

# Evaluating forest quality with Forest

---

03/1/2023

## Evaluating forest quality with Forest

- 1 Introduction
  - 1.1 Overview
  - 1.2 Flowchart of data analysis
  - 1.3 Package dependencies
  - 1.4 Supported os:
  - 1.5 Package installation
- 2 Standard workflow
  - 2.1 Quick Start
  - 2.2 建立模型
    - 2.2.1 Loading forest data
    - 2.2.2 Select forest data to create forestData object
    - 2.2.3 Summary of inputs and models
  - 2.3 用forestData绘制示例图
  - 2.4 用forestData得到森林的潜在生产力
    - 2.4.1 潜在生产力输出数据说明
  - 2.5 用forestData得到森林的现实生产力
    - 2.5.1 现实生产力输出数据说明
  - 2.6 Summary of potential and reality productivity

## 1 Introduction

---

### 1.1 Overview

R包forest是基于Institute of Forest Resource Information Techniques,Chinese Academy of Forestry的doctor Fu L的天然林立地质量评价方法。forest包依赖于天然林立地的数据，forest包中带有一份样例数据。实现的功能有，天然林立地树高分级的划分，树高模型、断面积生长模型、蓄积生长模型的建立，森林现实生产力与潜在生产力的计算。其中，树高模型可用Richard模型，Logistic模型，korf模型，Gompertz模型，Weibull模型与Schumacher模型构建，断面积生长模型和蓄积生长模型仅可用Richard模型构建。

### 1.2 Flowchart of data analysis

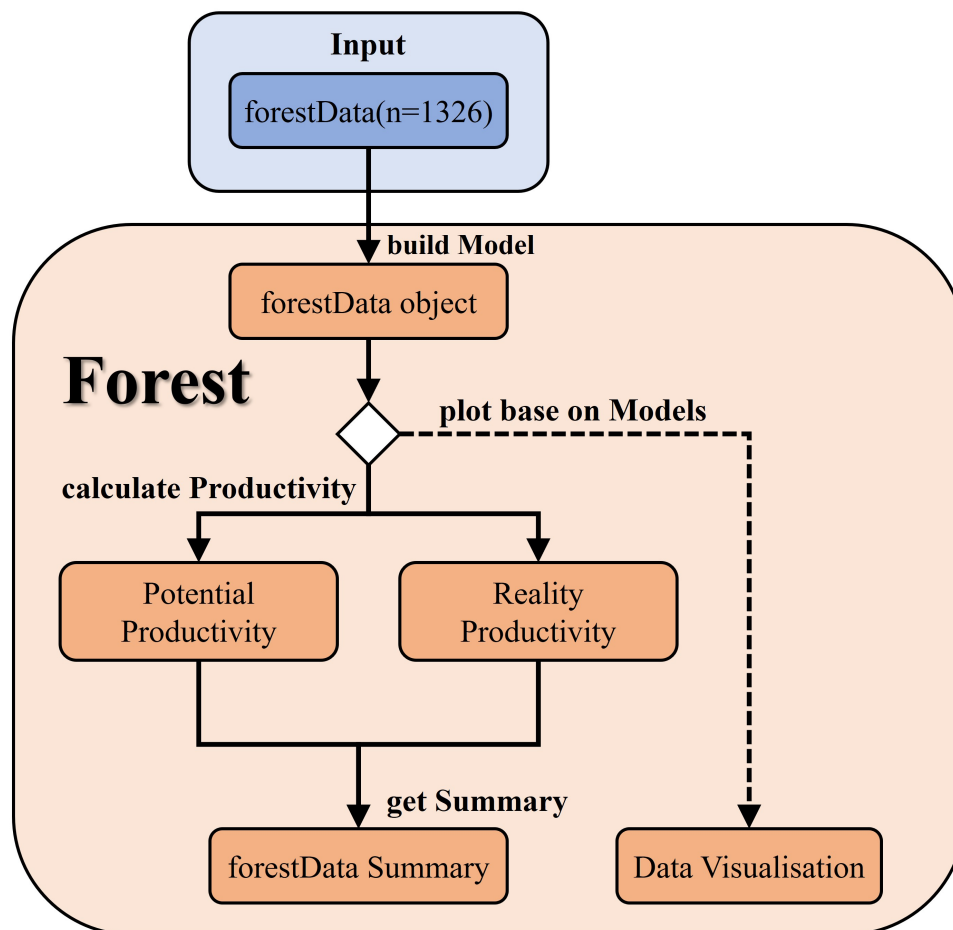


Figure 1. Flowchart of data analysis

### 1.3 Package dependencies

Package	Download Link
dplyr	<a href="https://CRAN.R-project.org/package=dplyr">https://CRAN.R-project.org/package=dplyr</a>
ggplot2	<a href="https://CRAN.R-project.org/package=ggplot2">https://CRAN.R-project.org/package=ggplot2</a>
nlme	<a href="https://CRAN.R-project.org/package=nlme">https://CRAN.R-project.org/package=nlme</a>

### 1.4 Supported os:

Windows, Linux and Mac OS are currently supported.

## 1.5 Package installation

You can install the released version of *localshiny* package from Cran or GitHub with the following command in R:

```
#install package dependencie
install.packages(c("dplyr", "ggplot2", "nlme"))

#install package
install.packages("forest")

#install.packages("devtools")

devtools::install_github("xxx/forest")
```

To ensure you have successfully installed *forest*, try loading it into your R session.

```
library(forest)
```

## 2 Standard workflow

### 2.1 Quick Start

这里我们展示的是林地质量评估的完整步骤。在class.plot的上游有一些步骤，在获得了AGE(Stand age of the tree), H(Height of the tree), S(Forest density index), BA(Basal area of the tree), Bio(Biomass of the tree)之后，应当自定义ID(Unique identifier for each tree)以及code(Codes for forest types)。这个代码块假设你有拥有了以上所述的数据。

```
#使用包中自带的forestData数据
data("forestData")
#使用forestData建立模型，并创建一个forestData class
forestData <- class.plot(forestData,model="Richards",
                        interval=5,number=5,a=19,b=0.1,c=0.8)
#绘制forestData的示例图
plot(forestData,model.type = "BA",plot.type = "Scatter",
     xlab = "AGE",ylab = "BA",legend.lab = "LastGroup",
     title = "Forest")
#计算forestData的潜在生产力
forestData <- potential.productivity(forestData)
#计算forestData的现实生产力
forestData <- reality.productivity(forestData)
#获得forestData的summary
summary(forestData)
```

## 2.2 建立模型

### 2.2.1 Loading forest data

为了建立一个准确的模型，好的数据是不可或缺的，在forest包中便内置了一个经过清洗的样例数据。

```
#loading forestData
data("forestData")
#you can also choose read.csv to load forestData
forestData <- read.csv("data\\forestData.csv")

head(select(forestData,ID,code,AGE,H,S,BA,Bio))
#output
```

	ID	code	AGE	H	S	BA	Bio
1	6100005337	1	45	11.9	1508.468	50.13462	474.4957
2	410001607	1	42	16.7	1490.493	47.22381	444.5069
3	6100005337	1	35	11.0	1401.944	46.64877	435.8741
4	6100005337	1	40	12.8	1303.489	44.15220	415.9098
5	410001607	1	38	15.2	1350.941	42.37152	400.3925
6	6220002848	1	88	11.2	1631.235	50.43886	395.2503

在后续分析所需要的数据中，ID(Unique identifier for each tree)，code(Codes for forest types)，AGE(Stand age of the tree)，H(Height of the tree)是不可或缺的，用以在后面的流程中建立树高模型(H Model)，并绘制相关示例图。

S(Forest density index)，BA(Basal area of the tree)，Bio(Biomass of the tree)是可选的数据，用以建立横截面面积模型(BA Model)与生物量模型(Bio Model)。但在后续现实生产力与潜在生产力的计算中，BA Model与Bio Model是必须的，也就是没有S，BA，Bio数据将无法计算生产力。

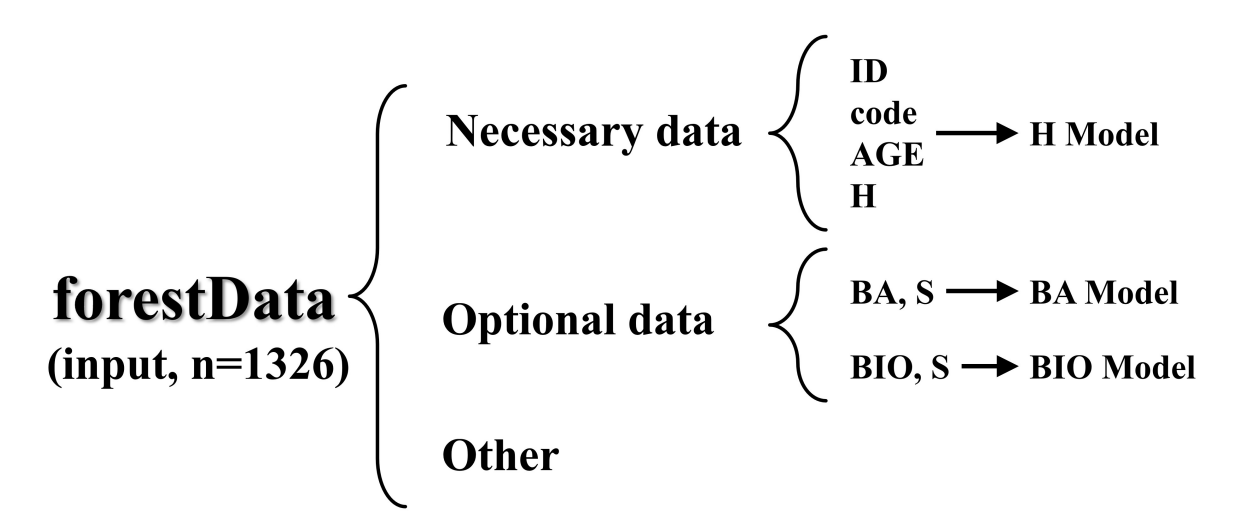


Figure 2. Description of forestData example

2.2.2 Select forest data to create forestData object

为了建立数学模型，并返回一个forestData object以供接下来的分析，需要使用class.plot函数。

class.plot至少需要"ID","code", "AGE","H"四列数据作为输入，以建立树高模型，同时会创建一个forestData类S3 method数据。如果输入数据中还有S，BA与Bio列的话，那么同时也会建立BA和Bio模型。

其中参数model可以选择"Logistic"、"Richards"、"Korf"、"Gompertz"、"Weibull"、"Schumacher"这六个中的一个，这会使用对应的model形式建立H model；interval 是树高分类的区间，interval=5就是创建一个以5 stand ages作为区间的初始树高分类；number是树高分类区间的最大值，number=5即初始树高分类数最多为5；a,b,c 是拟合模型的初始参数，当拟合出现错误时，可以多尝试一些初始参数作为尝试。

```
forestData <- class.plot(forestData,model="Richards",
                        interval=5,number=5,a=19,b=0.1,c=0.8)
```

由class.plot返回的forestData object包括继续分析所必须的数据(grading),由nlme函数返回的H model list, BA model list, Bio model list(Hmodel, BAmode, Biomodel)以及模型的参数 (estimateParameter)。

forestData Object

\$Input	Input data and tree height grading
\$Hmodel	nlme model of H data
\$BAmode	nlme model of BA data
\$Biomodel	nlme model of Bio data
\$output	Parameter of models

Figure 3. Description of forestData Object

2.2.3 Summary of inputs and models

为了得到输入数据与模型的summary，以了解模型的建立情况，可以使用forestData Object的summary函数。forestData Object的summary函数会返回一个summary.forestData Object并将summary打印在屏幕上。其中输出的第一段为模型各种输入数据的summary，第二，三，四段分别为H Model，BA Model，Bio Model的各项参数与参数的精简报告。

```
summary(forestData)
#output
#first paragraph
      H           S           BA           Bio
Min.   : 2.00   Min.   : 15.94   Min.   : 0.3017   Min.   : 1.224
1st Qu.: 7.70   1st Qu.: 360.38   1st Qu.: 9.3241   1st Qu.: 53.233
```

Median : 9.40	Median : 557.25	Median :14.9777	Median : 95.002
Mean : 9.83	Mean : 583.88	Mean :16.2648	Mean :109.322
3rd Qu.:11.90	3rd Qu.: 764.38	3rd Qu.:21.6455	3rd Qu.:147.737
Max. :17.70	Max. :1772.26	Max. :52.6455	Max. :474.496

## #second paragraph

Hmodel Parameters:

Nonlinear mixed-effects model fit by maximum likelihood

Model:  $H \sim 1.3 + a * (1 - \exp(-b * AGE))^c$

Data: data

AIC	BIC	logLik
2720.209	2746.159	-1355.105

Random effects:

Formula:  $a \sim 1 \mid \text{LASTGROUP}$

a Residual

StdDev: 3.6513 0.6616545

Fixed effects:  $a + b + c \sim 1$

	Value	Std.Error	DF	t-value	p-value
a	11.226213	1.6509803	1319	6.799726	0
b	0.020457	0.0029541	1319	6.924853	0
c	0.370395	0.0228807	1319	16.188147	0

Correlation:

	a	b
b	-0.137	
c	-0.122	0.949

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-4.13170023	-0.75823758	-0.03968202	0.74727148	4.97834758

Number of Observations: 1326

Number of Groups: 5

Concise Parameter Report:

Model Coefficients:

a1	a2	a3	a4	a5	b	c
6.331338	8.578689	10.91438	13.61481	16.69184	0.0204566	0.3703953

Model Evaluations:

pe	RMSE	R2	Var	TRE	AIC	BIC
-0.001864527	0.6604061	0.9455896	0.4364619	0.4185215	2720.209	2746.159
logLik						
-1355.105						

Model Formulas:

	Func	Spe
model1:	$H \sim 1.3 + a * (1 - \exp(-b * AGE))^c$	model1:pdDiag( $a \sim 1$ )

#third paragraph(similar in format to the second paragraph)

#Omitted here

```
BAmode1 Parameters:
```

```
.....
```

```
#fourth paragraph(Still similar in format to the second paragraph)
```

```
#Omitted here
```

```
Biomode1 Parameters:
```

```
.....
```

## 2.3 用forestData绘制示例图

forestData类的plot函数，使用经过class.plot函数处理过后的数据，可以分别绘制H Model,BA Model,Bio Model的Curve图，residual图，数据Scatter图，数据Scatter与模型拟合曲线图。

forestData Object的plot函数参数model.type可以选择H,BA或Bio，对应绘制的图使用的模型；plot.type可以选择Curve,residual,Scatter或Scatter\_Curve，对应绘制图的类型；xlab,ylab,legend.lab,title分别可以进行x轴，y轴，图例，标题上文字的更改。

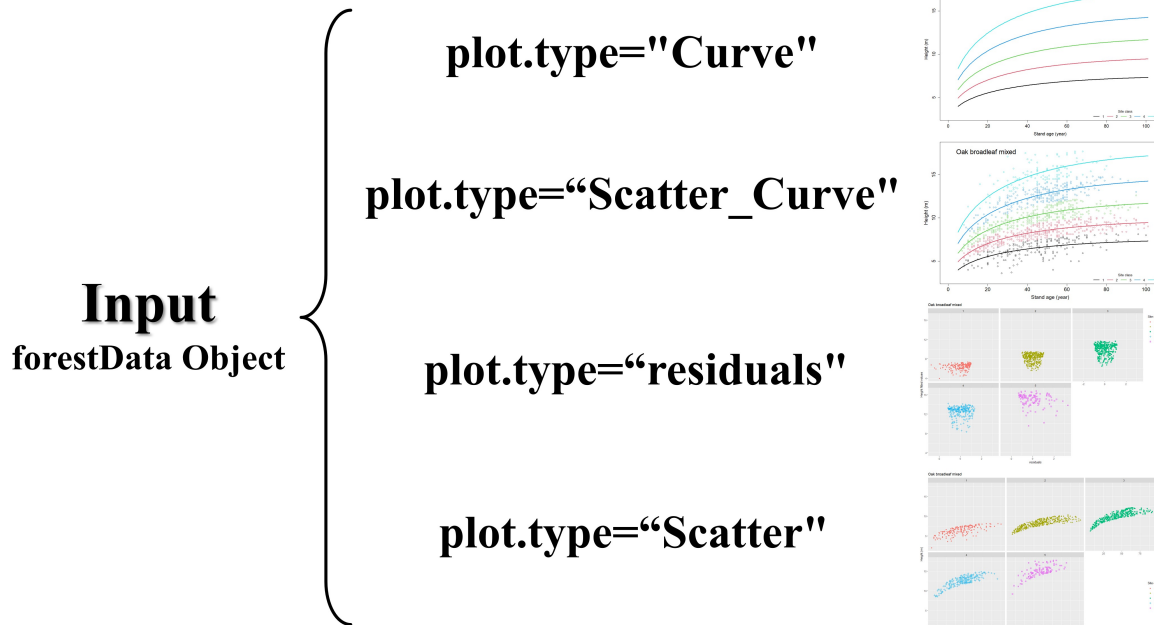


Figure 4. Output figure of forestData

```
#plot Figure 5
```

```
plot(forestData,model.type="H",  
     plot.type="Curve",  
     xlab="Stand age (year)",ylab="Height (m)",legend.lab="Site class",  
     title="Oak broadleaf")
```

```
#plot Figure 6
```

```
plot(forestData,model.type="H",  
     plot.type="Curve",  
     xlab="Age",ylab="Height",legend.lab="LASTGROUP",  
     title="Broadleaf")
```

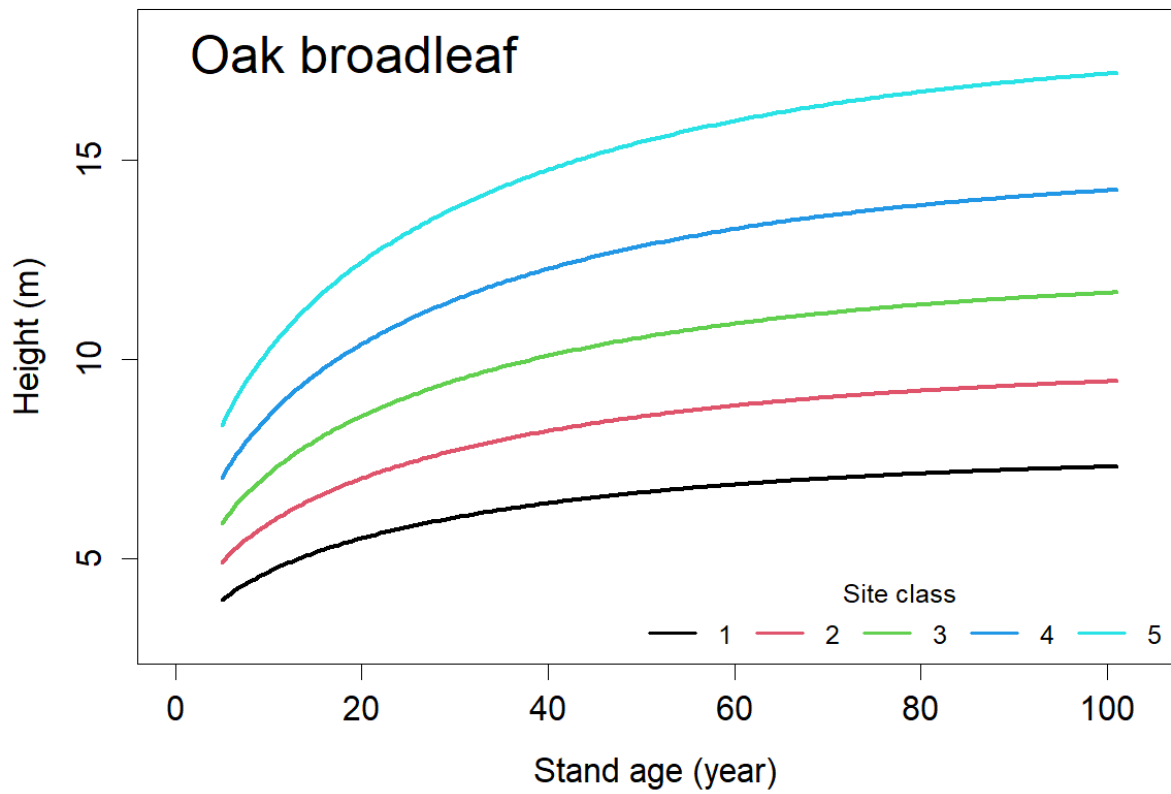


Figure 5. First example of figure

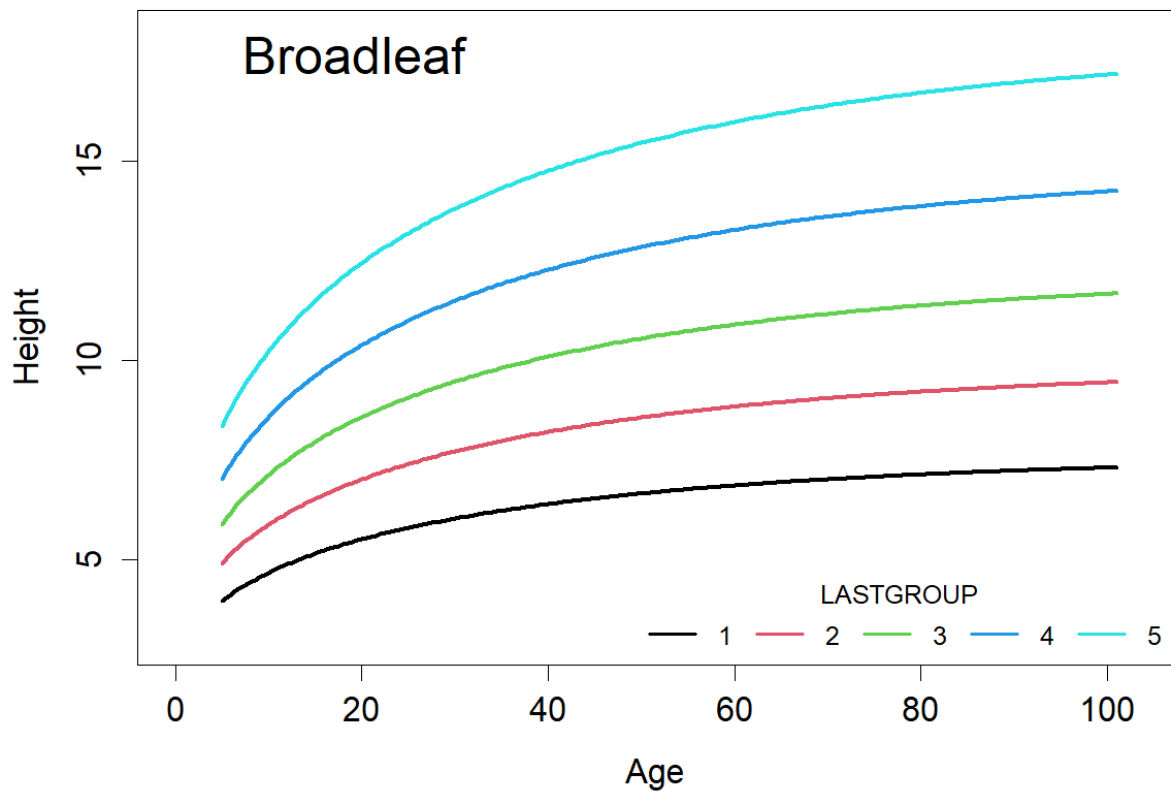


Figure 6. Second example of figure

在上述两张图例中，各项绘图参数不一致，导致绘出的图也有细微变化。图例1的xlab="Stand age (year)", 图例2的xlab="Age"; 图例1的ylab="Height (m)", 图例2的ylab="Height"; 图例1的legend.lab="Site class", 图例2的legend.lab="LASTGROUP"; 图例1的title="Oak broadleaf", 图例2的title="Broadleaf"。



## 2.4 用forestData得到森林的潜在生产力

为了建立森林潜在生产力模型，以得到森林的潜在生产力，需要使用potential.productivity函数。  
potential.productivity函数需要forestData Object数据作为输入，在forestData Object 中BA model与Bio model必须已经建立，输出结果将存储在forestData Object数据中。

```
forestData <- potential.productivity(forestData, code=1,
                                     age.min=5,age.max=150,
                                     left=0.05, right=100,
                                     e=1e-05, maxiter = 50)
```

其中参数code对应forestData中的code，这会使用对应code的森林数据计算其现实生产力或潜在生产力。  
age.min与age.max即森林stand age的最小值与最大值，潜在生产力的计算会在这最小值与最大值的区间中进行。  
left与right是拟合模型的初始参数，当拟合出现错误时，可以多尝试一些初始参数作为尝试。  
e为拟合模型的精度，当残差低于e时，认为模型收敛并停止拟合。  
maxiter为拟合模型的最大次数，当拟合次数等于maxiter时，认为模型收敛并停止拟合

### 2.4.1 潜在生产力输出数据说明

- Max\_GI：最大林分断面积
- Max\_MI：蓄积最大生长量
- N1：达到潜在生长量对应的林分株数
- D1：达到潜在生长里对应的林分平均直径
- S0：初始林分密度指数
- S1：达到潜在生长里对应的林分最佳密度指数
- G0：初始林分每公顷断面积
- G1：达到潜在生长量对应的林分每公顷断面积(1年以后)
- M0：初始林分每公顷蓄积
- M1：达到潜在生长量对应的林分每公顷蓄积

```
library(dplyr)
forestData$potential.productivity %>% head(.)
#output
      Max_GI  Max_MI      N1      D1      S0      S1      G0      G1
1  3.432031 25.28960 7314.484 7.871238 1509.526 1637.494 32.16056 35.59259
2  2.905191 21.52146 6715.212 8.179387 1492.373 1598.908 32.37989 35.28508
3  2.518421 18.74017 6241.259 8.455917 1476.469 1567.516 32.53121 35.04963
4  2.222457 16.60206 5854.617 8.707564 1461.918 1541.273 32.64190 34.86436
5  1.988672 14.90643 5530.422 8.938955 1448.291 1518.519 32.71869 34.70736
6  1.799336 13.52842 5253.364 9.153519 1435.503 1498.419 32.77100 34.57033

      M0      M1 LASTGROUP AGE
1 196.4822 221.7718        1  5
2 199.6247 221.1461        1  6
3 202.0687 220.8089        1  7
```

4	204.0589	220.6610	1	8
5	205.6789	220.5854	1	9
6	207.0203	220.5487	1	10

## 2.5 用forestData得到森林的现实生产力

为了建立森林现实生产力模型，以得到森林的现实生产力，需要使用reality.productivity函数。reality.productivity函数需要forestData Object数据作为输入，在forestData Object 中BA model与Bio model必须已经建立，输出结果将存储在forestData Object数据中。

```
forestData <- reality.productivity(forestData,
                                   left=0.05, right=100)
```

其中参数left与right是拟合模型的初始参数，当拟合出现错误时，可以多尝试一些初始参数作为尝试。

### 2.5.1 现实生产力输出数据说明

BAI：蓄积现实生产力

VI：蓄积潜在生产力

```
library(dplyr)
forestData$reality.productivity %>% head(.)
#output
```

	code	ID	AGE	H	class0	LASTGROUP	BA	S	Bio
1	1	6100005337	45	11.9	4	4	50.13462	1508.468	474.4957
2	1	410001607	42	16.7	5	5	47.22381	1490.493	444.5069
3	1	6100005337	35	11.0	3	4	46.64877	1401.944	435.8741
4	1	6100005337	40	12.8	4	4	44.15220	1303.489	415.9098
5	1	410001607	38	15.2	5	5	42.37152	1350.941	400.3925
6	1	6220002848	88	11.2	3	3	50.43886	1631.235	395.2503

	BAI	VI
1	0.36488249	2.8670220
2	0.42883352	3.6013437
3	0.57137875	4.7817218
4	0.51822786	4.4346054
5	0.55925908	4.9739993
6	0.07333166	0.3845029

## 2.6 Summary of potential and reality productivity

在得到森林潜在生产力与现实生产力后，可以使用forestData object的summary函数，以便了解潜在生产力与现实生产力。前四段在2.2.3中已经介绍，第五段即是潜在生产力与现实生产力的summary。

```
summary(forestData)
#output
#first paragraph
```

	H	S	BA	Bio
Min.	: 2.00	Min. : 15.94	Min. : 0.3017	Min. : 1.224
...				

### #fifth paragraph

Max_GI		Max_MI	
Min.	:0.1244	Min.	: 1.009
1st Qu.	:0.1757	1st Qu.	: 1.517
Median	:0.2591	Median	: 2.206
Mean	:0.4715	Mean	: 3.909
3rd Qu.	:0.4878	3rd Qu.	: 4.086
Max.	:3.9588	Max.	:33.858
BAI		VI	
Min.	:0.0000	Min.	:0.000
1st Qu.	:0.1388	1st Qu.	:1.028
Median	:0.2077	Median	:1.597
Mean	:0.2353	Mean	:1.846
3rd Qu.	:0.3116	3rd Qu.	:2.558
Max.	:0.8562	Max.	:7.309