

Package ‘mpspline2’

April 3, 2022

Type Package

Title Mass-Preserving Spline Functions for Soil Data

Version 0.1.6

Date 2022-04-03

Description

A low-dependency implementation of `GSIF::mpspline()` <<https://r-forge.r-project.org/scm/viewvc.php/pkg/R/mpspline.R?view=markup&revision=240&root=gsif>>, which applies a mass-preserving spline to soil attributes. Splining soil data is a safe way to make continuous down-profile estimates of attributes measured over discrete, often discontinuous depth intervals.

License GPL

Encoding UTF-8

Imports stats

Suggests testthat, covr

RoxygenNote 7.1.2

NeedsCompilation no

Author Lauren O'Brien [aut, cre] (<<https://orcid.org/0000-0002-7336-2171>>),
Brendan Malone [ctb] (<<https://orcid.org/0000-0002-0473-8518>>),
Tomislav Hengl [ctb] (<<https://orcid.org/0000-0002-9921-5129>>),
Tom Bishop [ctb],
David Rossiter [ctb],
Dylan Beaudette [ctb],
Andrew Brown [ctb]

Maintainer Lauren O'Brien <obr1soilau@gmail.com>

Repository CRAN

Date/Publication 2022-04-03 19:20:04 UTC

R topics documented:

<code>mpspline</code>	2
<code>mpspline_compact</code>	3
<code>mpspline_one</code>	4
<code>mpspline_tidy</code>	5

mpspline	<i>Spline discrete soils data - multiple sites</i>
----------	--

Description

This function implements the mass-preserving spline method of Bishop *et al* (1999) (doi: [10.1016/S00167061\(99\)000038](https://doi.org/10.1016/S00167061(99)000038)) for interpolating between measured soil attributes down a soil profile, across multiple sites' worth of data.

Usage

```
mpspline(
  obj = NULL,
  var_name = NULL,
  lam = 0.1,
  d = c(0, 5, 15, 30, 60, 100, 200),
  vlow = 0,
  vhigh = 1000
)
```

Arguments

obj	data.frame or matrix. Column 1 must contain site identifiers. Columns 2 and 3 must contain upper and lower sample depths, respectively. Subsequent columns will contain measured values for those depths.
var_name	length-1 character or length-1 integer denoting the column in obj in which target data is stored. If not supplied, the fourth column of the input object is assumed to contain the target data.
lam	number; smoothing parameter for spline. Defaults to 0.1.
d	sequential integer vector; denotes the output depth ranges in cm. Defaults to <code>c(0, 5, 15, 30, 60, 100, 200)</code> after the GlobalSoilMap specification, giving output predictions over intervals 0-5cm, 5-15cm, etc.
vlow	numeric; constrains the minimum predicted value to a realistic number. Defaults to 0.
vhigh	numeric; constrains the maximum predicted value to a realistic number. Defaults to 1000.

Value

A nested list of data for each input site. List elements are: Site ID, vector of predicted values over input intervals, vector of predicted values for each cm down the profile to `max(d)`, vector of predicted values over `d` (output) intervals, and root mean squared error.

Examples

```
dat <- data.frame("SID" = c( 1,  1,  1,  1,  2,  2,  2,  2),
                 "UD" = c( 0, 20, 40, 60,  0, 15, 45, 80),
                 "LD" = c(10, 30, 50, 70,  5, 30, 60, 100),
                 "VAL" = c( 6,  4,  3, 10, 0.1, 0.9, 2.5,  6),
                 stringsAsFactors = FALSE)
m1 <- mpspline(obj = dat, var_name = 'VAL')
```

mpspline_compact

Spline discrete soils data - multiple sites, compact output

Description

This function implements the mass-preserving spline method of Bishop *et al* (1999) (doi: [10.1016/S00167061\(99\)000038](https://doi.org/10.1016/S00167061(99)000038)) for interpolating between measured soil attributes down a soil profile, across multiple sites' worth of data. It returns a more compact output object than `mpspline()`.

Usage

```
mpspline_compact(
  obj = NULL,
  var_name = NULL,
  lam = 0.1,
  d = c(0, 5, 15, 30, 60, 100, 200),
  vlow = 0,
  vhigh = 1000
)
```

Arguments

<code>obj</code>	data.frame or matrix. Column 1 must contain site identifiers. Columns 2 and 3 must contain upper and lower sample depths, respectively. Subsequent columns will contain measured values for those depths.
<code>var_name</code>	length-1 character or length-1 integer denoting the column in <code>obj</code> in which target data is stored. If not supplied, the fourth column of the input object is assumed to contain the target data.
<code>lam</code>	number; smoothing parameter for spline. Defaults to 0.1.
<code>d</code>	sequential integer vector; denotes the output depth ranges in cm. Defaults to <code>c(0, 5, 15, 30, 60, 100, 200)</code> after the GlobalSoilMap specification, giving output predictions over intervals 0-5cm, 5-15cm, etc.
<code>vlow</code>	numeric; constrains the minimum predicted value to a realistic number. Defaults to 0.
<code>vhigh</code>	numeric; constrains the maximum predicted value to a realistic number. Defaults to 1000.

Value

A four-item list containing a matrix of predicted values over the input depth ranges, a matrix of predicted values over the output depth ranges, a matrix of 1cm predictions, and a matrix of RMSE and IQR-scaled RMSE values. Site identifiers are in rownames attributes.

Examples

```
dat <- data.frame("SID" = c( 1, 1, 1, 1, 2, 2, 2, 2),
                 "UD" = c( 0, 20, 40, 60, 0, 15, 45, 80),
                 "LD" = c(10, 30, 50, 70, 5, 30, 60, 100),
                 "VAL" = c( 6, 4, 3, 10, 0.1, 0.9, 2.5, 6),
                 stringsAsFactors = FALSE)
mpspline_compact(obj = dat, var_name = 'VAL')
```

mpspline_one

Spline discrete soils data - single site

Description

This function implements the mass-preserving spline method of Bishop *et al* (1999) (doi: [10.1016/S00167061\(99\)000038](https://doi.org/10.1016/S00167061(99)000038)) for interpolating between measured soil attributes down a single soil profile.

Usage

```
mpspline_one(
  site = NULL,
  var_name = NULL,
  lam = 0.1,
  d = c(0, 5, 15, 30, 60, 100, 200),
  vlow = 0,
  vhigh = 1000
)
```

Arguments

site	data frame containing data for a single soil profile. Column 1 must contain site identifiers. Columns 2 and 3 must contain upper and lower sample depths, respectively, measured in centimeters. Subsequent columns will contain measured values for those depths.
var_name	length-1 character or length-1 integer denoting the column in site in which target data is stored. If not supplied, the fourth column of the input object is assumed to contain the target data.
lam	number; smoothing parameter for spline. Defaults to 0.1.
d	sequential integer vector; denotes the output depth ranges in cm. Defaults to c(0, 5, 15, 30, 60, 100, 200) after the GlobalSoilMap specification, giving output predictions over intervals 0-5cm, 5-15cm, etc.

vlow	numeric; constrains the minimum predicted value to a realistic number. Defaults to 0.
vhigh	numeric; constrains the maximum predicted value to a realistic number. Defaults to 1000.

Value

A list with the following elements: Site ID, vector of predicted values over input intervals, vector of predicted values for each cm down the profile to $\max(d)$, vector of predicted values over d (output) intervals, and root mean squared error.

Examples

```
dat <- data.frame("SID" = c( 1,  1,  1,  1),
                 "UD"  = c( 0, 20, 40, 60),
                 "LD"  = c(10, 30, 50, 70),
                 "VAL" = c( 6,  4,  3, 10),
                 stringsAsFactors = FALSE)
mpspline_one(site = dat, var_name = 'VAL')
```

mpspline_tidy *Spline discrete soils data - multiple sites, tidy output*

Description

This function implements the mass-preserving spline method of Bishop *et al* (1999) (doi: [10.1016/S00167061\(99\)000038](https://doi.org/10.1016/S00167061(99)000038)) for interpolating between measured soil attributes down a soil profile, across multiple sites' worth of data. It returns an output object with tidy data formatting.

Usage

```
mpspline_tidy(
  obj = NULL,
  var_name = NULL,
  lam = 0.1,
  d = c(0, 5, 15, 30, 60, 100, 200),
  vlow = 0,
  vhigh = 1000
)
```

Arguments

obj	data.frame or matrix. Column 1 must contain site identifiers. Columns 2 and 3 must contain upper and lower sample depths, respectively, and be measured in centimeters. Subsequent columns will contain measured values for those depths.
var_name	length-1 character or length-1 integer denoting the column in obj in which target data is stored. If not supplied, the fourth column of the input object is assumed to contain the target data.

<code>lam</code>	number; smoothing parameter for spline. Defaults to 0.1.
<code>d</code>	sequential integer vector; denotes the output depth ranges in cm. Defaults to <code>c(0, 5, 15, 30, 60, 100, 200)</code> after the <code>GlobalSoilMap</code> specification, giving output predictions over intervals 0-5cm, 5-15cm, etc.
<code>vlow</code>	numeric; constrains the minimum predicted value to a realistic number. Defaults to 0.
<code>vhigh</code>	numeric; constrains the maximum predicted value to a realistic number. Defaults to 1000.

Value

A four-item list containing data frames of predicted values over the input depth ranges, the output depth ranges, 1cm-increment predictions, and RMSE and IQR-scaled RMSE values.

Examples

```
dat <- data.frame("SID" = c( 1,  1,  1,  1,  2,  2,  2,  2),
                 "UD"  = c( 0, 20, 40, 60,  0, 15, 45, 80),
                 "LD"  = c(10, 30, 50, 70,  5, 30, 60, 100),
                 "VAL" = c( 6,  4,  3, 10, 0.1, 0.9, 2.5,  6),
                 stringsAsFactors = FALSE)
mpspline_tidy(obj = dat, var_name = 'VAL')
```

Index

`mpspline`, 2
`mpspline()`, 3
`mpspline_compact`, 3
`mpspline_one`, 4
`mpspline_tidy`, 5