Package 'ChemometricsWithRData'

January 27, 2015

Type Package

Title Data for package ChemometricsWithR

Version 0.1.3

Date 2012-10-10

Author Ron Wehrens

Maintainer Ron Wehrens <ron.wehrens@gmail.com>

Description The package provides data sets used in the book ``Chemometrics with R - Multivariate Data Analysis in the Natural Sciences and Life Sciences" by Ron Wehrens, Springer (2011).

License GPL (>= 2)

Depends R (>= 2.10)

LazyLoad yes

Repository CRAN

Date/Publication 2012-10-12 16:12:34

NeedsCompilation no

X-CRAN-Original-Maintainer Ron Wehrens <ron.wehrens@fmach.it>

R topics documented:

bdata *HPLC-UV data of two chemical mixtures*

Description

Two chemical mixtures of three compounds have been measured using HPLC-UV. Two of the compounds are known: diazinon and parthion-ethyl, both organophosphorus pesticides. Each data matrix consists of 73 wavelengths and 40 time points. The challenge is to infer the pure spectra of the individual compounds, as well as their time profiles.

Usage

```
data(bdata)
```

Format

A list of four elements. The first two, `d1` and `d2`, are the mixture matrices of the two analytes and one unknown interferent. The last two, `sp1` and `sp2`, contain the pure spectra of the two analytes.

Source

Original matlab data files obtained from http://www.ub.edu/mcr/web_mcr/download_dataHPLC.html (bdataset.zip).

References

R. Wehrens. "Chemometrics with R - Multivariate Data Analysis in the Natural Sciences and Life Sciences". Springer, Heidelberg, 2011.

R. Tauler, S. Lacorte and D. Barcelo. "Application of multivariate curve self-modeling curve resolution for the quantitation of trace levels of organophosphorous pesticides in natural waters from interlaboratory studies". J. of Chromatogr. A, 730, 177-183 (1996).

Examples

```
data(bdata)
persp(bdata$d1, phi = 20, theta = 34, expand = .5,
      xlab = "Time", ylab = "Wavelength")
```

prostate *Averaged prostate data from package msProstate.*

Description

Prostate data, originally from package msProstate, where the three replicates have been averaged. Vector `prostate.type` is a factor, describing the type of sample: a control, a benign prostate enlargement or tumourous sample. Matrix `prostate` contains 327 rows (samples) and 10523 variables. Since package msProstate is orphaned at the end of 2012, the original data are also included here.

Usage

```
data(prostate)
```

Source

R package msProstate (orphaned end 2012).

References

R. Wehrens. "Chemometrics with R - Multivariate Data Analysis in the Natural Sciences and Life Sciences". Springer, Heidelberg, 2011.

See Also

Prostate2000Raw

Examples

```
data(prostate)
mz <- as.numeric(colnames(prostate))
plot(mz, prostate[1,], type = "h", xlab = "m/z", ylab = "Intensity",
     main = "Prostate data")
```

Prostate2000Raw *Prostate Cancer 2000 Raw Spectra*

Description

A data object of class `msSet`, consisting of 654 mass spectra (327 spectra in duplicate) from 2000 to 20000 Da, which were generated from patients with prostate cancer, benign prostatic hypertrophy, and normal controls. These spectra are already baseline corrected and normalized. Please see the references for more details.

Since the original package msProstate is orphaned at the end of 2012, the data are included in the ChemometricsWithR package so that the examples in the book are still executable. This manual page has been adapted to reflect this.

References

B.L. Adam, Y. Qu, J.W. Davis, M.D. Ward, M.A. Clements, L.H. Cazares, O.J. Semmes, P.F. Schellhammer, Y. Yasui, Z. Feng, and G.L. Wright, Jr., "Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men," *Cancer Research*, 62(13):3609–14, 2002.

Y. Qu, B.L. Adam, Y. Yasui, M.D. Ward, L.H. Cazares, P.F. Schellhammer, Z. Feng, O.J. Semmes, and G.L. Wright Jr., "Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients", *Clinical Chemistry*, 48(10):1835–43, 2002.

R. Wehrens, "Chemometrics with R - Multivariate Data Analysis in the Natural Sciences and Life Sciences". Springer, Heidelberg, 2011.

See Also

prostate.

Examples

```
## Examples have been changed from the original man page upon inclusion
## in the ChemometricsWithRData package
data("Prostate2000Raw")

## plot a few spectra, partially
matplot(Prostate2000Raw$mz[1:8000],
        Prostate2000Raw$intensity[1:8000, 1:5], type = "l",
        lty = 1, col = 1:5, xlab = "m/z", ylab = "response")
```

---

shootout *Shootout NIR data*

Description

NIR data from 654 tablets, measured at two different instruments. The data have been divided in training, test and validation sets.

Usage

```
data(shootout)
```

Format

Variable shootout is a list containing spectral data of tablets, measured on two instruments, as well as response variables.

Details

For every tablet, three reponse variables are measured: the amount of active ingredient (nominally 200 mg/tablet), the weight and the hardness. Typically, one wants to estimate the amount of active ingredient from the NIR spectra, a straightforward multivariate calibration problem. The goal of the shootout competition was to find the optimal way to transfer a calibration model of the first instrument to the second.

Source

http://www.idrc-chambersburg.org/shootout2002.html

References

R. Wehrens. "Chemometrics with R - Multivariate Data Analysis in the Natural Sciences and Life Sciences". Springer, Heidelberg, 2011.

Examples

```
data(shootout)
plot(seq(600, 1898, by = 2), shootout$calibrate.1[1,], type = "l",
xlab = "wavelength", ylab = "log(1/T)")
```

---

wines *Wine data*

Description

A data frame containing 177 rows and thirteen columns; object `vintages` contains the class labels. For compatibility with older versions of the package, variable `wine.classes` is retained, too.

These data are the results of chemical analyses of wines grown in the same region in Italy (Piedmont) but derived from three different cultivars: Nebbiolo, Barberas and Grignolino grapes. The wine from the Nebbiolo grape is called Barolo. The data contain the quantities of several constituents found in each of the three types of wines, as well as some spectroscopic variables.

Usage

```
data(wines)
```

Source

http://kdd.ics.uci.edu

References

M. Forina, C. Armanino, M. Castino and M. Ubigli. Vitis, 25:189-201 (1986)

Index