Package ‘mapsRinteractive’

May 13, 2019

Type Package
Title Local Adaptation and Evaluation of Digital Soil Maps
Version 1.0.0
Author Kristin Piikki & Mats Söderström
Maintainer Kristin Piikki <kristin.piikki@slu.se>
Description Local adaptation and evaluation of digital soil maps in raster format by use of point location soil property data.
License MIT + file LICENSE
Encoding UTF-8
LazyData true
RoxygenNote 6.1.1
Imports raster, gstat, rgdal, sp, rgeos
Suggests roxygen2
NeedsCompilation no
Repository CRAN
Date/Publication 2019-05-13 14:50:11 UTC

R topics documented:

check .................................................. 2
CLAYr .................................................. 3
CLAYs .................................................. 3
e ....................................................... 4
evaluate ............................................. 5
even ................................................... 5
kth .................................................... 6
mae .................................................. 7
me .................................................... 8
mri .................................................. 8
odd ................................................ 11
ordkrige ........................................... 12
Description

Checks attributes, geometries and projections of spatial datasets.

Usage

\[
\text{check}(x = \text{NULL}, y = \text{NULL}, z = \text{NULL}, \text{field} = \text{NULL}, \text{edge} = 0, \\
\text{filter} = 1, \text{resolution} = \text{NULL})
\]

Arguments

\begin{itemize}
  \item \textbf{x} \hspace{1cm} \text{Raster dataset. Required. Must be have a defined coordinate system. If the coordinate system is not cartesian, the data will be projected onto the Web Mercator (epsg: 3857) coordinate system before any analyses/tests.}
  \item \textbf{y} \hspace{1cm} \text{SpatialPolygonsDataFrame. Optional. Delineates the area within which the raster layer shall be locally adapted and evaluated. Must be projected. If not projected onto the same coordinate system as \textit{x}, it will be reprojected to the coordinate system of \textit{x}. If not provided, the analyses will be performed within the intersect of the raster and the sampled area.}
  \item \textbf{z} \hspace{1cm} \text{SpatialPointsDataFrame. Required. Must have at least one column with numerical data and these data must be of the same entity unit as the raster (specify this column by argument: \textit{field}). Must be projected. If not projected to the same coordinate system as \textit{x}, it will be reprojected to the coordinate system of \textit{x}.}
  \item \textbf{field} \hspace{1cm} \text{Character value. Required. Name of the column in \textit{y} with the data that shall be used to locally adapt and evaluate the raster}
  \item \textbf{edge} \hspace{1cm} \text{Numeric value. Optional. Specifies the width (m) of a buffer zone inside the edge of the polygon that is excluded from the analyses. Allowed values are within the closed range of 0-10000.}
  \item \textbf{filter} \hspace{1cm} \text{Positive integer. Optional. No of cells in the side of a square window for mean filtering of \textit{x}. Filtering is done before any resampling (see argument: \textit{resolution}). Allowed values are within the closed range of 1-20.}
  \item \textbf{resolution} \hspace{1cm} \text{Positive numeric value. Optional. The resolution (m) to which the imported raster shall be resampled before the adaptation. Allowed values are within the closed range of 0.1-10000. In addition, a resolution that means more than $1E+8$ raster cells is not allowed.}
\end{itemize}
Details

Intended for checking data in functions of mapsRinteractive.

Value

A list with checked and corrected datasets together with a vector of logged feedback.

---

**CLAYr**  
The Digital Soil Map of Sweden -topsoil clay content

**Description**


**Usage**

data(CLAYr)

**Format**

Raster layer

**References**


---

**CLAYs**  
SLU farm (Brogarden) soil sample data -topsoil clay content

**Description**

Projected coordinate system: Sweref99TM (epsg: 3006). Attribute: Lab analyzed topsoil (0-20 cm depth) clay content.

**Usage**

data(CLAYs)

**Format**

SpatialPointsDataFrame
References


Description

Calculates the Nash-Sutcliffe modelling efficiency (E) from observed and predicted values.

Usage

e(\text{observed, predicted})

Arguments

- \text{observed}: Numeric vector of observed values
- \text{predicted}: Numeric vector of predicted values. The length shall be the same as for observed.

Details

\[ E = 1 - \frac{\text{sum(\text{observed} - \text{predicted})}}{\text{sum(\text{observed} - \text{mean(\text{observed})})}} \]

Value

The Nash-Sutcliffe modelling efficiency (E) calculated from observed and predicted values.

References


Examples

\[ o<1:5 \]
\[ p<-c(2,2,4,3,5) \]
\[ e(\text{observed}=o, \text{predicted}=p) \]
evaluate

Description
Computes evaluation measures from observed and predicted data.

Usage
evaluate(df, observed, predicted)

Arguments
- df: Data.frame. Required. A data.frame with observed and predicted data.
- observed: Character value. Required. The name of the column in df with predicted data. The data must be of class numeric.
- predicted: Character value or vector. Required. The names of the column(s) in df with predicted data. The data must be of class numeric.

Value
A data.frame with evaluation statistics. For details, see mri function.

Examples
```r
df<-data.frame(obs=1:9, pred=c(2, 9, 10, 8, 3, 4, 6, 12, 1))
e<-evaluate(df, 'obs', 'pred')
print(e)
```

even

Description
Checks if an integer is even.

Usage
even(x)

Arguments
- x: Integer.
Value
Logical value (TRUE or FALSE). TRUE means that the value is even.

Examples
even(3)

Description
Identification of the kth highest/lowest value(s).

Usage
kth(x = NULL, k = 2, highest = TRUE, index = FALSE,
unique = FALSE, multiple = FALSE)

Arguments
x Numeric vector.
k Positive integer. The order of the value to find. Default = 2, which means that the next highest/lowest values is identified.
highest Logical. TRUE means that the kth highest value(s) is/are identified. FALSE means that the kth lowest value(s) is/are identified. Default = TRUE.
index Logical. TRUE means that the index/indices of the kth highest/lowest value(s) is/are returned. FALSE means that the kth highest/lowest value itself is returned. If ties exist and argument multiple = TRUE, the returned value is a vector, else it is a value. Default = FALSE.
unique Logical. TRUE means that duplicates are removed before the identification of the kth highest/lowest value(s). Default = FALSE
multiple Logical. TRUE means that, If ties exist a vector of all values in x that are equal to the kth highest/lowest values is returned. FALSE means that one random value from the vector of index values is returned. Default = FALSE

Details
NA values are removed.

Value
If index = FALSE: the kth highest/lowest value is returned.
If index = TRUE: the index of the kth highest/lowest value(s) is/are returned.
**Description**

Calculates the mean absolute error (MAE) from observed and predicted values.

**Usage**

```r
mae(observed, predicted)
```

**Arguments**

- `observed`: Numeric vector of observed values
- `predicted`: Numeric vector of predicted values. The length shall be the same as for observed.

**Details**

```r
mae = mean(abs(observed - predicted))
```

**Value**

The mean absolute error (MAE) calculated from the observed and the predicted values.

**Examples**

```r
o <- 1:5
p <- c(2, 2, 4, 3, 5)
mae(observed = o, predicted = p)
```
Description

Calculates the mean error (ME) from observed and predicted values.

Usage

\texttt{me(observed, predicted)}

Arguments

\begin{itemize}
  \item \texttt{observed} \hspace{1cm} Numeric vector of observed values
  \item \texttt{predicted} \hspace{1cm} Numeric vector of predicted values. The length shall be the same as for observed.
\end{itemize}

Details

\[ ME = \text{bias} = \text{mean}(\text{observed} - \text{predicted}) \]

Value

The mean error (ME) calculated from the observed and the predicted values.

Examples

\begin{verbatim}
o<-1:5
p<-c(2,2,4,3,5)
me(observed=o, predicted=p)
\end{verbatim}

Description

Local adaptation and evaluation of digital soil maps in raster format by use of point location soil property data.

Usage

\texttt{mri(x = NULL, y = NULL, z = NULL, field = NULL, edge = 0, filter = 1, resolution = NULL, md = "Sph", rg = NULL, ng = 0.1)}
Arguments

x  
Raster dataset. Required. Must be have a defined coordinate system. If the coordinate system is not cartesian, the data will be projected onto the Web Mercator (epsg: 3857) coordinate system before any analyses/tests.

y  
SpatialPolygonsDataFrame. Optional. Delineates the area within which the raster layer shall be locally adapted and evaluated. Must be projected. If not projected onto the same coordinate system as x, it will be reprojected to the coordinate system of x. If not provided, the analyses will be performed within the intersect of the raster and the sampled area.

z  
SpatialPointsDataFrame. Required. Must have at least one column with numerical data and these must be of the same entity unit as the raster (specify this column by argument: field). Must be projected. If not projected to the same coordinate system as x, it will be reprojected to the coordinate system of x.

field  
Character value. Required. Name of the column in y with the data that shall be used to locally adapt and evaluate the raster

edge  
Numeric value. Optional. Specifies the width (m) of a buffer zone inside the edge of the polygon that is excluded from the analyses. Allowed values are within the closed range of 0-10000.

filter  
Positive integer. Optional. No of cells in the side of a square window for mean filtering of x. Filtering is done before any resampling (see argument: resolution). Allowed values are within the closed range of 1-20.

resolution  
Positive numeric value. Optional. The resolution (m) to which the imported raster shall be resampled before the adaptation. Allowed values are within the closed range of 0.1-10000. In addition, a resolution that means more than 1E+8 raster cells is not allowed.

md  
Character value. Optional. Variogram model type for the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. Default is "Sph" (spherical model).

rg  
Numeric value. Optional. Range of the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. If no rg is specified it will be set to half of the square root of the mapping area: y (possibly shrunk by edge).

ng  
Numeric value. Optional. Nugget of the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. The nugget is expressed as a fraction of the sill. A ng = 0.1 means that the nugget is 10 is by default equal to the variance of the data to be kriged (i.e the point observations or the residuals). Allowed values of ng are within the closed range of 0-1.

Details

The mri function is intended for local adaptation and evaluation of large extent digital soil maps. A raster map and point location soil property are required. A SpatialPolygonsDataFrame can optionally be used to delineate the area for local adaptation and evaluation.

All spatial objects must have defined coordinate systems. If the defined coordinate systems are not the same, the point location data and the polygon data will be transformed to the coordinate system
of the raster. If the defined coordinate system of the raster is not a cartesian coordinate system, all spatial datasets will be projected onto the Web Mercator coordinate system (epsg: 3857).

The four maps are (created and) evaluated are: the original raster map, a map created solely based on the soil samples data (ordinary kriging using a standardized variogram), two maps based on a combination of the raster data and the point observations (regression kriging and residual kriging, both using standardized variograms).

The maps are evaluated by leave-one-out cross validation and a number of evaluation measures are computed: the Nash-Sutcliffe modelling efficiency (E), the mean absolute error (MAE; Janssen & Heuberger, 1995), the coefficient of determination of a linear regression between predicted and measured values (r2).

The mapped area is the intersection between the original raster map (argument: x), any provided SpatialPolygonsDataFrame (argument: y) and the buffered point locations. The buffer width is 1.5*(next largest distance) between one point and its nearest neighbour).

The mapsRInteractive algorithms have been described ad by Piikki et al.(2017) and Nijbroek et al. (2018) (before it was made available as an R package). More details can be found in these publications. It is also implemented in the open Swedish web application for precision agriculture markdata.se and the Sub-Saharan Africa Soil Data Manager.

On error: check that required data are provided (arguments x, y, z and field), check that all spatial datasets (arguments x, y, z) are projected, check that they do overlap and check that the arguments edge, filter and resolution have appropriate values.

Value

A list with:
1) ‘maps’. A raster stack with the original raster map (‘map’), the map, created by ordinary kriging of observed data (‘ordkrig’), by residual kriging (‘reskrig’) and by regression kriging (‘regkrig’).
2) ‘area’. SpatialPolygonsDataFrame with the polygon delineating the mapped area.
3) ‘pts’. SpatialPointsDataFrame with the point locations used for mapping, i.e points falling within the mapped area, excluding points with NA values in the observed values or the values extracted from the original map. The column names mean: obs = observed values. map = original map values. ordkrig_cv = values from the leave-one-out cross validation of the ordinary kriging. res = residuals (map - obs) reskrig_cv = values from the leave-one-out cross validation of the residual kriging. regpred = predicted values from the linear regression (obs = a*map + b) regres = residuals (regpred - obs) regkrig_cv = values from the leave-one-out cross validation of the regression kriging.
4) ‘evaluation’. a data.frame with evaluation statistics for the original map and the leave-one-out cross-validation of the other mapping methods.
5) ‘feedback’ a character vector with logged feedback on inputted and used data.

Author(s)

Kristin Piikki & Mats Söderström, <kristin.piikki@slu.se>

References

Cost Tradeoffs with Respect to Complexity in Otjozondjupa, Namibia. Sustainability, 10(5), 1610. doi:10.3390/su10051610


Examples

```r
# prepare raster dataset (the soil map to be adapted)
data('CLAYr')
CLAYr <- data.frame(CLAYr$POINT_X, CLAYr$POINT_Y, CLAYr$clay_percent) # rearrange columns
require(raster) # load required package
CLAYr <- rasterFromXYZ(CLAYr) # convert to raster
prj <- '+proj=utm +zone=33 +ellps=GRS80 +towgs84=0,0,0,0,0,0,0 +units=m +no_defs' # projection
 CRS(CLAYr) <- crs(prj) # define projection
names(CLAYr) <- 'clay_percent' # rename (not necessary)
# prepare example point location data
data('CLAYS')
CLAYS <- data.frame(CLAYS) # convert to data.frame
coordinates(CLAYS) <-~ POINT_X + POINT_Y # convert to SpatialPointsDataFrame
CRS(CLAYS) <- CRS(CLAYr) # define projection
# run local adaptation and evaluation
mri.out <- mri(x = CLAYr, z = CLAYS, field = 'clay_percent')
# check evaluation measures
print(mri.out$evaluation)
```

odd

Description

Checks if an integer is odd.

Usage

```r
odd(x)
```

Arguments

- **x**  
  Integer.

Value

Logical value (TRUE or FALSE). TRUE means that the value is odd.
Examples

```r
doordkrige(3)
```
**md**  Character value. Optional. Variogram model type for the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by `gstat::vgm`. Default is "Sph" (spherical model).

**rg**  Numeric value. Optional. Range of the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by `gstat::vgm`. If no rg is specified it will be set to half of the square root of the mapping area: \( y \) (possibly shrunk by edge).

**ng**  Numeric value. Optional. Nugget of the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by `gstat::vgm`. The nugget is expressed as a fraction of the sill. A \( ng = 0.1 \) means that the nugget is 10 is by default equal to the variance of the data to be kriged (i.e the point observations or the residuals). Allowed values of ng are within the closed range of 0-1.

**check_data**  Logical. Shall attributes, geometries and projections of the input data (arguments x, y and z) be checked. Default = TRUE.

---

**r2**

**Description**

Calculates the coefficient of determination \((r^2)\) for a linear regression model between predicted values and observed values.

**Usage**

\[
\text{r2(}\text{observed, predicted}\text{)}
\]

**Arguments**

- **observed**  Numeric vector of observed values
- **predicted**  Numeric vector of predicted values. The length shall be the same as for observed.
Value

Coefficient of determination ($r^2$) for a linear regression model between predicted values and observed values.

Examples

```r
o <- 1:5
p <- c(2, 2, 4, 3, 5)
r2( observed = o, predicted = p )
```

Description

Regression kriging using a standardized variogram.

Usage

```r
regkrige(x = NULL, y = NULL, z = NULL, field = NULL, edge = 0,
filter = 1, resolution = NULL, md = "sph", rg = NULL, ng = 0.1,
check_data = T)
```

Arguments

- **x**: Raster dataset. Required. Must be have a defined coordinate system. If the coordinate system is not cartesian, the data will be projected onto the Web Mercator (epsg: 3857) coordinate system before any analyses/tests.
- **y**: SpatialPolygonsDataFrame. Optional. Delineates the area within which the raster layer shall be locally adapted and evaluated. Must be projected. If not projected onto the same coordinate system as x, it will be reprojected to the coordinate system of x. If not provided, the analyses will be performed within the intersect of the raster and the sampled area.
- **z**: SpatialPointsDataFrame. Required. Must have at least one column with numerical data and these data must be of the same entity unit as the raster (specify this column by argument: field). Must be projected. If not projected to the same coordinate system as x, it will be reprojected to the coordinate system of x.
- **field**: Character value. Required. Name of the column in y with the data that shall be used to locally adapt and evaluate the raster.
- **edge**: Numeric value. Optional. Specifies the width (m) of a buffer zone inside the edge of the polygon that is excluded from the analyses. Allowed values are within the closed range of 0-10000.
- **filter**: Positive integer. Optional. No of cells in the side of a square window for mean filtering of x. Filtering is done before any resampling (see argument: resolution). Allowed values are within the closed range of 1-20.
reskrige

**resolution**  Positive numeric value. Optional. The resolution (m) to which the imported raster shall be resampled before the adaptation. Allowed values are within the closed range of 0.1-10000. In addition, a resolution that means more than 1E+8 raster cells is not allowed.

**md**  Character value. Optional. Variogram model type for the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. Default is "Sph" (spherical model).

**rg**  Numeric value. Optional. Range of the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. If no rg is specified it will be set to half of the square root of the mapping area: y (possibly shrunked by edge).

**ng**  Numeric value. Optional. Nugget of the standardized variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. The nugget is expressed as a fraction of the sill. A ng = 0.1 means that the nugget is 10 is by default equal to the variance of the data to be kriged (i.e the point observations or the residuals). Allowed values of ng are within the closed range of 0-1.

**check_data**  Logical. Shall attributes, geometries and projections of the input data (arguments x, y and z) be checked. Default = TRUE.

**Details**

This is the regression kriging function called by the mri function. For details, see documentation of the mri function.

**Value**

A list with 1) a raster layer with predicted values and 2) a SpatialPolygonsDataFrame with cross-validation data. For details, see mri function.

**Description**

Regression kriging using a standardized variogram.

**Usage**

```r
gskrig(x = NULL, y = NULL, z = NULL, field = NULL, edge = 0,
      filter = 1, resolution = NULL, md = "Sph", rg = NULL, ng = 0.1,
      check_data = T)
```
Arguments

**x**
- Raster dataset. Required. Must be have a defined coordinate system. If the coordinate system is not cartesian, the data will be projected onto the Web Mercator (epsg: 3857) coordinate system before any analyses/tests.

**y**
- SpatialPolygonsDataFrame. Optional. Delineates the area within which the raster layer shall be locally adapted and evaluated. Must be projected. If not projected onto the same coordinate system as x, it will be reprojected to the coordinate system of x. If not provided, the analyses will be performed within the intersect of the raster and the sampled area.

**z**
- SpatialPointsDataFrame. Required. Must have at least one column with numerical data and these data must be of the same entity unit as the raster (specify this column by argument: field). Must be projected. If not projected to the same coordinate system as x, it will be reprojected to the coordinate system of x.

**field**
- Character value. Required. Name of the column in y with the data that shall be used to locally adapt and evaluate the raster.

**edge**
- Numeric value. Optional. Specifies the width (m) of a buffer zone inside the edge of the polygon that is excluded from the analyses. Allowed values are within the closed range of 0-10000.

**filter**
- Positive integer. Optional. No of cells in the side of a square window for mean filtering of x. Filtering is done before any resampling (see argument: resolution). Allowed values are within the closed range of 1-20.

**resolution**
- Positive numeric value. Optional. The resolution (m) to which the imported raster shall be resampled before the adaptation. Allowed values are within the closed range of 0.1-10000. In addition, a resolution that means more than 1E+8 raster cells is not allowed.

**md**
- Character value. Optional. Variogram model type for the standardzed variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. Default is "Sph" (spherical model).

**rg**
- Numeric value. Optional. Range of the standardzed variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. If no rg is specified it will be set to half of the square root of the mapping area: y (possibly shrinked by edge).

**ng**
- Numeric value. Optional. Nugget of the standardzed variograms used for ordinary kriging interpolation of observed data or residuals. Variograms are generated by gstat::vgm. The nugget is expressed as a fraction of the sill. A ng = 0.1 means that the nugget is 10 is by default equal to the variance of the data to be kriged (i.e the point observations or the residuals). Allowed values of ng are within the closed range of 0-1.

**check_data**
- Logical. Shall attributes, geometries and projections of the input data (arguments x, y and z) be checked. Default = TRUE.

Details

This is the residual kriging function called by the mri function. For details, see documentation of the mri function.
### Description

Calculates the root mean square error (RMSE) from observed and predicted values.

### Usage

```r
rmse( observed, predicted )
```

### Arguments

- `observed` Numeric vector of observed values
- `predicted` Numeric vector of predicted values. The length shall be the same as for observed.

### Details

\[
\text{rmse} = \sqrt{\text{mean}(\text{observed} - \text{predicted})^2)}
\]

### Value

The root mean square error (RMSE) calculated from the observed and the predicted values.

### Examples

```r
o <- 1:5
p <- c(2, 2, 4, 3, 5)
rmse( observed = o, predicted = p )
```
Description
Create a `SpatialPolygonsDataFrame` from extent of a spatial object.

Usage

```r
gpd_from_extent(x)
```

Arguments

- `x`: A spatial object.

Details

If `x` is projected, the `SpatialPolygonsDataFrame` will also be projected.

Value

`SpatialPolygonsDataFrame`.

Examples

```r
require(raster) # load required package.
r1 <- raster::raster(extent(c(0,10,0,10)), res=1, vals=1:100) # create example raster.
spdf <- gpd_from_extent(r1) # convert the raster extent to SpatialPolygonsDataFrame.
plot(spdf) # Plot results.
```
Index

*Topic Internal
  check, 2
*Topic datasets
  CLAYr, 3
  CLAYs, 3
  check, 2
  CLAYr, 3
  CLAYs, 3
e, 4
evaluate, 5
even, 5
kth, 6
mae, 7
me, 8
mri, 8
odd, 11
ordkrige, 12
r2, 13
regkrige, 14
reskrige, 15
rmse, 17
spdf_from_extent, 18