbw.rect`, `nbw.seg` and `nbfires.extra`.

The object `nbfires` is a marked point pattern (an object of class `"ppp"`) consisting of all of the fires in the years 1987 to 2003 inclusive, with the omission of 1988. The marks consist of a data frame of auxiliary information about the fires; see *Details.* Patterns for individual years can be extracted using the function `split.ppp()`. (See **Examples**.)

The object `nbw.rect` is a rectangular window which covers central New Brunswick. It is provided for use in illustrative and 'practice' calculations inasmuch as the use of a rectangular window simplifies some computations considerably.

The object `nbw.seg` is a line segment pattern (object of class `"psp"`) consisting of all the boundary segments of the polygonal window of New Brunswick. The segments are classified into different types of boundary by `marks(nbw.seg)`. This is a data frame with three columns:

- The column `type` describes the physical type of the border. It is a factor with levels `"land"` (land border), `"river"` (river border), `"coast"` (coast of the mainland) and `"island"` (coast of the 5 islands). To plot this classification, type `plot(nbw.seg)`.

- The column `share` specifies the territory which shares the border with New Brunswick. It is a factor with levels `"Quebec"`, `"NovaScotia"`, `"USA"` and `"water"`. To plot this classification, type `plot(nbw.seg,which.marks="share")`.

- The column `full` specifies both the physical type of border and the adjacent territory. It is a factor with levels `"coast"`, `"island"`, `"landNovaScotia"`, `"landQuebec"`, `"riverQuebec"`, `"landUSA"`, `"riverUSAnorth"`, `"riverUSAsouth"`. To plot this classification, type `plot(nbw.seg,which.marks="ful`

For conformity with other datasets, `nbfires.extra` is a list containing all the supplementary data. It contains copies of `nbw.rect` and `nbw.seg`.

Details

The coordinates of the fire locations were provided in terms of latitude and longitude, to the nearest minute of arc. These were converted to New Brunswick stereographic projection coordinates (Thomson, Mephan and Steeves, 1977) which was the coordinate system in which the map of New Brunswick — which constitutes the observation window for the pattern — was obtained. The conversion was done using a C program kindly provided by Jonathan Beaudoin of the Department of Geodesy and Geomatics, University of New Brunswick.

Finally the data and window were rescaled since the use of the New Brunswick stereographic projection coordinate system resulted in having to deal with coordinates which are expressed as very large integers with a bewildering number of digits. Amongst other things, these huge numbers tended to create very untidy axis labels on graphs. The width of the bounding box of the window was made equal to 1000 units. In addition the lower left hand corner of this bounding box was shifted to the origin. The height of the bounding box was changed proportionately, resulting in a value of approximately 959.

In the final dataset `nbfires`, one coordinate unit is equivalent to 0.403716 kilometres. To convert the data to kilometres, use `rescale(nbfires)`.

The window for the fire patterns comprises 6 polygonal components, consisting of mainland New Brunswick and the 5 largest islands. Some lakes which should form holes in the mainland component are currently missing; this problem may be remedied in future releases. The window was

formed by 'simplifying' the map that was originally obtained. The simplification consisted in reducing (using an interactive visual technique) the number of polygon edges in each component. For instance the number of edges in the mainland component was reduced from over 138,000 to 500.

For some purposes it is probably better to use a discretized (mask type) window. See **Examples**.

Because of the coarseness of the coordinates of the original data (1 minute of longitude is approximately 1 kilometer at the latitude of New Brunswick), data entry errors, and the simplification of the observation window, many of the original fire locations appeared to be outside of the window. This problem was addressed by shifting the location of the 'outsider' points slightly, or deleting them, as seemed appropriate.

Note that the data contain duplicated points (two points at the same location). To determine which points are duplicates, use duplicated.ppp. To remove the duplication, use unique.ppp.

The columns of the data frame comprising the marks of nbfires are:

year This a *factor* with levels 1987, 1989, ..., 2002, 2003. Note that 1988 is not present in the levels.

fire.type A factor with levels forest, grass, dump, and other.

dis.date The discovery date of the fire, which is the nearest possible surrogate for the starting time of the fire. This is an object of class POSIXct and gives the starting discovery time of the fire to the nearest minute.

dis.julian The discovery date and time of the fire, expressed in 'Julian days', i.e. as a decimal fraction representing the number of days since the beginning of the year (midnight 31 December).

out.date The date on which the fire was judged to be 'out'. This is an object of class POSIXct and gives the 'out' time of the fire to the nearest minute.

out.julian The date and time at which the fire was judged to be 'out', expressed in Julian days.

cause General cause of the fire. This is a factor with levels unknown, rrds (railroads), misc (miscellaneous), ltning (lightning), for.ind (forest industry), incend (incendiary), rec (recreation), resid (resident), and oth.ind (other industry). Causes unknown, ltning, and incend are supposedly designated as 'final' by the New Brunswick Department of Natural Resources, meaning (it seems) "that's all there is to it". Other causes are apparently intended to be refined by being combined with "source of ignition". However cross-tabulating cause with ign.src — see below — reveals that very often these three 'causes' are associated with an "ignition source" as well.

ign.src Source of ignition, a factor with levels cigs (cigarette/match/pipe/ashes), burn.no.perm (burning without a permit), burn.w.perm (burning with a permit), presc.burn (prescribed burn), wood.spark (wood spark), mach.spark (machine spark), campfire, chainsaw, machinery, veh.acc (vehicle accident), rail.acc (railroad accident), wheelbox (wheelbox on railcars), hot.flakes (hot flakes off railcar wheels), dump.fire (fire escaping from a dump), ashes (ashes, briquettes, burning garbage, etc.)

fnl.size The final size of the fire (area burned) in hectares, to the nearest 10th hectare.

Note that due to data entry errors some of the "out dates" and "out times" in the original data sets were actually *earlier* than the corresponding "discovery dates" and "discover times". In such cases all corresponding entries of the marks data frame (i.e. dis.date, dis.julian, out.date, and out.julian) were set equal to NA. Also, some of the dates and times were missing (equal to NA) in the original data sets.

The 'ignition source' data were given as integer codes in the original data sets. The code book that I obtained gave interpretations for codes 1, 2, . . . , 15. However the actually also contained codes of 0, 16, 17, 18, and in one instance 44. These may simply be data entry errors. These uninterpretable values were assigned the level unknown. Many of the years had most, or sometimes all, of the ignition source codes equal to 0 (hence turning out as unknown, and many of the years had many missing values as well. These were also assigned the level unknown. Of the 7108 fires in nbfires, 4354 had an unknown ignition source. This variable is hence unlikely to be very useful.

There are also anomalies between cause and ign.src, e.g. cause being unknown but ign.src being cigs, burn.no.perm, mach.spark, hot.flakes, dump.fire or ashes. Particularly worrisome is the fact that the cause ltning (!!!) is associate with sources of ignition cigs, burn.w.perm, presc.burn, and wood.spark.

Source

The data were kindly provided by the New Brunswick Department of Natural Resources. Special thanks are due to Jefferey Betts for a great deal of assistance.

References

Turner, Rolf. Point patterns of forest fire locations. *Environmental and Ecological Statistics* **16** (2009) 197 – 223, DOI:10.1007/s10651-007-0085-1.

Thomson, D. B., Mephan, M. P., and Steeves, R. R. (1977) The stereographic double projection. Technical Report 46, University of New Brunswick, Fredericton, N. B., Canada URL: gge.unb.ca/Pubs/Pubs.html.

Examples

```
if(interactive()) {
  if(require(spatstat.geom)) {
# Get the year 2000 data.
X <- split(nbfires,"year")
Y.00 <- X[["2000"]]
# Plot all of the year 2000 data, marked by fire type.
plot(Y.00,which.marks="fire.type")
# Cut back to forest and grass fires.
Y.00 <- Y.00[marks(Y.00)$fire.type %in% c("forest","grass")]
# Plot the year 2000 forest and grass fires marked by fire duration time.
stt  <- marks(Y.00)$dis.julian
fin  <- marks(Y.00)$out.julian
marks(Y.00) <- cbind(marks(Y.00),dur=fin-stt)
plot(Y.00,which.marks="dur")
# Look at just the rectangular subwindow (superimposed on the entire window).
nbw.mask <- as.mask(Window(nbfires), dimyx=500)
plot(nbw.mask, col=c("green", "white"))
plot(Window(nbfires), border="red", add=TRUE)
plot(Y.00[nbw.rect],use.marks=FALSE,add=TRUE)
plot(nbw.rect,add=TRUE,border="blue")
  if(require(spatstat.explore)) {
    # Look at the K function for the year 2000 forest and grass fires.
    K.00 <- Kest(Y.00)
    plot(K.00)
    }
```

```
# Rescale to kilometres
NBF <- rescale(nbfires)
  }
}
```

---

nztrees   *New Zealand Trees Point Pattern*

Description

The data give the locations of trees in a forest plot.

They were collected by Mark and Esler (1970) and were extracted and analysed by Ripley (1981, pp. 169-175). They represent the positions of 86 trees in a forest plot approximately 140 by 85 feet.

Ripley discarded from his analysis the eight trees at the right-hand edge of the plot (which appear to be part of a planted border) and trimmed the window by a 5-foot margin accordingly.

Usage

```
data(nztrees)
```

Format

An object of class ″ppp″ representing the point pattern of tree locations. The Cartesian coordinates are in feet.

See `ppp.object` for details of the format of a point pattern object.

Note

To trim a 5-foot margin off the window, type nzsub <- nztrees[owin(c(0,148),c(0,95))]

Source

Mark and Esler (1970), Ripley (1981).

References

Ripley, B.D. (1981) *Spatial statistics*. John Wiley and Sons.

Mark, A.F. and Esler, A.E. (1970) An assessment of the point-centred quarter method of plotless sampling in some New Zealand forests. *Proceedings of the New Zealand Ecological Society* **17**, 106–110.

---

osteo   *Osteocyte Lacunae Data: Replicated Three-Dimensional Point Patterns*

Description

These data give the three-dimensional locations of osteocyte lacunae observed in rectangular volumes of solid bone using a confocal microscope.

There were four samples of bone, and ten regions were mapped in each bone, yielding 40 spatial point patterns. The data can be regarded as replicated observations of a three-dimensional point process, nested within bone samples.

Usage

```
data(osteo)
```

Format

A hyperframe with the following columns:

id	character string identifier of bone sample
shortid	last numeral in id
brick	serial number (1 to 10) of sampling volume within this bone sample
pts	three dimensional point pattern (class pp3)
depth	the depth of the brick in microns

Details

These data are three-dimensional point patterns representing the positions of *osteocyte lacunae*, holes in bone which were occupied by osteocytes (bone-building cells) during life.

Observations were made on four different skulls of Macaque monkeys iusing a three-dimensional microscope. From each skull, observations were collected in 10 separate sampling volumes. In all, there are 40 three-dimensional point patterns in the dataset.

The data were collected in 1984 by A. Baddeley, A. Boyde, C.V. Howard and S. Reid (see references) using the tandem-scanning reflected light microscope (TSRLM) at University College London. This was one of the first optical confocal microscopes available.

Each point pattern dataset gives the (x, y, z) coordinates (in microns) of all points visible in a three-dimensional rectangular box ("brick") of dimensions $81 \times 100 \times d$ microns, where d varies. The z coordinate is depth into the bone (depth of the focal plane of the confocal microscope); the (x, y) plane is parallel to the exterior surface of the bone; the relative orientation of the x and y axes is not important.

The bone samples were three intact skulls and one skull cap, all originally identified as belonging to the macaque monkey *Macaca fascicularis*, from the collection of the Department of Anatomy, University of London. Later analysis (Baddeley et al, 1993) suggested that the skull cap, given here as the first animal, was a different subspecies, and this was confirmed by anatomical inspection.

Sampling Procedure

The following extract from Baddeley et al (1987) describes the sampling procedure.

The parietal bones of three fully articulated adult Macaque monkey *(Macaca fascicularis)* skulls from the collection of University College London were used. The right parietal bone was examined, in each case, approximately 1 cm lateral to the sagittal suture and 2 cm posterior to the coronal suture. The skulls were mounted on plasticine on a moving stage placed beneath the TSRLM. Immersion oil was applied and a $\times 60$, NA 1.0 oil immersion objective lens (Lomo) was focussed at 10 microns below the cranial surface. The TV image was produced by a Panasonic WB 1850/B camera on a Sony PVM 90CE TV monitor.

A graduated rectangular counting frame 90×110 mm (representing 82×100 microns in real units) was marked on a Perspex overlay and fixed to the screen. The area of tissue seen within the frame defined a subfield: a guard area of 10 mm width was visible on all sides of the frame. Ten subfields were examined, arranged approximately in a rectangular grid pattern, with at least one field width separating each pair of fields. The initial field position was determined randomly by applying a randomly-generated coordinate shift to the moving stage. Subsequent fields were attained using the coarse controls of the microscope stage, in accordance with the rectangular grid pattern.

For each subfield, the focal plane was racked down from its initial 10 micron depth until all visible osteocyte lacunae had been examined. This depth d was recorded. The 3-dimensional sampling volume was therefore a rectangular box of dimensions $82 \times 100 \times d$ microns, called a "brick". For each visible lacuna, the fine focus racking control was adjusted until maximum brightness was obtained. The depth of the focal plane was then recorded as the z coordinate of the "centre point" of the lacuna. Without moving the focal plane, the x and y coordinates of the centre of the lacunar image were read off the graduated counting frame. This required a subjective judgement of the position of the centre of the 2-dimensional image. Profiles were approximately elliptical and the centre was considered to be well-defined. Accuracy of the recording procedure was tested by independent repetition (by the same operator and by different operators) and found to be reproducible to plus or minus 2 mm on the screen.

A lacuna was counted only if its (x, y) coordinates lay inside the 90×110 mm counting frame.

Source

Data were collected by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>.

References

Baddeley, A.J., Howard, C.V, Boyde, A. and Reid, S.A. (1987) Three dimensional analysis of the spatial distribution of particles using the tandem-scanning reflected light microscope. *Acta Stereologica* **6** (supplement II) 87–100.

Baddeley, A.J., Moyeed, R.A., Howard, C.V. and Boyde, A. (1993) Analysis of a three-dimensional point pattern with replication. *Applied Statistics* **42** (1993) 641–668.

Howard, C.V. and Reid, S. and Baddeley, A.J. and Boyde, A. (1985) Unbiased estimation of particle density in the tandem-scanning reflected light microscope. *Journal of Microscopy* **138** 203–212.

Examples

```
data(osteo)
if(require(spatstat.geom)) {
```

```
osteo
if(interactive()) {
  plot(osteo$pts[[1]], main="animal 1, brick 1")
  ape1 <- osteo[osteo$shortid==4, ]
  plot(ape1, tick.marks=FALSE)
  with(osteo, intensity(pts))
  plot(with(ape1, K3est(pts)))
}
}
```

---

paracou   *Kimboto trees at Paracou, French Guiana*

Description

This dataset is a point pattern of adult and juvenile Kimboto trees (*Pradosia cochlearia* or *P. pty-chandra*) recorded at Paracou in French Guiana. See Flores (2005).

The dataset paracou is a point pattern (object of class "ppp") containing the spatial coordinates of each tree, marked by age (a factor with levels adult and juvenile. The survey region is a rectangle approximately 400 by 525 metres. Coordinates are given in metres.

Note that the data contain duplicated points (two points at the same location). To determine which points are duplicates, use duplicated.ppp. To remove the duplication, use unique.ppp.

Usage

```
data(paracou)
```

Source

Data kindly contributed by Olivier Flores. All data belong to CIRAD https://www.cirad.fr and UMR EcoFoG http://www.ecofog.gf and are included in **spatstat** with permission. Original data sources: juvenile and some adult trees collected by Flores (2005); adult tree data sourced from CIRAD Paracou experimental plots dataset (2003 campaign).

References

Flores, O. (2005) *Déterminisme de la régénération chez quinze espèces d'arbres tropicaux en forêt guyanaise: les effets de l'environnement et de la limitation par la dispersion.* PhD Thesis, University of Montpellier 2, Montpellier, France.

Picard, N, Bar-Hen, A., Mortier, F. and Chadoeuf, J. (2009) The multi-scale marked area-interaction point process: a model for the spatial pattern of trees. *Scandinavian Journal of Statistics* **36** 23–41

Examples

```
  if(require(spatstat.geom)) {
plot(paracou, cols=2:3, chars=c(16,3))
  }
```

ponderosa	*Ponderosa Pine Tree Point Pattern*

Description

The data record the locations of 108 Ponderosa Pine (*Pinus ponderosa*) trees in a 120 metre square region in the Klamath National Forest in northern California, published as Figure 2 of Getis and Franklin (1987).

Franklin et al. (1985) determined the locations of approximately 5000 trees from United States Forest Service aerial photographs and digitised them for analysis. Getis and Franklin (1987) selected a 120 metre square subregion that appeared to exhibit clustering. This subregion is the ponderosa dataset.

In principle these data are equivalent to Figure 2 of Getis and Franklin (1987) but they are not exactly identical; some of the spatial locations appear to be slightly perturbed.

The data points identified as A, B, C on Figure 2 of Getis and Franklin (1987) correspond to points numbered 42, 7 and 77 in the dataset respectively.

Usage

```
data(ponderosa)
```

Format

Typing data(ponderosa) gives access to two objects, ponderosa and ponderosa.extra.

The dataset ponderosa is a spatial point pattern (object of class "ppp") representing the point pattern of tree positions. See ppp.object for details of the format. Spatial coordinates are given in metres.

The dataset ponderosa.extra is a list containing supplementary data. The entry id contains the index numbers of the three special points A, B, C in the point pattern. The entry plotit is a function that can be called to produce a nice plot of the point pattern.

Source

Prof. Janet Franklin, University of California, Santa Barbara

References

Franklin, J., Michaelsen, J. and Strahler, A.H. (1985) Spatial analysis of density dependent pattern in coniferous forest stands. *Vegetatio* **64**, 29–36.

Getis, A. and Franklin, J. (1987) Second-order neighbourhood analysis of mapped point patterns. *Ecology* **68**, 473–477.

Examples

```
  data(ponderosa)
 if(require(spatstat.geom)) {
 ponderosa.extra$plotit()
  }
```

---

pyramidal   *Pyramidal Neurons in Cingulate Cortex*

Description

Point patterns giving the locations of pyramidal neurons in micrographs from area 24, layer 2 of the cingulate cortex in the human brain. There is one point pattern from each of 31 human subjects. The subjects are divided into three groups: controls (12 subjects), schizoaffective (9 subjects) and schizophrenic (10 subjects).

Each point pattern is recorded in a unit square region; the unit of measurement is unknown.

These data were introduced and analysed by Diggle, Lange and Benes (1991).

Usage

```
data(pyramidal)
```

Format

pyramidal is a hyperframe with 31 rows, one row for each subject. It has a column named Neurons containing the point patterns of neuron locations, and a column named group which is a factor with levels "control", "schizoaffective", "schizophrenic" identifying the grouping of subjects.

Source

Peter Diggle's website.

References

Diggle, P.J., Lange, N. and Benes, F.M. (1991). Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. *Journal of the American Statistical Association* **86**, 618–625.

Examples

```
  if(require(spatstat.geom)) {
pyr <- pyramidal
pyr$grp <- abbreviate(pyramidal$group, minlength=7)
plot(pyr, quote(plot(Neurons, pch=16, main=grp)), main="Pyramidal Neurons")
  }
```

---

redwood   *California Redwoods Point Pattern (Ripley's Subset)*

Description

The data represent the locations of 62 seedlings and saplings of California Giant Redwood (*Sequoiadendron giganteum*) recorded in a square sampling region. They originate from Strauss (1975); the present data are a subset extracted by Ripley (1977) in a subregion that has been rescaled to a unit square. (The original physical size of the unit is approximately 63.1 feet).

Two versions of this dataset are provided: redwood and redwood3.

The dataset redwood was obtained from the **spatial** package. In this version the coordinates are given to 2 decimal places (multiples of 0.01 units) except for one point which has an x coordinate of 0.999, presumably to ensure that it is properly inside the window.

The dataset redwood3 was obtained from Peter Diggle's webpage. In this version the coordinates are given to 3 decimal places (multiples of 0.001 units). The ordering of the points is not the same in the two datasets.

There are many further analyses of this dataset. It is often used as a canonical example of a clustered point pattern (see e.g. Diggle, 1983).

The original, full redwood dataset is supplied in the spatstat.data package as redwoodfull.

Usage

```
data(redwood)
```

Format

An object of class "ppp" representing the point pattern of tree locations. The window has been rescaled to the unit square.

See [ppp.object](#) for details of the format of a point pattern object.

Source

Original data of Strauss (1975), subset extracted by Ripley (1977). Data obtained from Ripley's package **spatial** and from Peter Diggle's website.

References

Diggle, P.J. (1983) *Statistical analysis of spatial point patterns*. Academic Press.

Ripley, B.D. (1977) Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B* **39**, 172–212.

Strauss, D.J. (1975) A model for clustering. *Biometrika* **62**, 467–475.

See Also

[redwoodfull](#)

---

redwoodfull   *California Redwoods Point Pattern (Entire Dataset)*

Description

These data represent the locations of 195 seedlings and saplings of California Giant Redwood (*Sequoiadendron giganteum*) in a square sampling region.

They were described and analysed by Strauss (1975). This is the "**full**" dataset; most writers have analysed a subset extracted by Ripley (1977) which is available as redwood.

Strauss (1975) divided the sampling region into two subregions I and II demarcated by a diagonal line. The spatial pattern appears to be slightly regular in region I and strongly clustered in region II.

Strauss (1975) writes: "It was felt that the seedlings would be scattered fairly randomly, except that a number of tight clusters would form around some of the redwood tree stumps present in the plot. A discontinuity in the soil, very roughly demarked by the diagonal line in the figure, was expected to cause a difference in clustering behaviour between regions I and II. Moreover, almost all the redwood stumps were situated in region II."

The dataset redwoodfull contains the full point pattern of 195 trees. The window has been rescaled to the unit square. Its physical size is approximately 130 feet across.

The auxiliary information about the subregions is contained in redwoodfull.extra, which is a list with entries

rdiag	The coordinates of the diagonal boundary
	between regions I and II
regionI	Region I as a window object
regionII	Region II as a window object
regionR	Ripley's subrectangle (approximate)
plotit	Function to plot the full data and auxiliary markings

Ripley (1977) extracted a subset of these data, containing 62 points, lying within a square subregion which overlaps regions I and II. He rescaled that subset to the unit square. This subset has been re-analysed many times, and is the dataset usually known as "the redwood data" in the spatial statistics literature. The exact dataset used by Ripley is supplied in the **spatstat** library as redwood.

The approximate position of the square chosen by Ripley within the redwoodfull pattern is indicated by the window redwoodfull.extra$regionR. There are some minor inconsistencies with redwood since it originates from a different digitisation.

Usage

```
data(redwoodfull)
```

Format

The dataset redwoodfull is an object of class "ppp" representing the point pattern of tree locations. See ppp.object for details of the format of a point pattern object. The window has been rescaled to the unit square. Its physical size is approximately 128 feet across.

The dataset `redwoodfull.extra` is a list with entries

rdiag	coordinates of endpoints of a line, in format `list(x=numeric(2),y=numeric(2))`
regionI	a window object
regionII	a window object
regionR	a window object
plotit	Function with no arguments

Source

Strauss (1975). The plot of the data published by Strauss (1975) was scanned and digitised by Sandra Pereira, University of Western Australia, 2002.

References

Diggle, P.J. (1983) *Statistical analysis of spatial point patterns*. Academic Press.

Ripley, B.D. (1977) Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B* **39**, 172–212.

Strauss, D.J. (1975) A model for clustering. *Biometrika* **62**, 467–475.

See Also

[redwood](#)

Examples

```
    data(redwoodfull)
if(require(spatstat.geom)) {
    plot(redwoodfull)
    redwoodfull.extra$plotit()
    # extract the pattern in region II
    redwoodII <- redwoodfull[, redwoodfull.extra$regionII]
 }
```

---

residualspaper	*Data and Code From JRSS Discussion Paper on Residuals*

Description

This dataset contains the point patterns used as examples in the paper of Baddeley et al (2005). [Figure 2 is already available in **spatstat.data** as the [copper](#) dataset.]

R code is also provided to reproduce all the Figures displayed in Baddeley et al (2005). The component `plotfig` is a function, which can be called with a numeric or character argument specifying the Figure or Figures that should be plotted. See the Examples.

Usage

```
data(residualspaper)
```

Format

residualspaper is a list with the following components:

Fig1 The locations of Japanese pine seedlings and saplings from Figure 1 of the paper. A point pattern (object of class "ppp").

Fig3 The Chorley-Ribble data from Figure 3 of the paper. A list with three components, lung, larynx and incin. Each is a matrix with 2 columns giving the coordinates of the lung cancer cases, larynx cancer cases, and the incinerator, respectively. Coordinates are Eastings and Northings in km.

Fig4a The synthetic dataset in Figure 4 (a) of the paper.

Fig4b The synthetic dataset in Figure 4 (b) of the paper.

Fig4c The synthetic dataset in Figure 4 (c) of the paper.

Fig11 The covariate displayed in Figure 11. A pixel image (object of class "im") whose pixel values are distances to the nearest line segment in the copper data.

plotfig A function which will compute and plot any of the Figures from the paper. The argument of plotfig is either a numeric vector or a character vector, specifying the Figure or Figures to be plotted. See the Examples.

Source

Figure 1: Prof M. Numata. Data kindly supplied by Professor Y. Ogata with kind permission of Prof M. Tanemura.

Figure 3: Professor P.J. Diggle (rescaled by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>)

Figure 4 (a,b,c): Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>

References

Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005) Residual analysis for spatial point processes. *Journal of the Royal Statistical Society, Series B* **67**, 617–666.

Examples

```
if(FALSE) {
  data(residualspaper)

  if(require(spatstat.model)) {

  X <- residualspaper$Fig4a
  summary(X)
  plot(X)

  # reproduce all Figures
  residualspaper$plotfig()
```

```
    # reproduce Figures 1 to 10
    residualspaper$plotfig(1:10)

    # reproduce Figure 7 (a)
    residualspaper$plotfig("7a")
    }
  }
```

---

shapley   *Galaxies in the Shapley Supercluster*

Description

A point pattern recording the sky positions of 4215 galaxies in the Shapley Supercluster.

Usage

```
data(shapley)
```

Format

shapley is an object of class ″ppp″ representing the point pattern of galaxy locations (see `ppp.object`).

shapley.extra is a list containing additional data described under Notes.

Notes

This dataset comes from a survey by Drinkwater et al (2004) of the Shapley Supercluster, one of the most massive concentrations of galaxies in the local universe. The data give the sky positions of 4215 galaxies observed using the FLAIR-II spectrograph on the UK Schmidt Telescope (UKST). They were kindly provided by Dr Michael Drinkwater through the Centre for Astrostatistics at Penn State University.

Sky positions are given using the coordinates Right Ascension (degrees from 0 to 360) and Declination (degrees from -90 to 90).

The point pattern has three mark variables:

Mag Galaxy magnitude (a negative logarithmic measure of visible brightness).

V Recession velocity (km/sec) inferred from redshift, with corrections applied.

SigV Estimated standard error for V.

The region covered by the survey was approximately the UKST's standard quadrilateral survey fields 382 to 384 and 443 to 446. However, a few of the galaxy positions lie outside these fields.

The point pattern dataset shapley consists of all 4215 galaxy locations. The observation window for this pattern is a dilated copy of the convex hull of the galaxy positions, constructed so that all galaxies lie within the window.

Note that the data contain duplicated points (two points at the same location). To determine which points are duplicates, use `duplicated.ppp`. To remove the duplication, use `unique.ppp`.

The auxiliary dataset shapley.extra contains the following components:

UKSTfields a list of seven windows (objects of class "owin") giving the UKST standard survey
fields.

UKSTdomain the union of these seven fields, an object of class "owin".

plotit a function (called without arguments) that will plot the data and the survey fields in the
conventional astronomical presentation, in which Right Ascension is converted to hours and
minutes (1 hour equals 15 degrees) and Right Ascension decreases as we move to the right of
the plot.

Source

M.J. Drinkwater, Department of Physics, University of Queensland

References

Drinkwater, M.J., Parker, Q.A., Proust, D., Slezak, E. and Quintana, H. (2004) The large scale
distribution of galaxies in the Shapley Supercluster. *Publications of the Astronomical Society of
Australia* **21**, 89-96. DOI 10.1071/AS03057

Examples

```
data(shapley)
if(require(spatstat.geom)) {
shapley.extra$plotit(main="Shapley Supercluster")
}
```

---

simba   *Simulated data from a two-group experiment with replication within
                                each group.*

Description

The simba dataset contains simulated data from an experiment with a 'control' group and a 'treat-
ment' group, each group containing 5 experimental units.

The responses in the experiment are point patterns.

The responses in the control group are independent realisations of a Poisson point process with
intensity 80.

The responses in the treatment group are independent realisations of a Strauss process with activity
parameter $\beta = 100$, interaction parameter $\gamma = 0.5$ and interaction radius $R = 0.07$ in the unit
square.

Usage

```
data(simba)
```

Format

simba is a hyperframe with 10 rows, and columns named:

- Points containing the point patterns

- group factor identifying the experimental group, with levels control and treatment).

Source

Simulated data, generated by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>.

---

simdat *Simulated Point Pattern*

Description

This point pattern data set was simulated (using the Metropolis-Hastings algorithm) from a model fitted to the Numata Japanese black pine data set referred to in Baddeley and Turner (2000).

Usage

```
data(simdat)
```

Format

An object of class ″ppp″ in a square window of size 10 by 10 units.

See [ppp.object](#) for details of the format of a point pattern object.

Source

Rolf Turner <r.turner@auckland.ac.nz>

References

Baddeley, A. and Turner, R. (2000) Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics* **42**, 283–322.

---

simplenet   *Simple Example of Linear Network*

Description

A simple, artificially created, example of a linear network.

Usage

```
data(simplenet)
```

Format

simplenet is an object of class "linnet".

Source

Created by Adrian Baddeley <Adrian.Baddeley@curtin.edu.au>.

---

spiders   *Spider Webs on Mortar Lines of a Brick Wall*

Description

Data recording the locations of small spider webs on the network of mortar lines of a brick wall.

Usage

```
data("spiders")
```

Format

Object of class "lpp" representing a pattern of points on a linear network. Spatial coordinates are expressed in millimetres.

Details

The data give the positions of 48 webs of the urban wall spider *Oecobius navus* on the mortar lines of a brick wall, recorded by Voss (1999) and manually digitised by Mark Handcock. The mortar spaces provide the only opportunity for constructing webs (Voss 1999; Voss et al 2007) so this is a pattern of points on a network of lines.

The habitat preferences of this species were studied in detail by Voss et al (2007). Questions of interest include evidence for non-uniform density of webs and for interaction between nearby individuals.

Observations were made inside a square quadrat of side length 1.125 metres. The original hand-drawn map was digitised manually by Mark S. Handcock, and reformatted as a `spatstat` object by Ang Qi Wei.

The dataset `spiders` is an object of class `"lpp"` (point pattern on a linear network). Coordinates are given in millimetres. The linear network has 156 vertices and a total length of 20.22 metres.

Please cite Voss et al (2007) with any use of these data.

Source

Dr Sasha Voss. Coordinates manually recorded by M.S. Handcock and formatted by Q.W. Ang.

Please cite Voss et al (2007) with any use of these data.

References

Ang, Q.W. (2010) *Statistical methodology for events on a network.* Master's thesis, School of Mathematics and Statistics, University of Western Australia.

Voss, S. (1999) Habitat preferences and spatial dynamics of the urban wall spider: *Oecobius annulipes* Lucas. Honours Thesis, Department of Zoology, University of Western Australia.

Voss, S., Main, B.Y. and Dadour, I.R. (2007) Habitat preferences of the urban wall spider *Oecobius navus* (Araneae, Oecobiidae). *Australian Journal of Entomology* **46**, 261–268.

Examples

```
  if(require(spatstat.linnet)) {
plot(spiders, show.window=FALSE, pch=16)
   }
```

---

sporophores	*Sporophores Data*

Description

Spatial pattern of sporophores of three species of fungi around a tree.

Usage

```
data(sporophores)
```

Format

A multitype spatial point pattern (an object of class `"ppp"` with factor-valued marks indicating the species). Spatial coordinates are given in centimetres. Levels of the species variable are `"L laccata"`, `"L pubescens"` and `"Hebloma spp"`.

Details

Ford, Mason and Pelham (1980) studied the spatial locations of sporophores of three species of mycorrhizal fungi distributed around a young birch tree in agricultural soil. The dataset given here is the spatial pattern in the fifth year after the tree was planted. The species are *Laccaria laccata*, *Lactarius pubescens* and *Hebloma* spp.

Source

Data generously provided by Dr E.D. Ford. Please cite Ford et al (1980) in any use of these data.

References

Ford, E.D., Mason, P.A. and Pelham, J. (1980) Spatial patterns of sporophore distribution around a young birch tree in three successive years. *Transactions of the British Mycological Society* **75**, 287–296.

Examples

```
  if(require(spatstat.geom)) {
## reproduce Fig 1 in Ford et al (1980)
plot(sporophores, chars=c(16,1,2), cex=0.6, leg.args=list(cex=1.1))
points(0,0,pch=16, cex=2)
text(15,8,"Tree", cex=0.75)
  }
```

---

spruces   *Spruces Point Pattern*

Description

The data give the locations of Norwegian spruce trees in a natural forest stand in Saxonia, Germany. Each tree is marked with its diameter at breast height.

Usage

```
data(spruces)
```

Format

An object of class "ppp" representing the point pattern of tree locations in a 56 x 38 metre sampling region. Each tree is marked with its diameter at breast height. All values are given in metres.

See `ppp.object` for details of the format of a point pattern object. The marks are numeric.

These data have been analysed by Fiksel (1984, 1988), Stoyan et al (1987), Penttinen et al (1992) and Goulard et al (1996).

Source

Stoyan et al (1987). Original source unknown.

References

Fiksel, T. (1984) Estimation of parameterized pair potentials of marked and nonmarked Gibbsian point processes. *Elektron. Informationsverarb. u. Kybernet.* **20**, 270–278.

Fiksel, T. (1988) Estimation of interaction potentials of Gibbsian point processes. *Statistics* **19**, 77-86

Goulard, M., S\"arkk\"a, A. and Grabarnik, P. (1996) Parameter estimation for marked Gibbs point processes through the maximum pseudolikelihood method. *Scandinavian Journal of Statistics* **23**, 365–379.

Penttinen, A., Stoyan, D. and Henttonen, H. (1992) Marked point processes in forest statistics. *Forest Science* **38**, 806–824.

Stoyan, D., Kendall, W.S. and Mecke, J. (1987) *Stochastic Geometry and its Applications*. Wiley.

Examples

```
if(require(spatstat.geom)) {
   plot(spruces)
   # To reproduce Goulard et al. Figure 3
   # (Goulard et al: "influence zone radius equals 5 * stem diameter")
   # (help(plot.ppp) says: "size of symbol = diameter")
   plot(spruces, maxsize=10*max(spruces$marks))
   plot(unmark(spruces), add=TRUE)
}
```

---

stonetools   *Palaeolithic Stone Tools*

Description

This dataset is a spatial point pattern giving the locations of palaeolithic stone tools ('lithic' specimens) and animal bone fragments ('bone'), accurately surveyed in a layer of soil at David's Site, Olduvai Gorge, Tanzania. The surveyed layer is about 20 cm thick and approximately 1.85 million years old.

Details of the study, and data analysis, are reported by Diez-Martín et al. (2021). Please cite this article in any use of the data.

The data are presented as a two-dimensional point pattern with two columns of marks: the vertical position Z (numeric) and the type of artefact TYPE (factor with levels BONE and LITHIC). The window of observation is an irregular polygon, approximately 40 by 30 metres across. Spatial coordinates and vertical coordinate are expressed in metres. There are 3563 bone fragments and 1182 lithic specimens making a total of 4745 points.

Usage

```
data("stonetools")
```

Format

Marked spatial point pattern (object of class ″ppp″, see `ppp.object`) with two columns of marks, Z (numeric) and TYPE (factor with levels BONE and LITHIC). The window of observation is an irregular polygon. Spatial coordinate unit is metres.

Source

Dr. L. Cobo and Prof. F. Diez-Martín. Please cite Diez-Martín et al (2021) in any use of these data.

References

Diez-Martín, F., Cobo-Sánchez, L., Baddeley, A., Uribelarrea, D., Mabulla, A., Baquedano, E. and Domínguez-Rodrigo, M. (2021) Tracing the spatial imprint of Oldowan technological behaviors: A view from DS (Bed I, Olduvai Gorge, Tanzania). *PLOS ONE, Public Library of Science*, **16**, 1–47. `DOI: 10.1371/journal.pone.0254603`

Examples

```
  if(require(spatstat.geom)) {
plot(subset(stonetools, select=TYPE), cex=0.5, cols=2:3)
  }
```

---

swedishpines	*Swedish Pines Point Pattern*

Description

The data give the locations of pine saplings in a Swedish forest.

Usage

```
data(swedishpines)
```

Format

An object of class ″ppp″ representing the point pattern of tree locations in a rectangular plot 9.6 by 10 metres.

Cartesian coordinates are given in *decimetres* (multiples of 0.1 metre) rounded to the nearest decimetre. Type `rescale`(swedishpines) to get an equivalent dataset where the coordinates are expressed in metres.

See `ppp.object` for details of the format of a point pattern object.

Note

For previous analyses see Ripley (1981, pp. 172-175), Venables and Ripley (1997, p. 483), Baddeley and Turner (2000).

Source

Strand (1972), Ripley (1981)

References

Baddeley, A. and Turner, R. (2000) Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics* **42**, 283–322.

Ripley, B.D. (1981) *Spatial statistics*. John Wiley and Sons.

Strand, L. (1972). A model for stand growth. *IUFRO Third Conference Advisory Group of Forest Statisticians*, INRA, Institut National de la Recherche Agronomique, Paris. Pages 207–216.

Venables, W.N. and Ripley, B.D. (1997) *Modern applied statistics with S-PLUS*. Second edition. Springer Verlag.

Examples

```
if(require(spatstat.geom)) {
   swedishpines

   ## rescale to metres
   rescale(swedishpines)
}
```

---

urkiola	*Urkiola Woods Point Pattern*

Description

Locations of birch (*Betula celtiberica*) and oak (*Quercus robur*) trees in a secondary wood in Urkiola Natural Park (Basque country, northern Spain). They are part of a more extensive dataset collected and analysed by Laskurain (2008). The coordinates of the trees are given in meters.

Usage

```
data(urkiola)
```

Format

An object of class ″ppp″ representing the point pattern of tree locations. Entries include

x Cartesian x-coordinate of tree in metres

y Cartesian y-coordinate of tree in metres

marks factor indicating species of each tree

The levels of marks are birch and oak. See `ppp.object` for details of the format of a ppp object.

Source

N.A. Laskurain. Kindly formatted and communicated by M. de la Cruz Rot

References

Laskurain, N. A. (2008) *Dinámica espacio-temporal de un bosque secundario en el Parque Natural de Urkiola (Bizkaia).* Tesis Doctoral. Universidad del País Vasco /Euskal Herriko Unibertsitatea.

---

vesicles   *Vesicles Data*

Description

Point pattern of synaptic vesicles observed in rat brain tissue.

Usage

```
data(vesicles)
```

Format

The dataset `vesicles` is a point pattern (object of class `"ppp"`) representing the location of the synaptic vesicles. The window of the point pattern represents the region of presynapse where synaptic vesicles were observed in this study. There is a hole in the window, representing the region occupied by mitochondria, where synaptic vesicles do not occur.

The dataset `vesicles.extra` is a list with entries

presynapse	outer polygonal boundary of presynapse
mitochondria	polygonal boundary of mitochondria
mask	binary mask representation of vesicles window
activezone	line segment pattern representing the active zone.

All coordinates are in nanometres (nm).

Details

As part of a study on the effects of stress on brain function, Khanmohammadi et al (2014) analysed the spatial pattern of synaptic vesicles in 45-nanometre-thick sections of rat brain tissue visualised in transmission electron microscopy.

To investigate the influence of stress, Khanmohammadi et al (2014) study the distribution of the synaptic vesicles in the pre-synaptic neuron in relation to the active zone of the presynaptic membrane. They hypothesize that the synaptic vesicle density is a decreasing function of distance to the active zone.

The boundaries for the active zone, mitochondria, pre- and post synaptic terminals, and the centre of the synaptic vesicles were annotated by hand on the image.

Raw Data

For demonstration and training purposes, the raw data files for this dataset are also provided in the
spatstat.data package installation:

`vesicles.txt`	spatial locations of vesicles
`presynapse.txt`	vertices of `presynapse`
`mitochondria.txt`	vertices of `mitochondria`
`vesiclesimage.tif`	greyscale microscope image
`vesiclesmask.tif`	binary image of mask
`activezone.txt`	coordinates of `activezone`

The files are in the folder `rawdata/vesicles` in the **spatstat.data** installation directory. The precise
location of the files can be obtained using `system.file`, as shown in the examples.

Source

Nicoletta Nava, Mahdieh Khanmohammadi and Jens Randel Nyengaard. Experiment performed
by Nicoletta Nava at the Stereology and Electron Microscopy Laboratory, Aarhus University, Den-
mark. Images were annotated by Mahdieh Khanmohammadi at the Department of Computer Sci-
ence, University of Copenhagen. Jens Randel Nyengaard provided supervision and guidance, and
curated the data.

References

Khanmohammadi, M., Waagepetersen, R., Nava, N., Nyengaard, J.R. and Sporring, J. (2014)
Analysing the distribution of synaptic vesicles using a spatial point process model. *5th ACM Con-
ference on Bioinformatics, Computational Biology and Health Informatics*, Newport Beach, CA,
USA, September 2014.

Examples

```
  if(require(spatstat.geom)) {
plot(vesicles)
with(vesicles.extra,
    plot(activezone, add=TRUE, col="red"))
  }

## read coordinates of vesicles from raw data, for training purposes
vf <- system.file("rawdata", "vesicles", "vesicles.txt",
                  package="spatstat.data")
if(!any(nzchar(vf)))
   stop("Could not find raw data file vesicles.txt")
vdf <- read.table(vf, header=TRUE)
```

waka	*Trees in Waka national park*

Description

This dataset is a spatial point pattern of trees recorded at Waka National Park, Gabon. See Balinga et al (2006).

The dataset waka is a point pattern (object of class ″ppp″) containing the spatial coordinates of each tree, marked by the tree diameter at breast height dbh. The survey region is a 100 by 100 metre squawre. Coordinates are given in metres, while the dbh is in centimetres.

Usage

```
data(waka)
```

Source

Nicolas Picard

References

Balinga, M., Sunderland, T., Walters, G., Issemb\'e, Y., Asaha, S. and Fombod, E. (2006) *A vegetation assessment of the Waka national park, Gabon.* Herbier National du Gabon, LBG, MBG, WCS, FRP and Simthsonian Institution, Libreville, Gabon. CARPE Report, 154 pp. http: //carpe.umd.edu/

Picard, N., Bar-Hen, A., Mortier, F. and Chadoeuf, J. (2009) The multi-scale marked area-interaction point process: a model for the spatial pattern of trees. *Scandinavian Journal of Statistics* **36** 23–41

Examples

```
data(waka)
  if(require(spatstat.geom)) {
plot(waka, markscale=0.01)
title(sub="Tree diameters to scale")
plot(waka, markscale=0.04)
title(sub="Tree diameters 4x scale")
  }
```

---

waterstriders	*Waterstriders data. Three independent replications of a point pattern formed by insects.*

Description

The territorial behaviour of an insect group called waterstriders was studied in a series of laboratory experiments by Dr Matti Nummelin (University of Helskini). The data were analysed in the pioneering PhD thesis of Antti Penttinen (1984).

The dataset `waterstriders` is a list of three point patterns. Each point pattern gives the locations of larvae of the waterstrider *Limnoporus (Gerris) rufoscutellatus* (larval stage V) in a homogeneous area about 48 cm square. The point patterns can be assumed to be independent.

It is known that this species of waterstriders exhibits territorialism at older larvae stages and at the adult stage. Therefore, if any deviation from Complete Spatial Randomness exists in these three point patterns, it is expected to be towards inhibition.

The data were obtained from photographs which were scanned manually. The waterstriders are in a pool which is larger than the picture. A guard area (width about 2.5 cm) has been deleted because it is a source of inhomogeneity to interactions.

Penttinen (1984, chapter 5) fitted a pairwise interaction model with a Strauss/hardcore interaction (see `StraussHard`) with hard core radius 1.5 cm and interaction radius 4.5 cm.

Usage

```
data(waterstriders)
```

Format

`waterstriders` is a list of three point patterns (objects of class `"ppp"`). It also has class `"listof"` so that it can be plotted and printed directly. The point pattern coordinates are in centimetres.

Source

Data were collected by Dr. Matti Nummelin (University of Helsinki, Finland). Data kindly provided by Prof. Antti Penttinen, University of Jyv\"askyl\"a, Finland.

References

Penttinen, A. (1984) Modelling interaction in spatial point patterns: parameter estimation by the maximum likelihood method. *Jyv\"askyl\"a Studies in Computer Science, Economics and Statistics* **7**, University of Jyv\"askyl\"a, Finland.

Index