

# Package ‘mpspline2’

November 23, 2025

**Type** Package

**Title** Mass-Preserving Spline Functions for Soil Data

**Version** 0.1.9

**Date** 2025-11-20

**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** 7.3.3

**URL** <https://github.com/obrl-soil/mpspline2>

**NeedsCompilation** no

**Author** Lauren O'Brien [aut, cre] (ORCID:

<<https://orcid.org/0000-0002-7336-2171>>),

Brendan Malone [ctb] (ORCID: <<https://orcid.org/0000-0002-0473-8518>>),

Tomislav Hengl [ctb] (ORCID: <<https://orcid.org/0000-0002-9921-5129>>),

Tom Bishop [ctb],

David Rossiter [ctb],

Dylan Beaudette [ctb],

Andrew Brown [ctb] (ORCID: <<https://orcid.org/0000-0002-4565-533X>>)

**Maintainer** Lauren O'Brien <obrlsoilau@gmail.com>

**Repository** CRAN

**Date/Publication** 2025-11-22 23:50:02 UTC

## Contents

mpspline . . . . .	2
mpspline_compact . . . . .	3
mpspline_one . . . . .	5
<b>Index</b>	<b>7</b>

---

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	character or integer vector denoting the column(s) 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 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 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**

```
# single variable
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')

# multiple variables
dat_multi <- 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),
                       "VAL1" = c( 6, 4, 3, 10, 0.1, 0.9, 2.5, 6),
                       "VAL2" = c( 5, 3, 2, 9, 0.2, 1.0, 2.0, 5),
                       stringsAsFactors = FALSE)
m_multi <- mpspline(obj = dat_multi, var_name = c('VAL1', 'VAL2'))
# access results: m_multi[['VAL1']][['1']]
```

---

mpspline\_compact

*Spline discrete soils data - multiple sites*


---

**Description**

These functions implement 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. `mpspline_compact()` returns results as matrices while `mpspline_tidy()` returns results as data frames.

**Usage**

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

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, measured in centimeters. Subsequent columns will contain measured values for those depths.
var_name	character or integer vector denoting the column(s) 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 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

mpspline\_compact() returns a four-item list containing matrices: predicted values over input depth ranges, output depth ranges, 1cm predictions, and RMSE values. Site identifiers are in row-names.

mpspline\_tidy() returns a four-item list of data frames with the same predictions but in tidy format, with an added VARIABLE column when processing multiple variables.

### 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)

# single variable
result <- mpspline_compact(obj = dat, var_name = 'VAL')

# multiple variables
dat_multi <- 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),
                       "VAL1" = c( 6,  4,  3, 10, 0.1, 0.9, 2.5,  6),
                       "VAL2" = c( 5,  3,  2,  9, 0.2, 1.0, 2.0,  5),
                       stringsAsFactors = FALSE)

```

```

result_multi <- mpspline_compact(obj = dat_multi, var_name = c('VAL1', 'VAL2'))

# Single variable with tidy output
result <- mpspline_tidy(obj = dat, var_name = 'VAL')

# Multiple variables
dat_multi <- 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),
                        "VAL1" = c( 6, 4, 3, 10, 0.1, 0.9, 2.5, 6),
                        "VAL2" = c( 5, 3, 2, 9, 0.2, 1.0, 2.0, 5),
                        stringsAsFactors = FALSE)
result_multi <- mpspline_tidy(obj = dat_multi, var_name = c('VAL1', 'VAL2'))
subset(result_multi$est_dcm, VARIABLE == 'VAL1')

```

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	character or integer vector denoting the column(s) 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.

<code>vlow</code>	numeric; constrains the minimum predicted value to a realistic number. Defaults to 0.
<code>whigh</code>	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')
```

# Index

`mpspline`, [2](#)  
`mpspline_compact`, [3](#)  
`mpspline_one`, [5](#)  
`mpspline_tidy` (`mpspline_compact`), [3](#)