# Package 'BLA'

January 20, 2025

**Type** Package

**Title** Boundary Line Analysis

**Version** 1.0.1

**Description** Fits boundary line models to datasets as proposed by Webb (1972) <doi:10.1080/00221589.1972.11514472> and makes statistical inferences about their parameters. Provides additional tools for testing datasets for evidence of boundary presence and selecting initial starting values for model optimization prior to fitting the boundary line models. It also includes tools for conducting post-hoc analyses such as predicting boundary values and identifying the most limiting factor (Miti, Milne, Giller, Lark (2024) <doi:10.1016/j.fcr.2024.109365>). This ensures a comprehensive analysis for datasets that exhibit upper boundary structures.

**License** GPL (>= 3)

**URL** https://chawezimiti.github.io/BLA/

**BugReports** https://github.com/chawezimiti/BLA/issues

**Depends** R (>= 4.0)

**Imports** data.table, MASS, mvtnorm, numDeriv, stats

**Suggests** knitr, rmarkdown, testthat (>= 3.0.0)

**VignetteBuilder** knitr

**Config/testthat/edition** 3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.3.1

**NeedsCompilation** no

**Author** Chawezi Miti [cre, aut, cph] (<https://orcid.org/0000-0001-7776-8366>),
Richard M Lark [aut] (<https://orcid.org/0000-0003-2571-8521>),
Alice E Milne [aut] (<https://orcid.org/0000-0002-4509-0578>),
Ken E Giller [aut] (<https://orcid.org/0000-0002-5998-4652>),
Victor O Sadras [ctb],
University of Nottingham/Rothamsted Research [fnd]

**Maintainer** Chawezi Miti <chawezi.miti@nottingham.ac.uk>

**Repository** CRAN

**Date/Publication** 2024-05-28 16:50:02 UTC

# Contents



---

`blbin` *Binning method for determining the boundary line model*

---

### Description

This function fits a boundary model to the upper bounds of a scatter plot of x and y based on the binning method. The data are first divided into equal sized sections in the x-axis and a boundary point in each section is selected based on a set criteria (e.g. 0.90, 0.95 or 0.99 percentile of y among other criteria). A model is then fitted to the resulting boundary points by the least squares method. This is done using optimization procedure and hence requires some starting guess parameters for the proposed model.

### Usage

```
blbin(x,y,bins, model="explore", equation=NULL, start, tau=0.95,
     optim.method="Nelder-Mead", xmin=min(bound$x), xmax=max(bound$x),plot=TRUE,
       bp_col="red", bp_pch=16, bl_col="red", lwd=1,line_smooth=1000,...)
```

**Arguments**

x
: A numeric vector of values for the independent variable.

y
: A numeric vector of values for the response variable.

bins
: A numeric vector of length 3 or 4 that determines the size of sections. The first and second values give the range of the data to be binned while the third and fourth values give the width of the bins and the step size respectively. If only three values are provided, the step size is assumed to be equal to bin width.

model
: Selects the functional form of the boundary line. It includes "explore" as default, "blm" for linear model, "lp" for linear plateau model, "mit" for the Mitscherlich model, "schmidt" for the Schmidt model, "logistic" for logistic model, "logisticND" for logistic model proposed by Nelder (1961), "inv-logistic" for the inverse logistic model, "double-logistic" for the double logistic model, "qd" for quadratic model and the "trapezium" for the trapezium model. The "explore" is used to check the position of boundary points in each bin so that the correct model can be applied. For custom models, set model = "other".

equation
: A custom model function writen in the form of an R function. Applies only when argument model="other", else it is NULL.

start
: A numeric vector of initial starting values for optimization in fitting the boundary model. Its length and arrangement depend on the suggested model:

  - For the "blm" model, it is a vector of length 2 arranged as intercept and slope.
  - For the "lp" model, it is a vector of length 3 arranged as intercept, slope and maximum response.
  - For the "logistic" and "inv-logistic" models, it is a vector of length 3 arranged as the scaling parameter, shape parameter and maximum response.
  - For the "logisticND" model proposed by Nelder (1961), it is a vector of length 3 arranged as the scaling parameter, shape parameter and maximum response.
  - For the "double-logistic" model, it is a vector of length 6 arranged as the scaling parameter one, shape parameter one, maximum response, maximum response, scaling parameter two and shape parameter two.
  - For the "qd" model, it is a vector of length 3 arranged as constant, linear coefficient and quadratic coefficient.
  - For the "trapezium" model, it is a vector of length 3 arranged as intercept one, slope one, maximum response, intercept two and slope two.
  - For the "mit" model, it is a vector of length 3 arranged as the intercept, shape parameter and the maximum response.
  - For the "schmidt" model, it is a vector of length 3 arranged as scaling parameter, shape parameter (x-value at maximum response ) and maximum response.

tau
: A percentile value (0-1) that represents the boundary point within each bin (default is tau = 0.95).

| | |
|---|---|
| optim.method | Describes the method used to optimize the model as in the optim() function. The methods include "Nelder-Mead", "BFGS", "CG", "L-BFGS-B", "SANN" and "Brent". |
| xmin | Numeric value that describes the minimum x value to which the boundary line is to be fitted (default is min(x)). |
| xmax | A numeric value that describes the maximum x value to which the boundary line is to be fitted (default is max(x)). xmin and xmax determine the subset of the data set used to fit boundary model. |
| plot | If TRUE, a plot is part of the output. If FALSE, plot is not part of output (default is TRUE). |
| bp_col | Selects the color of the boundary points. |
| bp_pch | Point character as pch of the plot() function. It controls the shape of the boundary points on plot (bp_pch = 16 as default). |
| bl_col | Colour of the boundary line. |
| lwd | Determines the thickness of the boundary line on the plot (default is 1). |
| line_smooth | Parameter that describes the smoothness of the boundary line. (default is 1000). The higher the value, the smoother the line. |
| ... | Additional graphical parameters as in the par() function. |

## Details

Some inbuilt models are available for the blbin() function. The "explore" option for the argument model generates a plot showing the location of the boundary points selected by the binning procedure. This helps to identify which model type is suitable to fit as a boundary line. The suggest model forms are as follows:

1. Linear model ("blm")
$$y = \beta_1 + \beta_2 x$$
where $\beta_1$ is the intercept and $\beta_2$ is the slope.
2. Linear plateau model ("lp")
$$y = \min(\beta_1 + \beta_2 x, \beta_0)$$
where $\beta_1$ is the intercept , $\beta_2$ is the slope and $\beta_0$ is the maximum response.
3. The logistic ("logistic") and inverse logistic ("inv-logistic") models
$$y = \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
$$y = \beta_0 - \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
where $\beta_1$ is a scaling parameter , $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
4. Logistic model ("logisticND") (Nelder (2009))
$$y = \frac{\beta_0}{1 + (\beta_1 \times e^{-\beta_2 x})}$$
where $\beta_1$ is a scaling parameter, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

5. Double logistic model (`"double-logistic"`)

$$y = \frac{\beta_{0,1}}{1 + e^{\beta_2(\beta_1 - x)}} - \frac{\beta_{0,2}}{1 + e^{\beta_4(\beta_3 - x)}}$$

   where $\beta_1$ is a scaling parameter one, $\beta_2$ is shape parameter one, $\beta_{0,1}$ and $\beta_{0,2}$ are the maximum response , $\beta_3$ is a scaling parameter two and $\beta_4$ is a shape parameter two.

6. Quadratic model (`"qd"`)

$$y = \beta_1 + \beta_2 x + \beta_3 x^2$$

   where $\beta_1$ is a constant, $\beta_2$ is a linear coefficient and $\beta_3$ is the quadratic coefficient.

7. Trapezium model (`"trapezium"`)

$$y = \min(\beta_1 + \beta_2 x, \beta_0, \beta_3 + \beta_4 x)$$

   where $\beta_1$ is the intercept one, $\beta_2$ is the slope one, $\beta_0$ is the maximum response, $\beta_3$ is the intercept two and $\beta_3$ is the slope two.

8. Mitscherlich model (`"mit"`)

$$y = \beta_0 - \beta_1 * \beta_2^x$$

   where $\beta_1$ is the intercept, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

9. Schmidt model (`"schmidt"`)

$$y = \beta_0 + \beta_1(x - \beta_2)^2$$

   where $\beta_1$ is ascaling parameter, $\beta_2$ is a shape parameter (x-value at maximum response ) and $\beta_0$ is the maximum response .

10. Custom model ("other") This option allows you to create your own model form using the function `function()`. The custom model should be assigned to the argument `equation`. Note that the parameters for the custom model should be `a` and `b` for a two parameter model; `a`, `b` and `c` for a three parameter model; `a`, `b`, `c` and `d` for a four parameter model and so on.

The function `blbin()` utilities the optimization procedure of the `optim()` function to determine the model parameters. There is a tendency for optimization algorithms to settle at a local optimum. To remove the risk of settling for local optimum parameters, it is advised that the function is run using several starting values and the results with the smallest error (residue mean square) can be taken as a representation of the global optimum.

The common errors encountered due to poor start values

1. function cannot be evaluated at initial parameters
2. initial value in 'vmmin' is not finite

### Value

A list of length 5 consisting of the fitted model, equation form, parameters of the boundary line, the residue mean square and the boundary points. Additionally, a graphical representation of the boundary line on the scatter plot is produced.

### Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

## References


Casanova, D., Goudriaan, J., Bouma, J., & Epema, G. (1999). Yield gap analysis in relation to soil properties in direct-seeded flooded rice.

Nelder, J.A. 1961. The fitting of a generalization of the logistic curve. Biometrics 17: 89–110.

Phillips, B.F. & Campbell, N.A. 1968. A new method of fitting the von Bertelanffy growth curve using data on the whelk. Dicathais, Growth 32: 317–329.

Schmidt, U., Thöni, H., & Kaupenjohann, M. (2000). Using a boundary line approach to analyze N2O flux data from agricultural soils. Nutrient Cycling in Agro-ecosystems, 57, 119-129.


## Examples

```
x<-log(SoilP$P)
y<-SoilP$yield
start<-c(4,3,13.6, 35, -5)
bins<-c(1.6,4.74,0.314)

blbin(x,y, bins=bins, start=start,model = "trapezium", tau=0.99,
       xlab=expression("Phosphorus/ln(mg L"^-1*")"),
       ylab=expression("Yield/ t ha"^-1), pch=16,
       col="grey", bp_col="grey")
```

---

`ble_profile` *Likelihood profile for various measurement error values*

---

## Description

Estimates the standard deviation of measurement error (`sign`) of the response variable, an input of the `cbvn()` function, when a measured value is not available (Lark & Milne, 2016). `sigh` is fixed at each of a set of values in turn, and remaining parameters are estimated conditional on `sigh` by maximum likelihood. The maximized likelihoods for the sequence of values constitutes a likelihood profile. The value of `sigh` where the profile is maximized is selected.

## Usage

```
ble_profile(data, sigh, model="lp", equation=NULL,  start, UpLo="U",
            optim.method="BFGS", plot=TRUE)
```

## Arguments

| | |
|---|---|
| `data` | A dataframe with two numeric columns, independent (x) and dependent (y) variables respectively. |
| `sigh` | A vector of the suggested standard deviations of the measurement error values. |

model
: Selects the functional form of the boundary line. It includes "blm" for linear model, "lp" for linear plateau model, "mit" for the Mitscherlich model, "schmidt" for the Schmidt model, "logistic" for logistic model, "logisticND" for logistic model proposed by Nelder (1961), "inv-logistic" for the inverse logistic model, "double-logistic" for the double logistic model, "qd" for quadratic model and the "trapezium" for the trapezium model. For custom models, set model = "other".

equation
: A custom model function writen in the form of an R function. Applies only when argument model="other", else it is NULL.

start
: A numeric vector of initial starting values for optimization in fitting the boundary model. Its length and arrangement depend on the suggested model:

- For the "blm" model, it is a vector of length 7 arranged as the intercept, the slope, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "lp" model, it is a vector of length 8 arranged as the intercept, the slope, the maximum or plateau response, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "mit" model, it is a vector of length 8 arranged as the intercept, shape parameter, the maximum or plateau response, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "logistic", "inv-logistic" and "logisticND" models, it is a vector of length 8 arranged as scaling parameter, shape parameter, the maximum or plateau value, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "double-logistic" model, it is a vector of length 11 arranged as scaling parameter, shape parameter, maximum response, maximum response, scaling parameter two, shape parameter two, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "trapezium" model, it is a vector of length 10 arranged as intercept one, slope one, maximum response, intercept two, slope two, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "qd" model, it is a vector of length 8 arranged as a constant, linear coefficient, quadratic coefficient,mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "schmidt" model, it is a vector of length 8 arranged the scaling parameter, shape parameter (x-value at maximum response ), maximum response, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.

UpLo
: Selects the type of boundary. "U" fits the upper boundary and "L" fits the lower boundary.

optim.method
: Describes the method used to optimize the model as in the optim() function. The methods include "Nelder-Mead", "BFGS", "CG", "L-BFGS-B", "SANN" and "Brent".

plot If `TRUE`, a plot is part of the output. If `FALSE`, plot is not part of output (default is `TRUE`).

**Details**

Some inbuilt models are available for the `cbvn()` function. The suggest model forms are as follows:

1. Linear model (`"blm"`)
$$y = \beta_1 + \beta_2 x$$
where $\beta_1$ is the intercept and $\beta_2$ is the slope.
2. Linear plateau model (`"lp"`)
$$y = \min(\beta_1 + \beta_2 x, \beta_0)$$
where $\beta_1$ is the intercept , $\beta_2$ is the slope and $\beta_0$ is the maximum response.
3. The logistic (`"logistic"`) and inverse logistic (`"inv-logistic"`) models
$$y = \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
$$y = \beta_0 - \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
where $\beta_1$ is a scaling parameter , $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
4. Logistic model (`"logisticND"`) (Nelder (1961))
$$y = \frac{\beta_0}{1 + (\beta_1 \times e^{-\beta_2 x})}$$
where $\beta_1$ is a scaling parameter, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
5. Double logistic model (`"double-logistic"`)
$$y = \frac{\beta_{0,1}}{1 + e^{\beta_2(\beta_1 - x)}} - \frac{\beta_{0,2}}{1 + e^{\beta_4(\beta_3 - x)}}$$
where $\beta_1$ is a scaling parameter one, $\beta_2$ is shape parameter one, $\beta_{0,1}$ and $\beta_{0,2}$ are the maximum response , $\beta_3$ is a scaling parameter two and $\beta_4$ is a shape parameter two.
6. Quadratic model (`"qd"`)
$$y = \beta_1 + \beta_2 x + \beta_3 x^2$$
where $\beta_1$ is a constant, $\beta_2$ is a linear coefficient and $\beta_3$ is the quadratic coefficient.
7. Trapezium model (`"trapezium"`)
$$y = \min(\beta_1 + \beta_2 x, \beta_0, \beta_3 + \beta_4 x)$$
where $\beta_1$ is the intercept one, $\beta_2$ is the slope one, $\beta_0$ is the maximum response, $\beta_3$ is the intercept two and $\beta_3$ is the slope two.
8. Mitscherlich model (`"mit"`)
$$y = \beta_0 - \beta_1 * \beta_2^x$$
where $\beta_1$ is the intercept, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

9. Schmidt model ("schmidt")

$$y = \beta_0 + \beta_1(x - \beta_2)^2$$

where $\beta_1$ is ascaling parameter, $\beta_2$ is a shape parameter (x-value at maximum response ) and $\beta_0$ is the maximum response .

The function `ble_profile()` utilities the optimization procedure of the `optim()` function to determine the model parameters. There is a tendency for optimization algorithms to settle at a local optimum. To remove the risk of settling for local optimum parameters, it is advised that the function is run using several starting values and the results with the largest likelihood can be taken as a representation of the global optimum.

The common errors encountered due to poor start values

1. function cannot be evaluated at initial parameters
2. initial value in 'vmmin' is not finite

## Value

A list of length 2 containing the suggested standard deviations of measurement error values and the corresponding log-likelihood values. additionally, a likelihood profile plot (log-likelihood against the standard deviation of measurement error) is produced.

## Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

## References

Lark, R. M., & Milne, A. E. (2016). Boundary line analysis of the effect of water filled pore space on nitrous oxide emission from cores of arable soil. European Journal of Soil Science, 67 , 148-159.

Nelder, J.A. 1961. The fitting of a generalization of the logistic curve. Biometrics 17: 89–110.

## Examples

```
x<-evapotranspiration$`ET(mm)`
y<-evapotranspiration$`yield(t/ha)`
data<-data.frame(x,y)
start<-c(0.5,0.02,289.6,2.4,83.7,1.07,0.287)
sigh <- c(0.6,0.7,0.8,0.9)

ble_profile(data,start=start,model = "blm", sigh = sigh)
```

blqr — *Boundary line model determination using quantile regression*

## Description

This function fits a boundary model to the upper bounds of a scatter plot of x and y by estimating the conditional quantile (0-1) of the response variable, y, across values of the predictor variables, x. This is achieved using optimization procedure and hence requires some starting guess parameters of a proposed model.

## Usage

```
blqr(x,y,model, equation=NULL,start,tau=0.95,optim.method="Nelder-Mead",
     xmin=min(bound$x),xmax=max(bound$x), plot=TRUE,line_col="red",lwd=1,
     line_smooth=1000,...)
```

## Arguments

| | |
|---|---|
| x | A numeric vector of values for the independent variable. |
| y | A numeric vector of values for the response variable. |
| model | Selects the functional form of the boundary line. It includes "blm" for linear model, "lp" for linear plateau model, "mit" for the Mitscherlich model, "schmidt" for the Schmidt model, "logistic" for logistic model, "logisticND" for logistic model proposed by Nelder (1961), "inv-logistic" for the inverse logistic model, "double-logistic" for the double logistic model, "qd" for quadratic model and the "trapezium" for the trapezium model.For custom models, set model = "other". |
| equation | A custom model function writen in the form of an R function. Applies only when argument model="other", else it is NULL. |
| start | A numeric vector of initial starting values for optimization in fitting the boundary model. Its length and arrangement depend on the suggested model:<br>• For the "blm" model, it is a vector of length 2 arranged as intercept and slope.<br>• For the "lp" model, it is a vector of length 3 arranged as intercept, slope and maximum response.<br>• For the "logistic" and "inv-logistic" models, it is a vector of length 3 arranged as the scaling parameter, shape parameter and maximum response.<br>• For the "logisticND" model proposed by Nelder (1961), it is a vector of length 3 arranged as the scaling parameter, shape parameter and maximum response.<br>• For the "double-logistic" model, it is a vector of length 6 arranged as the scaling parameter one, shape parameter one, maximum response, maximum response, scaling parameter two and shape parameter two.<br>• For the "qd" model, it is a vector of length 3 arranged as constant, linear coefficient and quadratic coefficient. |

- For the "trapezium" model, it is a vector of length 3 arranged as intercept one, slope one, maximum response, intercept two and slope two.
- For the "mit" model, it is a vector of length 3 arranged as the intercept, shape parameter and the maximum response.
- For the "schmidt" model, it is a vector of length 3 arranged as scaling parameter, shape parameter (x-value at maximum response ) and maximum response.

`tau` The quantile value (0- 1) that represents the boundary (`default is tau = 0.95`).

`optim.method` Describes the method used to optimize the model as in the `optim()` function. The methods include "Nelder-Mead", "BFGS", "CG", "L-BFGS-B", "SANN" and "Brent".

`xmin` Numeric value that describes the minimum `x` value to which the boundary line is to be fitted (default is `min(x)`).

`xmax` A numeric value that describes the maximum `x` value to which the boundary line is to be fitted (default is `max(x)`). `xmin` and `xmax` determine the subset of the data set used to fit boundary model.

`plot` If `TRUE`, a plot is part of the output. If `FALSE`, plot is not part of output (default is `TRUE`).

`line_col` Selects the color of the boundary line.

`lwd` Determines the thickness of the boundary line on the plot (default is 1).

`line_smooth` Parameter that describes the smoothness of the boundary line. (default is 1000). The higher the value, the smoother the line.

`...` Additional graphical parameters.

## Details

Some inbuilt models are available for the `blqr()` function. The suggest model forms are as follows:

1. Linear model ("blm")
$$y = \beta_1 + \beta_2 x$$
where $\beta_1$ is the intercept and $\beta_2$ is the slope.
2. Linear plateau model ("lp")
$$y = \min(\beta_1 + \beta_2 x, \beta_0)$$
where $\beta_1$ is the intercept , $\beta_2$ is the slope and $\beta_0$ is the maximum response.
3. The logistic ("logistic") and inverse logistic ("inv-logistic") models
$$y = \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
$$y = \beta_0 - \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
where $\beta_1$ is a scaling parameter , $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

4. Logistic model ("logisticND") (Nelder (1961))

$$y = \frac{\beta_0}{1 + (\beta_1 \times e^{-\beta_2 x})}$$

   where $\beta_1$ is a scaling parameter, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

5. Double logistic model ("double-logistic")

$$y = \frac{\beta_{0,1}}{1 + e^{\beta_2(\beta_1 - x)}} - \frac{\beta_{0,2}}{1 + e^{\beta_4(\beta_3 - x)}}$$

   where $\beta_1$ is a scaling parameter one, $\beta_2$ is a shape parameter one, $\beta_{0,1}$ and $\beta_{0,2}$ are the maximum response , $\beta_3$ is a scaling parameter two and $\beta_4$ is a shape parameter two.

6. Quadratic model ("qd")

$$y = \beta_1 + \beta_2 x + \beta_3 x^2$$

   where $\beta_1$ is a constant, $\beta_2$ is a linear coefficient and $\beta_3$ is the quadratic coefficient.

7. Trapezium model ("trapezium")

$$y = \min(\beta_1 + \beta_2 x, \beta_0, \beta_3 + \beta_4 x)$$

   where $\beta_1$ is the intercept one, $\beta_2$ is the slope one, $\beta_0$ is the maximum response, $\beta_3$ is the intercept two and $\beta_3$ is the slope two.

8. Mitscherlich model ("mit")

$$y = \beta_0 - \beta_1 * \beta_2^x$$

   where $\beta_1$ is the intercept, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

9. Schmidt model ("schmidt")

$$y = \beta_0 + \beta_1(x - \beta_2)^2$$

   where $\beta_1$ is ascaling parameter, $\beta_2$ is a shape parameter (x-value at maximum response ) and $\beta_0$ is the maximum response .

10. Custom model ("other") This option allows you to create your own model form using the function `function()`. The custom model should be assigned to the argument `equation`. Note that the parameters for the custom model should be `a` and `b` for a two parameter model; `a`, `b` and `c` for a three parameter model; `a`, `b`, `c` and `d` for a four parameter model and so on.

The function `blbin()` utilities the optimization procedure of the `optim()` function to determine the model parameters. There is a tendency for optimization algorithms to settle at a local optimum. To remove the risk of settling for local optimum parameters, it is advised that the function is run using several starting values and the results with the smallest error (weighted residue sum square) can be taken as a representation of the global optimum.

The common errors encountered due to poor start values

1. function cannot be evaluated at initial parameters
2. initial value in 'vmmin' is not finite

## Value

A list of length 5 consisting of the fitted model, equation form, parameters of the boundary line, the weighted residue sum square. Additionally, a graphical representation of the boundary line on the scatter plot is produced.

### Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

### References

Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment, 1(8), 412-420.

Nelder, J.A. 1961. The fitting of a generalization of the logistic curve. Biometrics 17: 89–110.

Phillips, B.F. & Campbell, N.A. 1968. A new method of fitting the von Bertelanffy growth curve using data on the whelk. Dicathais, Growth 32: 317–329.

Schmidt, U., Thöni, H., & Kaupenjohann, M. (2000). Using a boundary line approach to analyze N2O flux data from agricultural soils. Nutrient Cycling in Agroecosystems, 57, 119-129.

### Examples

```
x<-log(SoilP$P)
y<-SoilP$yield
start<-c(4,3,13.6)

blqr(x,y, start=start,model = "lp", tau=0.99,
      xlab=expression("ET mm ha"^-1),
      ylab=expression("Wheat yield/ ton ha"^-1),
      pch=16, col="grey")
```

---

## bolides *Boundary line determination technique*

---

### Description

This function selects upper bounding points of a scatter plot of x and y based on the boundary line determination technique proposed by Schnug et al. (1995). A model is then fitted to the resulting boundary points by the least squares method. This is done using optimization procedure and hence requires some starting values for the model parameters for the proposed model.

### Usage

```
bolides(x,y,model="explore", equation=NULL, start, optim.method="Nelder-Mead",
        xmin=min(bound$x), xmax=max(bound$x), plot=TRUE,bp_col="red", bp_pch=16,
         bl_col="red" ,lwd=1,line_smooth=1000,...)
```

### Arguments

| | |
|---|---|
| x | A numeric vector of values for the independent variable. |
| y | A numeric vector of values for the response variable. |

`model` Selects the functional form of the boundary line. It includes `"explore"` as default, `"blm"` for linear model, `"lp"` for linear plateau model, `"mit"` for the Mitscherlich model, `"schmidt"` for the Schmidt model, `"logistic"` for logistic model, `"logisticND"` for logistic model proposed by Nelder (1961), `"inv-logistic"` for the inverse logistic model, `"double-logistic"` for the double logistic model, `"qd"` for quadratic model and the `"trapezium"` for the trapezium model. The `"explore"` is used to check the position of boundary points so that the correct `model` can be applied. For custom models, set `model = "other"`.

`equation` A custom model function writen in the form of an R function. Applies only when argument `model="other"`, else it is `NULL`.

`start` A numeric vector of initial starting values for optimization in fitting the boundary model. Its length and arrangement depend on the suggested model:

- For the `"blm"` model, it is a vector of length 2 arranged as intercept and slope.
- For the `"lp"` model, it is a vector of length 3 arranged as intercept, slope and maximum response.
- For the `"logistic"` and `"inv-logistic"` models, it is a vector of length 3 arranged as the scaling parameter, shape parameter and maximum response.
- For the `"logisticND"` model proposed by Nelder (1961), it is a vector of length 3 arranged as the scaling parameter, shape parameter and maximum response.
- For the `"double-logistic"` model, it is a vector of length 6 arranged as the scaling parameter one, shape parameter one, maximum response, maximum response, scaling parameter two and shape parameter two.
- For the `"qd"` model, it is a vector of length 3 arranged as constant, linear coefficient and quadratic coefficient.
- For the `"trapezium"` model, it is a vector of length 3 arranged as intercept one, slope one, maximum response, intercept two and slope two.
- For the `"mit"` model, it is a vector of length 3 arranged as the intercept, shape parameter and the maximum response.
- For the `"schmidt"` model, it is a vector of length 3 arranged as scaling parameter, shape parameter (x-value at maximum response ) and maximum response.

`optim.method` Describes the method used to optimize the model as in the `optim()` function. The methods include `"Nelder-Mead"`, `"BFGS"`, `"CG"`, `"L-BFGS-B"`, `"SANN"` and `"Brent"`.

`xmin` Numeric value that describes the minimum `x` value to which the boundary line is to be fitted (default is `min(x)`).

`xmax` A numeric value that describes the maximum `x` value to which the boundary line is to be fitted (default is `max(x)`). `xmin` and `xmax` determine the subset of the data set used to fit boundary model.

`plot` If `TRUE`, a plot is part of the output. If `FALSE`, plot is not part of output (default is `TRUE`).

`bp_col` Selects the color of the boundary points.

| | |
|---|---|
| bp_pch | Point character as pch of the plot() function. It controls the shape of the boundary points on plot (bp_pch = 16 as default). |
| bl_col | Colour of the boundary line. |
| lwd | Determines the thickness of the boundary line on the plot (default is 1). |
| line_smooth | Parameter that describes the smoothness of the boundary line. (default is 1000). The higher the value, the smoother the line. |
| ... | Additional graphical parameters as in the par() function. |

## Details

Some inbuilt models are available for the bolides() function. The "explore" option for the argument model generates a plot showing the ocation of the boundary points selected by the binning procedure. This helps to identify which model type is suitable to fit as a boundary line. The suggest model forms are as follows:

1. Linear model ("blm")
$$y = \beta_1 + \beta_2 x$$
where $\beta_1$ is the intercept and $\beta_2$ is the slope.
2. Linear plateau model ("lp")
$$y = \min(\beta_1 + \beta_2 x, \beta_0)$$
where $\beta_1$ is the intercept , $\beta_2$ is the slope and $\beta_0$ is the maximum response.
3. The logistic ("logistic") and inverse logistic ("inv-logistic") models
$$y = \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
$$y = \beta_0 - \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
where $\beta_1$ is a scaling parameter , $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
4. Logistic model ("logisticND") (Nelder (1961))
$$y = \frac{\beta_0}{1 + (\beta_1 \times e^{-\beta_2 x})}$$
where $\beta_1$ is a scaling parameter, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
5. Double logistic model ("double-logistic")
$$y = \frac{\beta_{0,1}}{1 + e^{\beta_2(\beta_1 - x)}} - \frac{\beta_{0,2}}{1 + e^{\beta_4(\beta_3 - x)}}$$
where $\beta_1$ is a scaling parameter one, $\beta_2$ is a shape parameter one, $\beta_{0,1}$ and $\beta_{0,2}$ are the maximum response , $\beta_3$ is a scaling parameter two and $\beta_4$ is a shape parameter two.
6. Quadratic model ("qd")
$$y = \beta_1 + \beta_2 x + \beta_3 x^2$$
where $\beta_1$ is a constant, $\beta_2$ is a linear coefficient and $\beta_3$ is the quadratic coefficient.

7. Trapezium model ("trapezium")

$$y = \min(\beta_1 + \beta_2 x, \beta_0, \beta_3 + \beta_4 x)$$

   where $\beta_1$ is the intercept one, $\beta_2$ is the slope one, $\beta_0$ is the maximum response, $\beta_3$ is the intercept two and $\beta_3$ is the slope two.

8. Mitscherlich model ("mit")

$$y = \beta_0 - \beta_1 * \beta_2^x$$

   where $\beta_1$ is the intercept, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.

9. Schmidt model ("schmidt")

$$y = \beta_0 + \beta_1 (x - \beta_2)^2$$

   where $\beta_1$ is ascaling parameter, $\beta_2$ is a shape parameter (x-value at maximum response ) and $\beta_0$ is the maximum response .

10. Custom model ("other") This option allows you to create your own model form using the function `function()`. The custom model should be assigned to the argument `equation`. Note that the parameters for the custom model should be `a` and `b` for a two parameter model; `a`, `b` and `c` for a three parameter model; `a`, `b`, `c` and `d` for a four parameter model and so on.

The function `bolides()` utilities the optimization procedure of the `optim()` function to determine the model parameters. There is a tendency for optimization algorithms to settle at a local optimum. To remove the risk of settling for local optimum parameters, it is advised that the function is run using several starting values and the results with the smallest error (residue mean square) can be taken as a representation of the global optimum.

The common errors encountered due to poor start values

1. function cannot be evaluated at initial parameters
2. initial value in 'vmmin' is not finite

## Value

A list of length 5 consisting of the fitted model, equation form, parameters of the boundary line, the residue mean square and the boundary points. Additionally, a graphical representation of the boundary line on the scatter plot is produced.

## Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

## References

Nelder, J.A. 1961. The fitting of a generalization of the logistic curve. Biometrics 17: 89–110.

Phillips, B.F. & Campbell, N.A. 1968. A new method of fitting the von Bertelanffy growth curve using data on the whelk. Dicathais, Growth 32: 317–329.

Schmidt, U., Thöni, H., & Kaupenjohann, M. (2000). Using a boundary line approach to analyze N2O flux data from agricultural soils. Nutrient Cycling in Agroecosystems, 57, 119-129. Schnug, E., Heym, J. M., & Murphy, D. P. L. (1995). Boundary line determination technique (BOLIDES). In P. C. Robert, R. H. Rust, & W. E. Larson (Eds.), site specific management for agricultural systems (p. 899-908). Wiley Online Library.

### Examples

```
x<-log(SoilP$P)
y<-SoilP$yield
start<-c(4,3,13.6,35,-5)

bolides(x,y,start=start,model = "trapezium",
        xlab=expression("Phosphorus/ln(mg L"^-1*")"),
        ylab=expression("Yield/ t ha"^-1), pch=16,
        col="grey", bp_col="grey")
```

---

cbvn *Fitting boundary line using censored bivariate normal model*

---

### Description

This function fits a response model to the upper limits of a scatter plot of of x and y to determine the most efficient response of y as a function of x (given a measurement error of y) based on a censored distribution (Milne et al., 2016). The location of censor in the data cloud is determined based on the maximum likelihood approach. This is done using optimization procedure and hence requires some starting guess parameters for the proposed model. It then compares the results with an uncensored normal bivariate distribution to access the appropriateness of the censored model.

### Usage

```
cbvn(data,model="lp", equation=NULL, start, sigh, UpLo="U", optim.method="BFGS",
      Hessian=FALSE, plot=TRUE, line_smooth=1000, lwd=2, l_col="red",...)
```

### Arguments

| | |
|---|---|
| data | A dataframe with two numeric columns, independent (x) and dependent (y) variables respectively. |
| model | Selects the functional form of the boundary line. It includes "blm" for linear model, "lp" for linear plateau model, "mit" for the Mitscherlich model, "schmidt" for the Schmidt model, "logistic" for logistic model, "logisticND" for logistic model proposed by Nelder (1961), "inv-logistic" for the inverse logistic model, "double-logistic" for the double logistic model, "qd" for quadratic model and the "trapezium" for the trapezium model. For custom models, set model = "other". |
| equation | A custom model function writen in the form of an R function. Applies only when argument model="other", else it is NULL. |
| start | A numeric vector of initial starting values for optimization in fitting the boundary model. Its length and arrangement depend on the suggested model: |

- For the "blm" model, it is a vector of length 7 arranged as the intercept, the slope, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "lp" model, it is a vector of length 8 arranged as the intercept, the slope, the maximum or plateau response, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "mit" model, it is a vector of length 8 arranged as the intercept, shape parameter, the maximum or plateau response, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "logistic", "inv-logistic" and "logisticND" models, it is a vector of length 8 arranged as scaling parameter, shape parameter, the maximum or plateau value, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "double-logistic" model, it is a vector of length 11 arranged as scaling parameter, shape parameter, maximum response, maximum response, scaling parameter two, shape parameter two, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "trapezium" model, it is a vector of length 10 arranged as intercept one, slope one, maximum response, intercept two, slope two, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "qd" model, it is a vector of length 8 arranged as a constant, linear coefficient, quadratic coefficient,mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.
- For the "schmidt" model, it is a vector of length 8 arranged the scaling parameter, shape parameter (x-value at maximum response ), maximum response, mean of x, mean of y, standard deviation of x, standard deviation of y and the correlation of x and y.

sigh — Standard deviation of the measurement error.

UpLo — Selects the type of boundary. "U" fits the upper boundary and "L" fits the lower boundary.

optim.method — Describes the method used to optimize the model as in the optim() function. The methods include "Nelder-Mead", "BFGS", "CG", "L-BFGS-B", "SANN" and "Brent".

Hessian — If True, the hessian matrix is part of the output (default is FALSE‘).

plot — If TRUE, a plot is part of the output. If FALSE, plot is not part of output (default is TRUE).

line_smooth — Parameter that describes the smoothness of the boundary line. (default is 1000). The higher the value, the smoother the line.

lwd — Determines the thickness of the boundary line on the plot (default is 1).

l_col — Selects the color of the boundary line.

... — Additional graphical parameters as in the par() function.

### Details

Some inbuilt models are available for the `cbvn()` function. The suggest model forms are as follows:

1. Linear model (`"blm"`)
$$y = \beta_1 + \beta_2 x$$
where $\beta_1$ is the intercept and $\beta_2$ is the slope.
2. Linear plateau model (`"lp"`)
$$y = \min(\beta_1 + \beta_2 x, \beta_0)$$
where $\beta_1$ is the intercept , $\beta_2$ is the slope and $\beta_0$ is the maximum response.
3. The logistic (`"logistic"`) and inverse logistic (`"inv-logistic"`) models
$$y = \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
$$y = \beta_0 - \frac{\beta_0}{1 + e^{\beta_2(\beta_1 - x)}}$$
where $\beta_1$ is a scaling parameter , $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
4. Logistic model (`"logisticND"`) (Nelder (1961))
$$y = \frac{\beta_0}{1 + (\beta_1 \times e^{-\beta_2 x})}$$
where $\beta_1$ is a scaling parameter, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
5. Double logistic model (`"double-logistic"`)
$$y = \frac{\beta_{0,1}}{1 + e^{\beta_2(\beta_1 - x)}} - \frac{\beta_{0,2}}{1 + e^{\beta_4(\beta_3 - x)}}$$
where $\beta_1$ is a scaling parameter one, $\beta_2$ is shape parameter one, $\beta_{0,1}$ and $\beta_{0,2}$ are the maximum response , $\beta_3$ is a scaling parameter two and $\beta_4$ is a shape parameter two.
6. Quadratic model (`"qd"`)
$$y = \beta_1 + \beta_2 x + \beta_3 x^2$$
where $\beta_1$ is a constant, $\beta_2$ is a linear coefficient and $\beta_3$ is the quadratic coefficient.
7. Trapezium model (`"trapezium"`)
$$y = \min(\beta_1 + \beta_2 x, \beta_0, \beta_3 + \beta_4 x)$$
where $\beta_1$ is the intercept one, $\beta_2$ is the slope one, $\beta_0$ is the maximum response, $\beta_3$ is the intercept two and $\beta_3$ is the slope two.
8. Mitscherlich model (`"mit"`)
$$y = \beta_0 - \beta_1 * \beta_2^x$$
where $\beta_1$ is the intercept, $\beta_2$ is a shape parameter and $\beta_0$ is the maximum response.
9. Schmidt model (`"schmidt"`)
$$y = \beta_0 + \beta_1(x - \beta_2)^2$$
where $\beta_1$ is a scaling parameter, $\beta_2$ is a shape parameter (x-value at maximum response ) and $\beta_0$ is the maximum response .

The function `cbvn()` utilities the optimization procedure of the `optim()` function to determine the model parameters. There is a tendency for optimization algorithms to settle at a local optimum. To remove the risk of settling for local optimum parameters, it is advised that the function is run using several starting values and the results with the smallest likelihood (or AIC) can be taken as a representation of the global optimum.

The common errors encountered due to poor start values

1. function cannot be evaluated at initial parameters
2. initial value in 'vmmin' is not finite

## Value

A list of length 5 consisting of the fitted model, equation form, parameters of the boundary line, AIC (for boundary line model and a null model) and a hessian matrix. Additionally, a graphical representation of the boundary line on the scatter plot is produced.

## Author(s)

1. Chawezi Miti <chawezi.miti@nottingham.ac.uk>
2. Richard Murray Lark <murray.lark@nottingham.ac.uk>

## References

Nelder, J.A. 1961. The fitting of a generalization of the logistic curve. Biometrics 17: 89–110.

Lark, R. M., & Milne, A. E. (2016). Boundary line analysis of the effect of water filled pore space on nitrous oxide emission from cores of arable soil. European Journal of Soil Science, 67 , 148-159.

Lark, R. M., Gillingham, V., Langton, D., & Marchant, B. P. (2020). Boundary line models for soil nutrient concentrations and wheat yield in national-scale datasets. European Journal of Soil Science, 71 , 334-351.

Milne, A. E., Ferguson, R. B., & Lark, R. M. (2006). Estimating a boundary line model for a biological response by maximum likelihood.Annals of Applied Biology, 149, 223–234.

Phillips, B.F. & Campbell, N.A. 1968. A new method of fitting the von Bertelanffy growth curve using data on the whelk. Dicathais, Growth 32: 317–329.

Schmidt, U., Thöni, H., & Kaupenjohann, M. (2000). Using a boundary line approach to analyze N2O flux data from agricultural soils. Nutrient Cycling in Agroecosystems, 57, 119-129.

## Examples

```
x<-evapotranspiration$`ET(mm)`
y<-evapotranspiration$`yield(t/ha)`
data<-data.frame(x,y)
start<-c(0.5,0.02,289.6,2.4,83.7,1.07,0.287)

cbvn(data, start=start, model = "blm", sigh=0.51,
        xlab=expression("ET/ mm ha"^-1),
        ylab=expression("Yield/ ton ha"^-1),
        pch=16, col="grey", line_smooth = 100)
```

## `evapotranspiration` *Evapotranspiration data*

### Description

This is a dataset compiled by Sadras & Angus (2006) that is comprises measures of wheat yield and estimated evapotranspiration (ET) from sites in China, the Mediterranean regions of Europe, North America, and Australia. For more details about this dataset refer to Sadras & Angus (2006).

### Usage

```
evapotranspiration
```

### Format

A data.frame with 691 rows and 3 columns:

**Region** Location of measurement.

**ET (mm)** Evapotranspiration measurement(mm) measured.

**Yield (t/ha)** Wheat yield (ton/ha) measured.

### Details

This data set should only be used for illustration purposes for this package and should not be used in any form publication without permission from the owners.

### Source

Sadras, V. O., & Angus, J. F. (2006). Benchmarking water-use efficiency of rainfed wheat in dry environments. Australian Journal of Agricultural Research, 57 , 847-856.

### Examples

```
data(evapotranspiration)
```

## `expl_boundary` *Testing evidence of boundary existence in dataset*

### Description

This function determines the probability of having bounding effects in a scatter plot of of `x` and `y` based on the clustering of points at the upper edge of the scatter plot (Miti et al.2024). It tests the hypothesis of larger clustering at the upper bounds of a scatter plot against a null bivariate normal distribution with no bounding effect (random scatter at upper edges). It returns the probability (p-value) of the observed clustering given that it a realization of an unbounded bivariate normal distribution.

### Usage

```
expl_boundary(x, y, shells = 10, simulations = 1000, plot = TRUE, ...)
```

### Arguments

| | |
|---|---|
| x | A numeric vector of values for the independent variable. |
| y | A numeric vector of values for the response variable. |
| shells | A numeric value indicating the number of boundary peels (default is 10). |
| simulations | The number of simulations for the null bivariate normally distributed data sets used to test the hypothesis (default is 1000). |
| plot | If TRUE, a plot is part of the output. If FALSE, plot is not part of output (default is TRUE). |
| ... | Additional graphical parameters as with the par() function. |

### Details

It is recommended that any outlying observations, as identified by the bagplot() function of the aplpack package are removed from the data. This is also implemented in the simulation step in the expl_boundary() function.

### Value

A dataframe with the p-values of obtaining the observed standard deviation of the euclidean distances of vertices in the upper peels to the center of the dataset for the left and right sections of the dataset.

### Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

### References

Eddy, W. F. (1982). Convex hull peeling, COMPSTAT 1982-Part I: Proceedings in Computational Statistics, 42-47. Physica-Verlag, Vienna.

Miti. c., Milne. A. E., Giller. K. E. and Lark. R. M (2024). Exploration of data for analysis using boundary line methodology. Computers and Electronics in Agriculture 219 (2024) 108794.

### Examples

```
x<-evapotranspiration$`ET(mm)`
y<-evapotranspiration$`yield(t/ha)`
expl_boundary(x,y,10,100) # recommendation is to set simulations to greater than 1000
```

## limfactor — *Determination of the most limiting factor to biological response*

### Description

This function determines the most limiting factor based on von Liebig law of the minimum given results of the predicted boundary line values for the different factors of interest. Boundary lines for various factors are fitted and the factor that predicts the minimum response for a particular point is considered as the most limiting factor (Casanova et al. 1995).

### Usage

```
limfactor(...)
```

### Arguments

| | |
|---|---|
| `...` | vectors with predicted values from the boundary line models for each factor being evaluated. |

### Value

A dataframe consisting of the most limiting factor and the minimum predicted response

### Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

### Examples

```
N<-rnorm(10,50,5)#assuming these are predicted responses using the fitted BL for N,P,K
K<-rnorm(10,50,4)
P<-rnorm(10,50,6)

limfactor(N,K,P)
```

## predictBL — *Predict boundary response*

### Description

This function predicts the most efficient response at a level of factor, x, given the parameters of the fitted boundary line.

### Usage

```
predictBL(object, x)
```

### Arguments

| | |
|---|---|
| `object` | An output in form of a list from the boundary line fitting using the `blqr()`, `blbin()`, `bolides()` or `cbvn()` functions. |
| `x` | A numeric vector of values for the factor with which response is to be predicted. |

### Value

A vector predicted value of response.

### Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

### Examples

```
x<-evapotranspiration$`ET(mm)`
y<-evapotranspiration$`yield(t/ha)`
z<-bolides(x,y, start = c(0.5,0.02), model= "blm", xmax = 350)

Results<-predictBL(z,x)

head(Results) # prediction for first 6 lines
```

---

## `soil` *Soil survey data*

---

### Description

This data set is a subset of a data set that was assembled by AgSpace Agriculture Ltd in a soil survey conducted on farms in various management units across England. The survey included measurements of wheat yield as well as various soil parameters. This particular dataset only contains the soil properties pH and phosphorus (P).

### Usage

```
soil
```

### Format

A data.frame with 6110 rows and 3 columns:

**yield** Wheat yield (ton/ha) measured.

**pH** pH measurement.

**P** soil p (mg/kg).

### Details

This data set should only be used for illustration purposes for this package and should not be used in any form publication without permission from the owners.

### Examples

```
data(soil)
```

---

## SoilP — *Soil Phosphorus data*

---

### Description

This data set is a subset of a data set that was assembled by AgSpace Agriculture Ltd in a soil survey conducted on farms in various management units across England.

### Usage

```
SoilP
```

### Format

A data.frame with 6020 rows and 2 columns:

**yield** Wheat yield (ton/ha) measured.

**P** Phosphorus (mg/l).

### Examples

```
data(SoilP)
```

---

## SoilpH — *Soil pH data*

---

### Description

This data set is a subset of a data set that was assembled by AgSpace Agriculture Ltd in a soil survey conducted on farms in various management units across England.

### Usage

```
SoilpH
```

### Format

A data.frame with 6047 rows and 2 columns:

**yield** Wheat yield (ton/ha) measured.

**pH** pH measurement.

### Examples

```
data(SoilpH)
```

---

## startValues — *Starting values for optimization functions*

---

### Description

This functions helps to determine initial values for a selected boundary line model when using the functions `blbin()`, `blqr()`, `bolides()`, `cbvn()` and `ble_profile()` to determine model parameters.

### Usage

```
startValues(model = "explore", p = NULL, digits = 2, ...)
```

### Arguments

| | |
|---|---|
| `model` | Selects the functional form of the boundary line. It includes `"blm"` for linear model, `"lp"` for linear plateau model, `"mit"` for the Mitscherlich model, `"schmidt"` for the Schmidt model, `"logistic"` for logistic model, `"logisticND"` for logistic model proposed by Nelder (1961), `"inv-logistic"` for the inverse logistic model, `"double-logistic"` for the double logistic model, `"qd"` for quadratic model, `"trapezium"` for the trapezium model and `"explore"` for function use exploration. The default is `"explore"`. |
| `p` | The number of selected points used to obtain start values for the logistic mitcherlich and schmidt models. It is `NULL` for other models. |
| `digits` | Number of decimal points for logistic type models (default is 2). |
| `...` | Additional graphical parameters. Applies to the logistic, mitcherlich and schmidt models to control the text on the plot. |

### Details

This function uses the `locator()` function. Once the model is selected, the points that make up the boundary points are selected using mouse click on the plots.

### Value

A list containing the parameters of the suggested model.

### Author(s)

Chawezi Miti <chawezi.miti@nottingham.ac.uk>

### References

Fekedulegn, D., Mac Siurtain, M.P., & Colbert, J.J. 1999. Parameter estimation of nonlinear growth models in forestry. Silva Fennica 33(4): 327–336.

Lark, R. M., & Milne, A. E. (2016). Boundary line analysis of the effect of water filled pore space on nitrous oxide emission from cores of arable soil. European Journal of Soil Science, 67 , 148-159.

### Examples

```
startValues(model="explore")
```

---

## summastat *Summary statistics*

---

### Description

A function to calculate summary statistics of a set of data.

### Usage

```
summastat(x, sigf, varname, plot = TRUE)
```

### Arguments

| | |
|---|---|
| `x` | A vector of numeric values. |
| `sigf` | The number of significant figures to report (optional). |
| `varname` | The name of the variable (optional), character so in quotes e.g. "Clay content". If not used then the variable is called x on plots. |
| `plot` | If `TRUE`, a plot is part of the output. If `FALSE`, plot is not part of output (default is `TRUE`). |

### Value

A matrix containing the mean value, median value, first and third quartiles, sample variance, sample standard deviation, coefficient of skewness, octile skewness, coefficient of kurtosis and the number of probable outliers in a data set. A histogram with a boxplot over it and QQ plot of the variable x if `plot=TRUE`.

### Author(s)

Richard Murray Lark <murray.lark@nottingham.ac.uk>

## Examples

```
x<-evapotranspiration$`ET(mm)`
summastat(x,2)
```

# Index