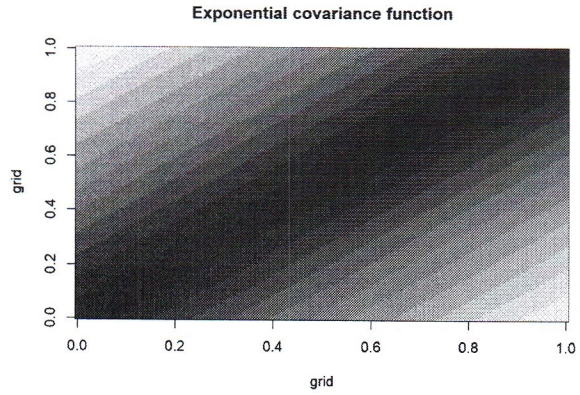
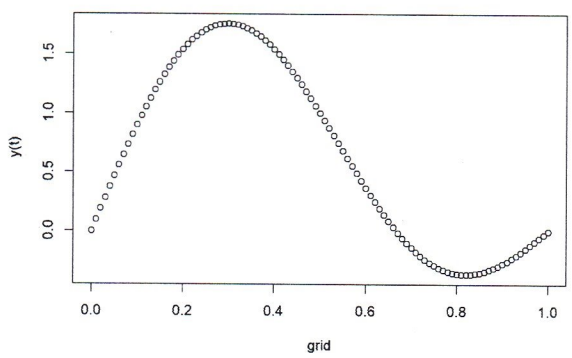
```
### ------------------------------------------------------------------------
### ------------------------------------------------------------------------
### DEPTH MEASURES - FUNCTIONAL VERSION
### ------------------------------------------------------------------------
### ------------------------------------------------------------------------
# Depth-based techniques find a natural application in Functional Statistics.
# We define functional data to be populations of real-valued continuous functions over a uni
# or multivariate domain, with a given degree of smoothness...

# A natural way to reason (and actually comprehend why they need specific and peculiar
# techniques) is to think about them as relatively dense LONGITUDINAL data: unlike normal
# data, functional data are not invariant to permutations of their dimensions(i.e if I switch
# two columns of a standard multivariate dataset, nothing actually changes... if I do it with a
# functional one, I will create a bloody mess.)

# I have structured this tutorial as a VERY fast primer to understand how to do some
# methodological and applied work with functional data.
# Since Functional Data Analysis is one of the core research areas of our department, you'll
# see several examples about applications of the techniques that are shown in the course...
# So, the very first part of this tutorial will be probably useful also further on.
# In this tutorial you'll learn how to:

# Simulate Functional Data
# How to compute a functional boxplot (and choose the proper depth for it...)
# How to build and interpret an outliergram the difference between shape/phase outliers
# and magnitude outliers a real world exercise...

# The package we will be using is the "roahd" package, that allows us faster simulation tasks, and
# the creation of outliergrams... To work with functional data the other main package is "fda",
# I will show you a bit about it for boxplots and so on...

#install.packages("roahd")
library(roahd)
```

```
### ------------------------------------------------------------------------
### ------------------------------------------------------------------------
### EXAMPLE 1 - How to simulate functional data + Boxplot + Outliergram
### ------------------------------------------------------------------------
### ------------------------------------------------------------------------
grid  = seq( 0, 1, length.out = 101 )   — grid on which data are defined
alpha = 0.2
beta  = 0.2
psi   = exp_cov_function( grid, alpha, beta )
image(psi,
      main = 'Exponential covariance function',
      xlab = 'grid', ylab = 'grid')
```

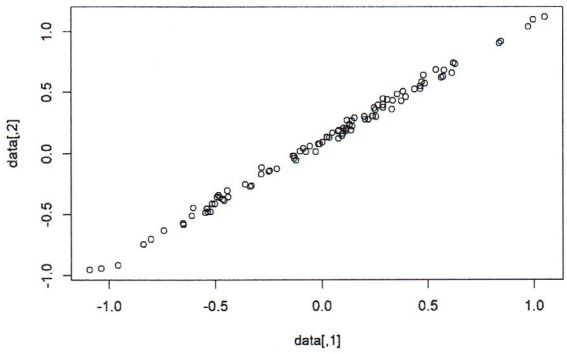



```
mu = sin(pi*grid)+sin(2*pi*grid)
plot(grid,mu, ylab = "y(t)")
```



```
# to simulate data you use...
n     = 100
set.seed(26111992)
data = generate_gauss_fdata(n,mu,Cov=psi)
# the output of t is actually a n X p matrix, where P is equal to the length of the grid
# instead of having a "proper" functional datum, I have its evaluation on a relatively
# fine grid...
# if I try to plot it...
plot(data)
```

we generate from a multivariate normal in practice since "functional data does not really exist"

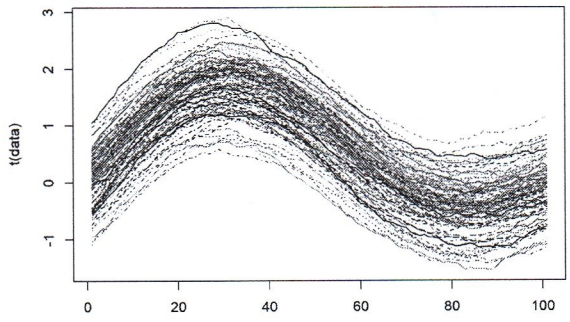



```
# what to do?
# either you keep them as a matrix, and remember to do this...
matplot(t(data),type='l')
```

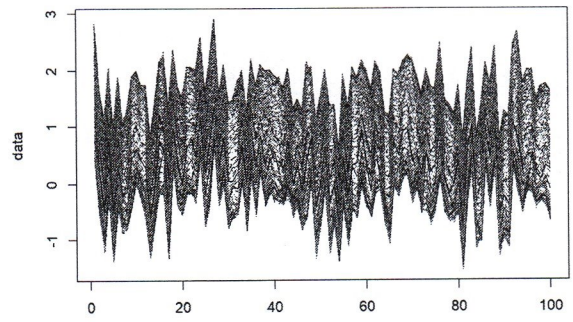
data =

	time-1	time-2	..	time-k
observation-1				
observation-2				
⋮				
observation-j				

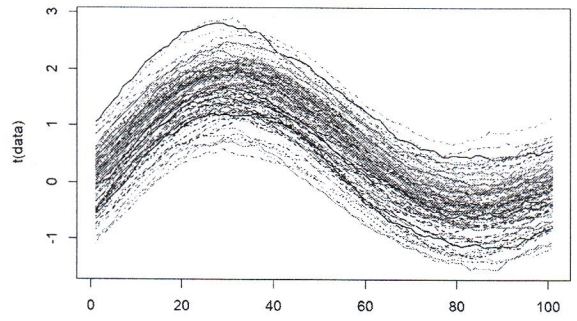
matplot wants observations on the columns
⟹ t(data)



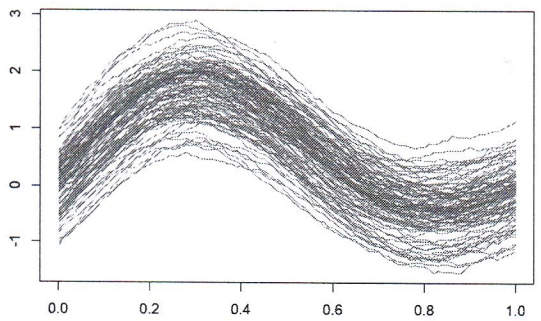

```
# don't forget the transpose, nor the type... otherwise...
matplot(data,type='l')
```




```
matplot(t(data),type='l')
```




```
# or, you fully exploit the power of the package...
fdata = fData(grid,data)
plot(fdata)
```




```
# you can compute depths with
BD(data)
```

```
##   [1] 0.02000000 0.18808081 0.05232323 0.13555556 0.02000000 0.15131313
##   [7] 0.06666667 0.12040404 0.16666667 0.09494949 0.16424242 0.14525253
##  [13] 0.03616162 0.13919192 0.10000000 0.05151515 0.05030303 0.05151515
##  [19] 0.15636364 0.16080808 0.14444444 0.09070707 0.11111111 0.03878788
##  [25] 0.13515152 0.07939394 0.02000000 0.15737374 0.10060606 0.17232323
##  [31] 0.19090909 0.19373737 0.09636364 0.05636364 0.03171717 0.16848485
##  [37] 0.03818182 0.10404040 0.12303030 0.06969697 0.15818182 0.18000000
##  [43] 0.10000000 0.11494949 0.13313131 0.11757576 0.05070707 0.09494949
##  [49] 0.05313131 0.09272727 0.12303030 0.06424242 0.15777778 0.02000000
##  [55] 0.09555556 0.07737374 0.09919192 0.13111111 0.10909091 0.11616162
##  [61] 0.18363636 0.07959596 0.12242424 0.07959596 0.08222222 0.10909091
##  [67] 0.13515152 0.09636364 0.07898990 0.07979798 0.13555556 0.08888889
##  [73] 0.06242424 0.17111111 0.20383838 0.05070707 0.19656566 0.14666667
##  [79] 0.10343434 0.19191919 0.02000000 0.14606061 0.05232323 0.03636364
##  [85] 0.07656566 0.04585859 0.15010101 0.03696970 0.02000000 0.10606061
##  [91] 0.08060606 0.03737374 0.02000000 0.10888889 0.08888889 0.08060606
##  [97] 0.17353535 0.15656566 0.17151515 0.17272727
```

```
MBD(data)
```

```
##   [1] 0.05719172 0.49653165 0.15976998 0.41158716 0.10031203 0.50471647
##   [7] 0.21503150 0.29264326 0.49960396 0.37481948 0.48817082 0.48499850
##  [13] 0.11867387 0.36584258 0.35317332 0.32794679 0.19312931 0.14647265
##  [19] 0.49721172 0.44727473 0.44336634 0.39178118 0.43681368 0.06240824
##  [25] 0.42253225 0.35573757 0.02623262 0.47286129 0.33650765 0.38392039
##  [31] 0.43922592 0.47504150 0.36330633 0.29880788 0.25878388 0.49344734
##  [37] 0.29682768 0.48142214 0.38613261 0.44941894 0.49735174 0.48808681
##  [43] 0.39169317 0.37477948 0.46695270 0.31890989 0.39058906 0.35276528
##  [49] 0.14329633 0.37268327 0.39447345 0.23693369 0.36784678 0.02997700
##  [55] 0.47637364 0.25546355 0.48403040 0.48865887 0.27104110 0.48541054
##  [61] 0.46270427 0.30912091 0.44830683 0.29656366 0.22154415 0.38610461
##  [67] 0.47769377 0.25992399 0.21277128 0.39889389 0.44842684 0.45694369
##  [73] 0.43167317 0.45565157 0.49539154 0.16483448 0.45467947 0.36253025
##  [79] 0.31687369 0.49307131 0.04170217 0.47351335 0.16186219 0.14755676
##  [85] 0.26747275 0.22022002 0.50583658 0.10202820 0.06994099 0.27106111
##  [91] 0.21666367 0.11248525 0.17698770 0.49314331 0.45107111 0.39074507
##  [97] 0.46877288 0.48143014 0.50100010 0.46247625
```

```
BD(fdata)
```

```
##   [1] 0.02000000 0.18808081 0.05232323 0.13555556 0.02000000 0.15131313
##   [7] 0.06666667 0.12040404 0.16666667 0.09494949 0.16424242 0.14525253
##  [13] 0.03616162 0.13919192 0.10000000 0.05151515 0.05030303 0.05151515
##  [19] 0.15636364 0.16080808 0.14444444 0.09070707 0.11111111 0.03878788
##  [25] 0.13515152 0.07939394 0.02000000 0.15737374 0.10060606 0.17232323
##  [31] 0.19090909 0.19373737 0.09636364 0.05636364 0.03171717 0.16848485
##  [37] 0.03818182 0.10404040 0.12303030 0.06969697 0.15818182 0.18000000
##  [43] 0.10000000 0.11494949 0.13313131 0.11757576 0.05070707 0.09494949
##  [49] 0.05313131 0.09272727 0.12303030 0.06424242 0.15777778 0.02000000
##  [55] 0.09555556 0.07737374 0.09919192 0.13111111 0.10909091 0.11616162
##  [61] 0.18363636 0.07959596 0.12242424 0.07959596 0.08222222 0.10909091
##  [67] 0.13515152 0.09636364 0.07898990 0.07979798 0.13555556 0.08888889
##  [73] 0.06242424 0.17111111 0.20383838 0.05070707 0.19656566 0.14666667
##  [79] 0.10343434 0.19191919 0.02000000 0.14606061 0.05232323 0.03636364
##  [85] 0.07656566 0.04585859 0.15010101 0.03696970 0.02000000 0.10606061
##  [91] 0.08060606 0.03737374 0.02000000 0.10888889 0.08888889 0.08060606
##  [97] 0.17353535 0.15656566 0.17151515 0.17272727
```

```
MBD(fdata)
```

```
##   [1] 0.05719172 0.49653165 0.15976998 0.41158716 0.10031203 0.50471647
##   [7] 0.21503150 0.29264326 0.49960396 0.37481948 0.48817082 0.48499850
##  [13] 0.11867387 0.36584258 0.35317332 0.32794679 0.19312931 0.14647265
##  [19] 0.49721172 0.44727473 0.44336634 0.39178118 0.43681368 0.06240824
##  [25] 0.42253225 0.35573757 0.02623262 0.47286129 0.33650765 0.38392039
##  [31] 0.43922592 0.47504150 0.36330633 0.29880788 0.25878388 0.49344734
##  [37] 0.29682768 0.48142214 0.38613261 0.44941894 0.49735174 0.48808681
##  [43] 0.39169317 0.37477948 0.46695270 0.31890989 0.39058906 0.35276528
##  [49] 0.14329633 0.37268327 0.39447345 0.23693369 0.36784678 0.02997700
##  [55] 0.47637364 0.25546355 0.48403040 0.48865887 0.27104110 0.48541054
##  [61] 0.46270427 0.30912091 0.44830683 0.29656366 0.22154415 0.38610461
##  [67] 0.47769377 0.25992399 0.21277128 0.39889389 0.44842684 0.45694369
##  [73] 0.43167317 0.45565157 0.49539154 0.16483448 0.45467947 0.36253025
##  [79] 0.31687369 0.49307131 0.04170217 0.47351335 0.16186219 0.14755676
##  [85] 0.26747275 0.22022002 0.50583658 0.10202820 0.06994099 0.27106111
##  [91] 0.21666367 0.11248525 0.17698770 0.49314331 0.45107111 0.39074507
##  [97] 0.46877288 0.48143014 0.50100010 0.46247625
```

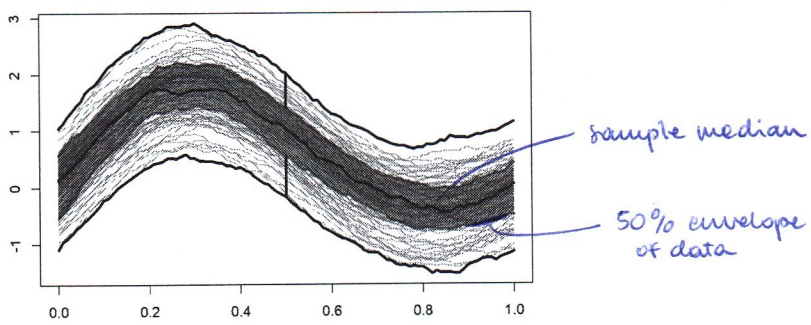
```
# with respective relative versions
BD_relative(mu,data)
```

```
## [1] 0.1878788
```

```
MBD_relative(mu,data)
```

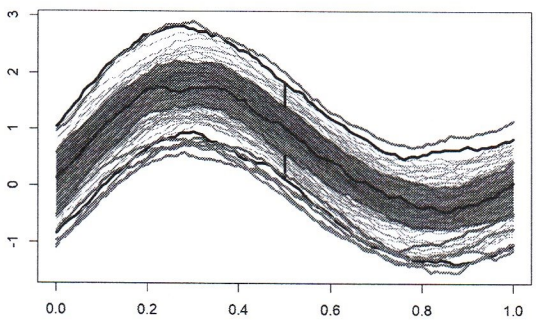
```
## [1] 0.5013081
```

```
# functional boxplot
# it works only with fdata....
fbplot(fdata)
```



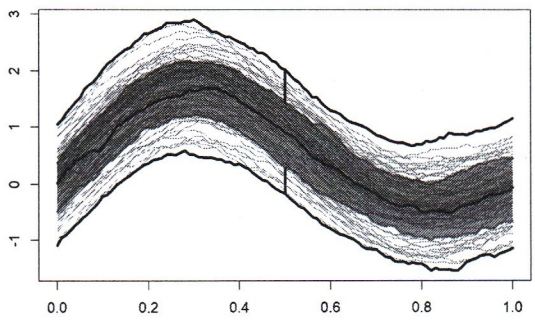
```
## $Depth
##   [1] 0.05719172 0.49653165 0.15976998 0.41158716 0.10031203 0.50471647
##   [7] 0.21503150 0.29264326 0.49960396 0.37481948 0.48817082 0.48499850
##  [13] 0.11867387 0.36584258 0.35317332 0.32794679 0.19312931 0.14647265
##  [19] 0.49721172 0.44727473 0.44336634 0.39178118 0.43681368 0.06240824
##  [25] 0.42253225 0.35573757 0.02623262 0.47286129 0.33650765 0.38392039
##  [31] 0.43922592 0.47504150 0.36330633 0.29880788 0.25878388 0.49344734
##  [37] 0.29682768 0.48142214 0.38613261 0.44941894 0.49735174 0.48808681
##  [43] 0.39169317 0.37477948 0.46695270 0.31890989 0.39058906 0.35276528
##  [49] 0.14329633 0.37268327 0.39447345 0.23693369 0.36784678 0.02997700
##  [55] 0.47637364 0.25546355 0.48403040 0.48865887 0.27104110 0.48541054
##  [61] 0.46270427 0.30912091 0.44830683 0.29656366 0.22154415 0.38610461
##  [67] 0.47769377 0.25992399 0.21277128 0.39889389 0.44842684 0.45694369
##  [73] 0.43167317 0.45565157 0.49539154 0.16483448 0.45467947 0.36253025
##  [79] 0.31687369 0.49307131 0.04170217 0.47351335 0.16186219 0.14755676
##  [85] 0.26747275 0.22022002 0.50583658 0.10202820 0.06994099 0.27106111
##  [91] 0.21666367 0.11248525 0.17698770 0.49314331 0.45107111 0.39074507
##  [97] 0.46877288 0.48143014 0.50100010 0.46247625
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## integer(0)
```

```
# the automated selection of F is actually pretty easy...
fbplot(fdata,adjust = list(VERBOSE=T))
```



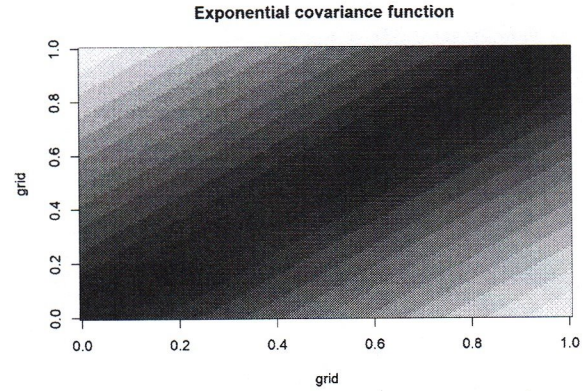

```
## $Depth
##   [1] 0.05719172 0.49653165 0.15976998 0.41158716 0.10031203 0.50471647
##   [7] 0.21503150 0.29264326 0.49960396 0.37481948 0.48817082 0.48499850
##  [13] 0.11867387 0.36584258 0.35317332 0.32794679 0.19312931 0.14647265
##  [19] 0.49721172 0.44727473 0.44336634 0.39178118 0.43681368 0.06240824
##  [25] 0.42253225 0.35573757 0.02623262 0.47286129 0.33650765 0.38392039
##  [31] 0.43922592 0.47504150 0.36330633 0.29880788 0.25878388 0.49344734
##  [37] 0.29682768 0.48142214 0.38613261 0.44941894 0.49735174 0.48808681
##  [43] 0.39169317 0.37477948 0.46695270 0.31890989 0.39058906 0.35276528
##  [49] 0.14329633 0.37268327 0.39447345 0.23693369 0.36784678 0.02997700
##  [55] 0.47637364 0.25546355 0.48403040 0.48865887 0.27104110 0.48541054
##  [61] 0.46270427 0.30912091 0.44830683 0.29656366 0.22154415 0.38610461
##  [67] 0.47769377 0.25992399 0.21277128 0.39889389 0.44842684 0.45694369
##  [73] 0.43167317 0.45565157 0.49539154 0.16483448 0.45467947 0.36253025
##  [79] 0.31687369 0.49307131 0.04170217 0.47351335 0.16186219 0.14755676
##  [85] 0.26747275 0.22022002 0.50583658 0.10202820 0.06994099 0.27106111
##  [91] 0.21666367 0.11248525 0.17698770 0.49314331 0.45107111 0.39074507
##  [97] 0.46877288 0.48143014 0.50100010 0.46247625
##
## $Fvalue
## [1] 1.03026
##
## $ID_outliers
## [1] 13 27 54 81 84 89
```

```
fbplot(fdata,Depths = 'BD')
```

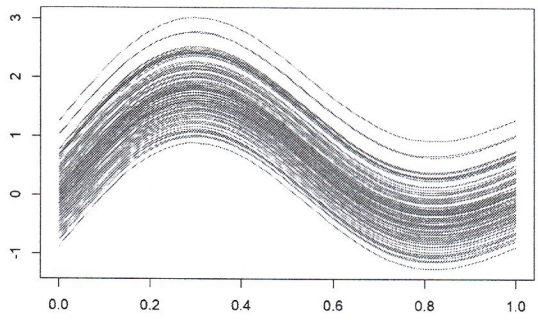



```
## $Depth
##   [1] 0.02000000 0.18808081 0.05232323 0.13555556 0.02000000 0.15131313
##   [7] 0.06666667 0.12040404 0.16666667 0.09494949 0.16424242 0.14525253
##  [13] 0.03616162 0.13919192 0.10000000 0.05151515 0.05030303 0.05151515
##  [19] 0.15636364 0.16080808 0.14444444 0.09070707 0.11111111 0.03878788
##  [25] 0.13515152 0.07939394 0.02000000 0.15737374 0.10060606 0.17232323
##  [31] 0.19090909 0.19373737 0.09636364 0.05636364 0.03171717 0.16848485
##  [37] 0.03818182 0.10404040 0.12303030 0.06969697 0.15818182 0.18000000
##  [43] 0.10000000 0.11494949 0.13313131 0.11757576 0.05070707 0.09494949
##  [49] 0.05313131 0.09272727 0.12303030 0.06424242 0.15777778 0.02000000
##  [55] 0.09555556 0.07737374 0.09919192 0.13111111 0.10909091 0.11616162
##  [61] 0.18363636 0.07959596 0.12242424 0.07959596 0.08222222 0.10909091
##  [67] 0.13515152 0.09636364 0.07898990 0.07979798 0.13555556 0.08888889
##  [73] 0.06242424 0.17111111 0.20383838 0.05070707 0.19656566 0.14666667
##  [79] 0.10343434 0.19191919 0.02000000 0.14606061 0.05232323 0.03636364
##  [85] 0.07656566 0.04585859 0.15010101 0.03696970 0.02000000 0.10606061
##  [91] 0.08060606 0.03737374 0.02000000 0.10888889 0.08888889 0.08060606
##  [97] 0.17353535 0.15656566 0.17151515 0.17272727
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## integer(0)
```

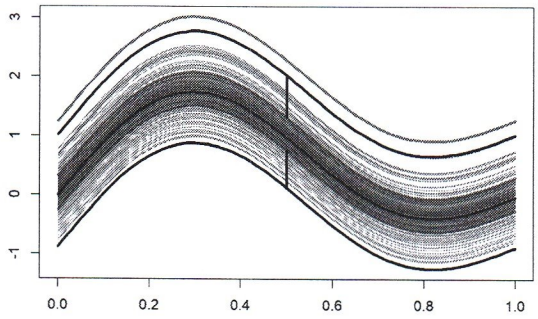
```
# BD goes bonkers with very rough data... the situation is way better with smoother data...
psi2 = exp_cov_function( grid, alpha, 0.001 )
image(psi2,
      main = 'Exponential covariance function',
      xlab = 'grid', ylab = 'grid')
```




```
data_smooth  = generate_gauss_fdata(n,mu,Cov=psi2)
fdata_smooth = fData(grid,data_smooth)
plot(fdata_smooth)
```

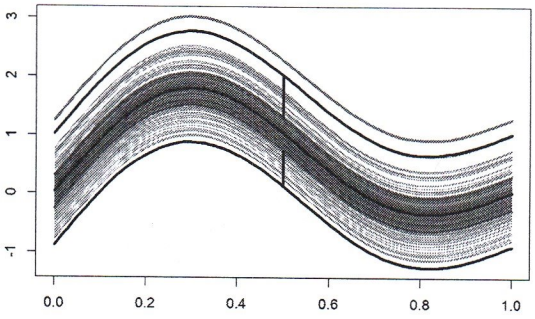



```
fbplot(fdata_smooth,Depths = 'BD')
```



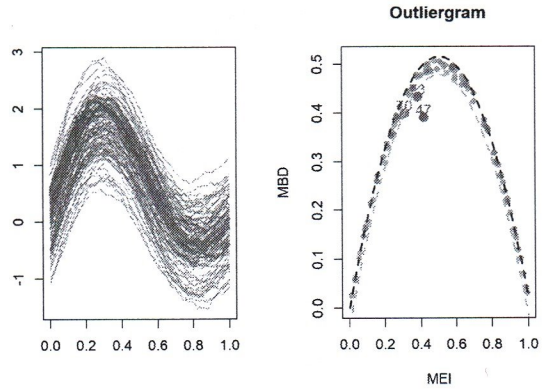

```
## $Depth
##   [1] 0.29818182 0.30363636 0.22848485 0.46444444 0.30363636 0.44181818
##   [7] 0.41030303 0.43212121 0.18363636 0.07696970 0.42080808 0.37878788
##  [13] 0.28787879 0.28828283 0.44989899 0.07636364 0.02000000 0.25757576
##  [19] 0.14303030 0.29818182 0.07696970 0.44545455 0.43333333 0.46444444
##  [25] 0.44585859 0.38363636 0.16707071 0.47555556 0.42424242 0.21555556
##  [31] 0.14585859 0.24323232 0.38363636 0.41595960 0.28181818 0.43898990
##  [37] 0.14585859 0.38868687 0.44727273 0.44585859 0.35777778 0.45636364
##  [43] 0.46606061 0.44666667 0.30363636 0.03939394 0.03959596 0.44020202
##  [49] 0.25757576 0.42080808 0.44666667 0.46626263 0.42727273 0.44666667
##  [55] 0.49272727 0.45555556 0.43333333 0.48060606 0.44424242 0.02000000
##  [61] 0.19575758 0.09515152 0.11393939 0.41454545 0.45838384 0.28505051
##  [67] 0.03959596 0.39636364 0.47252525 0.31555556 0.32707071 0.27454545
##  [73] 0.03939394 0.44222222 0.41595960 0.40727273 0.36909091 0.46444444
##  [79] 0.37393939 0.32707071 0.22848485 0.45434343 0.32323232 0.11292929
##  [85] 0.03939394 0.19979798 0.33818182 0.36383838 0.24323232 0.42424242
##  [91] 0.46444444 0.49272727 0.40464646 0.48060606 0.13151515 0.14585859
##  [97] 0.19575758 0.35777778 0.47454545 0.09595960
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## [1] 17
```

```
fbplot(fdata_smooth,Depths = 'MBD')
```



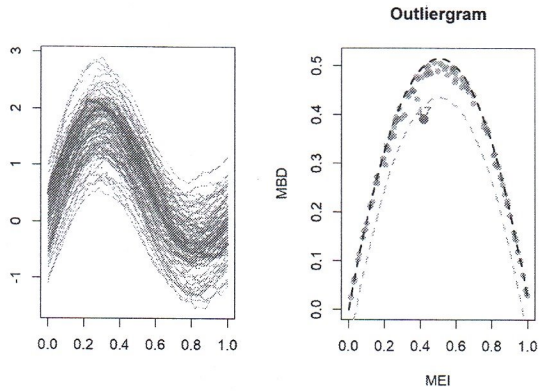
```
## $Depth
##   [1] 0.30187219 0.32881488 0.23712571 0.50402440 0.34214021 0.48990699
##   [7] 0.42924892 0.48085009 0.18363636 0.09563156 0.47257726 0.38383638
##  [13] 0.32463846 0.28828283 0.49708771 0.12141014 0.02000000 0.26432443
##  [19] 0.20050205 0.31057706 0.08165817 0.50346435 0.47564156 0.50482448
##  [25] 0.47176918 0.40422642 0.16707071 0.50596860 0.43490149 0.21555556
##  [31] 0.15787379 0.25896390 0.40797080 0.42270827 0.29937994 0.46854485
##  [37] 0.18069207 0.40199420 0.51432943 0.44585859 0.36513851 0.51427343
##  [43] 0.51449345 0.45993199 0.33586759 0.06224822 0.04421442 0.50791279
##  [49] 0.27062506 0.45847985 0.51328133 0.51463346 0.46606461 0.45319932
##  [55] 0.51224122 0.51262926 0.44621862 0.49233123 0.48417042 0.02000000
##  [61] 0.21595960 0.12053805 0.12392639 0.45834383 0.48227023 0.29448745
##  [67] 0.05477548 0.42979698 0.49721572 0.34038804 0.35309731 0.27454545
##  [73] 0.05343134 0.50049605 0.42396840 0.41395940 0.39480148 0.50557256
##  [79] 0.39497750 0.35946595 0.23964196 0.51132913 0.32947095 0.13255526
##  [85] 0.06149215 0.19979798 0.37187119 0.38476848 0.24729873 0.43794179
##  [91] 0.50338034 0.51199720 0.48984698 0.48867887 0.14270427 0.17390339
##  [97] 0.21813981 0.37021502 0.51302930 0.09874787
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## [1] 17
```

```
outliergram(fdata)
```



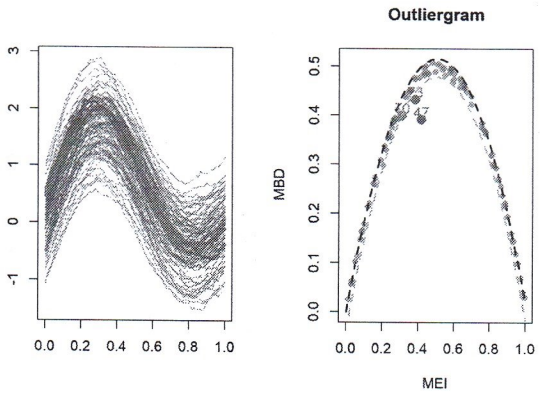

```
## $Fvalue
## [1] 1.5
##
## $d
##   [1] 8.593731e-04 6.030346e-03 1.131004e-03 4.531859e-03 6.315879e-04
##   [6] 1.020862e-02 4.820957e-03 1.630064e-03 7.622941e-03 1.155672e-02
##  [11] 2.093611e-02 1.318472e-02 2.481991e-03 1.506408e-03 1.428380e-02
##  [16] 2.455457e-02 4.293974e-03 1.585307e-03 1.490597e-02 1.384166e-02
##  [21] 8.286096e-03 3.205782e-02 1.812443e-02 1.976039e-04 1.904357e-02
##  [26] 2.993107e-02 8.373115e-05 2.480707e-02 5.388064e-03 3.060742e-03
##  [31] 2.044521e-03 6.515543e-03 8.762896e-03 2.037293e-02 1.479451e-02
##  [36] 1.072266e-02 3.388557e-02 2.416285e-02 7.015989e-03 2.414400e-02
##  [41] 6.194441e-03 1.634520e-02 1.648414e-02 1.400790e-02 3.255387e-02
##  [46] 2.010696e-03 1.088825e-01 1.011630e-02 6.584025e-04 7.771747e-03
##  [51] 6.313107e-03 7.462766e-03 1.442481e-03 7.050210e-05 1.896079e-02
##  [56] 3.945424e-03 2.641627e-02 2.611227e-02 1.983802e-03 1.822582e-02
##  [61] 4.686013e-03 8.812446e-03 1.865260e-02 1.150365e-02 5.974459e-04
##  [66] 1.411052e-02 4.579983e-03 4.581805e-03 2.013706e-03 4.068565e-02
##  [71] 2.456119e-02 2.563484e-02 5.329230e-02 1.320120e-02 6.113047e-03
##  [76] 4.391093e-03 9.717407e-03 3.518094e-03 1.056288e-02 9.364818e-03
##  [81] 2.276267e-04 1.145212e-02 5.445297e-04 3.393844e-03 5.136632e-03
##  [86] 1.188792e-02 7.616326e-03 9.051994e-04 5.807709e-04 4.449475e-03
##  [91] 3.341681e-03 3.251810e-04 1.084770e-02 2.015073e-02 1.993261e-02
##  [96] 1.023164e-02 1.101128e-02 2.942476e-02 1.344428e-02 9.583691e-03
##
## $ID_outliers
## [1] 47 70 73
```

```
outliergram(fdata, adjust=list(VERBOSE=T))
```



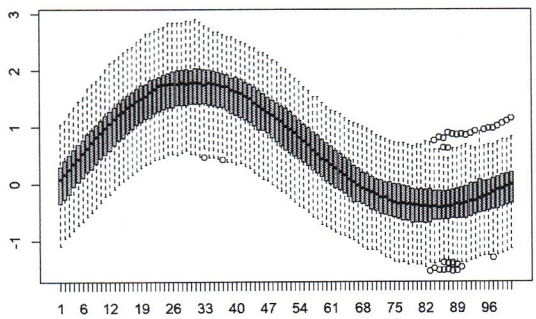

```
## $Fvalue
## [1] 4.712206
##
## $d
##   [1] 8.593731e-04 6.030346e-03 1.131004e-03 4.531859e-03 6.315879e-04
##   [6] 1.020862e-02 4.820957e-03 1.630064e-03 7.622941e-03 1.155672e-02
##  [11] 2.093611e-02 1.318472e-02 2.481991e-03 1.506408e-03 1.428380e-02
##  [16] 2.455457e-02 4.293974e-03 1.585307e-03 1.490597e-02 1.384166e-02
##  [21] 8.286096e-03 3.205782e-02 1.812443e-02 1.976039e-04 1.904357e-02
##  [26] 2.993107e-02 8.373115e-05 2.480707e-02 5.388064e-03 3.060742e-03
##  [31] 2.044521e-03 6.515543e-03 8.762896e-03 2.037293e-02 1.479451e-02
##  [36] 1.072266e-02 3.388557e-02 2.416285e-02 7.015989e-03 2.414400e-02
##  [41] 6.194441e-03 1.634520e-02 1.648414e-02 1.400790e-02 3.255387e-02
##  [46] 2.010696e-03 1.088825e-01 1.011630e-02 6.584025e-04 7.771747e-03
##  [51] 6.313107e-03 7.462766e-03 1.442481e-03 7.050210e-05 1.896079e-02
##  [56] 3.945424e-03 2.641627e-02 2.611227e-02 1.983802e-03 1.822582e-02
##  [61] 4.686013e-03 8.812446e-03 1.865260e-02 1.150365e-02 5.974459e-04
##  [66] 1.411052e-02 4.579983e-03 4.581805e-03 2.013706e-03 4.068565e-02
##  [71] 2.456119e-02 2.563484e-02 5.329230e-02 1.320120e-02 6.113047e-03
##  [76] 4.391093e-03 9.717407e-03 3.518094e-03 1.056288e-02 9.364818e-03
##  [81] 2.276267e-04 1.145212e-02 5.445297e-04 3.393844e-03 5.136632e-03
##  [86] 1.188792e-02 7.616326e-03 9.051994e-04 5.807709e-04 4.449475e-03
##  [91] 3.341681e-03 3.251810e-04 1.084770e-02 2.015073e-02 1.993261e-02
##  [96] 1.023164e-02 1.101128e-02 2.942476e-02 1.344428e-02 9.583691e-03
##
## $ID_outliers
## [1] 47
```

```
outliergram(fdata,p_check = 0.1)
```

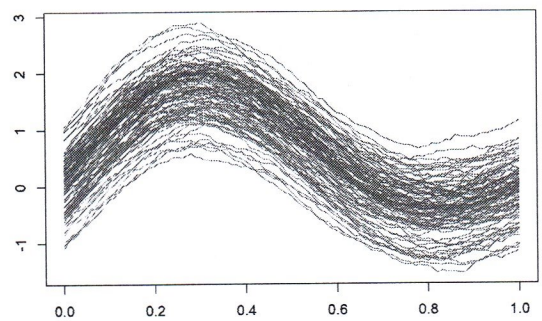


```
## $Fvalue
## [1] 1.5
##
## $d
##   [1] 8.593731e-04 6.030346e-03 1.131004e-03 4.531859e-03 6.315879e-04
##   [6] 1.020862e-02 4.820957e-03 1.630064e-03 7.622941e-03 1.155672e-02
##  [11] 2.093611e-02 1.318472e-02 2.481991e-03 1.506408e-03 1.428380e-02
##  [16] 2.455457e-02 4.293974e-03 1.585307e-03 1.490597e-02 1.384166e-02
##  [21] 8.286096e-03 3.205782e-02 1.812443e-02 1.976039e-04 1.904357e-02
##  [26] 2.993107e-02 8.373115e-05 2.480707e-02 5.388064e-03 3.060742e-03
##  [31] 2.044521e-03 6.515543e-03 8.762896e-03 2.037293e-02 1.479451e-02
##  [36] 1.072266e-02 3.388557e-02 2.416285e-02 7.015989e-03 2.414400e-02
##  [41] 6.194441e-03 1.634520e-02 1.648414e-02 1.400790e-02 3.255387e-02
##  [46] 2.010696e-03 1.088825e-01 1.011630e-02 6.584025e-04 7.771747e-03
##  [51] 6.313107e-03 7.462766e-03 1.442481e-03 7.050210e-05 1.896079e-02
##  [56] 3.945424e-03 2.641627e-02 2.611227e-02 1.983802e-03 1.822582e-02
##  [61] 4.686013e-03 8.812446e-03 1.865260e-02 1.150365e-02 5.974459e-04
##  [66] 1.411052e-02 4.579983e-03 4.581805e-03 2.013706e-03 4.068565e-02
##  [71] 2.456119e-02 2.563484e-02 5.329230e-02 1.320120e-02 6.113047e-03
##  [76] 4.391093e-03 9.717407e-03 3.518094e-03 1.056288e-02 9.364818e-03
##  [81] 2.276267e-04 1.145212e-02 5.445297e-04 3.393844e-03 5.136632e-03
##  [86] 1.188792e-02 7.616326e-03 9.051994e-04 5.807709e-04 4.449475e-03
##  [91] 3.341681e-03 3.251810e-04 1.084770e-02 2.015073e-02 1.993261e-02
##  [96] 1.023164e-02 1.101128e-02 2.942476e-02 1.344428e-02 9.583691e-03
##
## $ID_outliers
## [1] 47 70 73
```

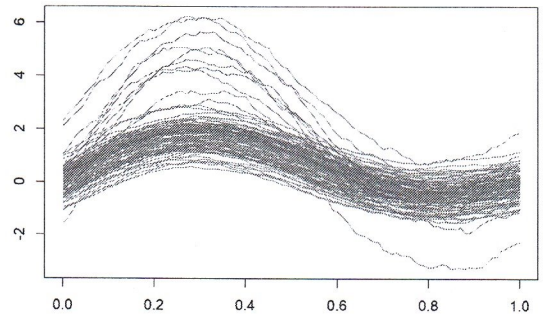
```
boxplot(data)
```



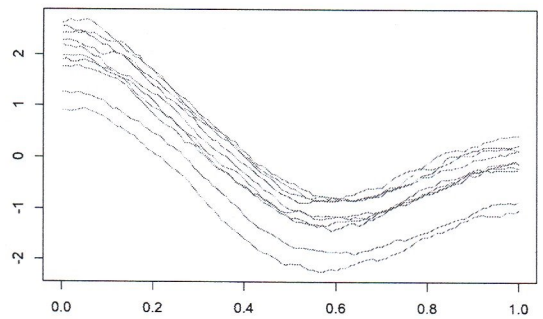

```
### ---------------------------------------------------------------------------
### ---------------------------------------------------------------------------
### EXAMPLE 2 - Functional outlier detection
### ---------------------------------------------------------------------------
### ---------------------------------------------------------------------------
outlier_share = .10
n_act         = n*outlier_share

# Let's generate 2 cases...
plot(fdata)
```



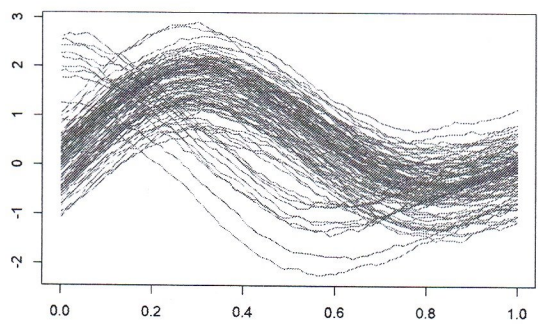

```
# magnitude outliers
# let's inflate the last 10 obs by a number generated from a unif(2,3)
fdata_temp = fdata[1:90,]
mag_temp   = fdata[91:100,] * runif(10,2,3)
fdata_mag  = append_fData(fdata_temp,mag_temp)
plot(fdata_mag)
```



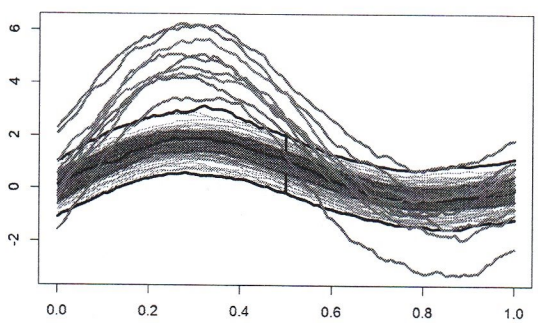

```
# shape_outlier
shift_q  = 1.5
mu_warp  = mu = sin(pi*grid+shift_q)+sin(2*pi*grid+shift_q)
shp_temp = generate_gauss_fdata(10,mu_warp, Cov=psi)
shp_temp = fData(grid,shp_temp)
plot(shp_temp)
```




```
fdata_shape = append_fData(fdata_temp,shp_temp)
plot(fdata_shape)
```

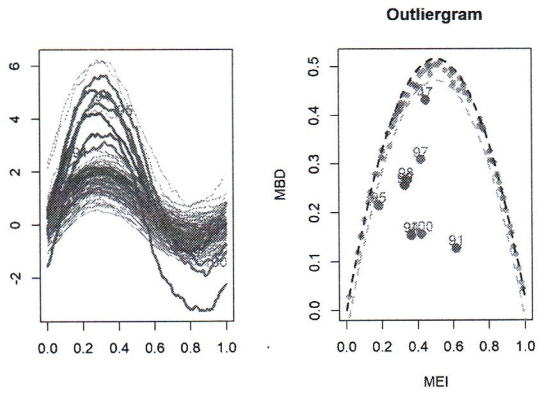



```
# now, boxplots!
fbplot(fdata_mag)
```



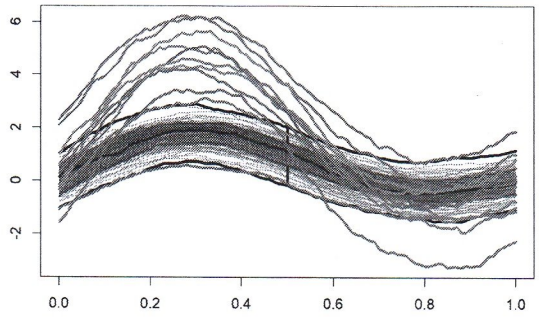
```
## $Depth
##   [1] 0.15155716 0.49304730 0.17154715 0.44284228 0.11557356 0.50774077
##   [7] 0.22760076 0.30161216 0.50522052 0.40827083 0.48219822 0.48753475
##  [13] 0.13126313 0.37234723 0.40059806 0.37719972 0.20371037 0.21928793
##  [19] 0.49156716 0.44595060 0.46281628 0.42413641 0.46461246 0.15279328
##  [25] 0.42972897 0.40783878 0.11814981 0.46549655 0.37640364 0.39155316
##  [31] 0.44165017 0.47504150 0.39858586 0.30673667 0.31079308 0.49347135
##  [37] 0.33600360 0.49284328 0.41968797 0.46092809 0.50079208 0.48013001
##  [43] 0.42024002 0.37634763 0.45936394 0.32662666 0.43114511 0.38872887
##  [49] 0.15804180 0.37934393 0.42494449 0.24604660 0.37410741 0.04435044
##  [55] 0.49008701 0.26066407 0.50009201 0.50280028 0.32625063 0.50112411
##  [61] 0.46213621 0.36167817 0.45923592 0.30545255 0.22938094 0.41154315
##  [67] 0.48901090 0.31372537 0.27781378 0.44008601 0.45616762 0.45042704
##  [73] 0.45787179 0.46000000 0.48888689 0.24042604 0.45388339 0.36939094
##  [79] 0.32421442 0.49241124 0.05545555 0.48170217 0.23132513 0.15851385
##  [85] 0.27316932 0.27725773 0.50342434 0.18416842 0.08526253 0.27820182
##  [91] 0.12800280 0.02744874 0.10026403 0.15428143 0.21408341 0.22867287
##  [97] 0.30816482 0.26544054 0.25510751 0.15580958
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## [1]  91  92  93  94  95  96  98  99 100
```

```
outliergram(fdata_mag)
```



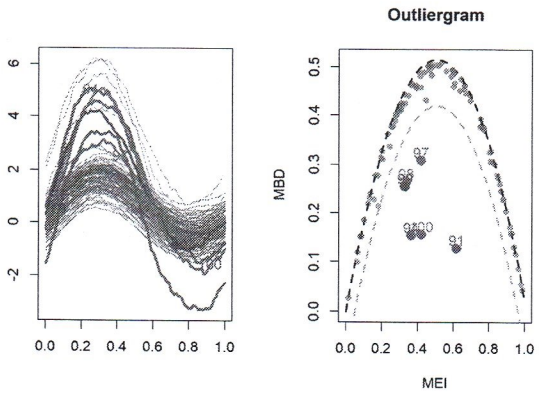

```
## $Fvalue
## [1] 1.5
##
## $d
##   [1] 2.105478e-03 6.000481e-03 1.309874e-03 4.854109e-03 4.433315e-04
##   [6] 5.820859e-03 7.046328e-03 2.079139e-03 7.162736e-03 1.793205e-02
##  [11] 3.092919e-02 6.679005e-03 2.681298e-03 2.053354e-03 7.671223e-03
##  [16] 1.673359e-02 2.689338e-03 4.830384e-03 1.740942e-02 1.503505e-02
##  [21] 1.108970e-02 2.829285e-02 1.115478e-02 3.065415e-03 1.237973e-02
##  [26] 1.463548e-02 3.006043e-03 2.964003e-02 1.158254e-02 1.300526e-03
##  [31] 3.763663e-03 3.945939e-03 1.503952e-02 2.397658e-02 2.064440e-02
##  [36] 5.576439e-03 4.196230e-02 1.913680e-02 8.682136e-03 2.867221e-02
##  [41] 9.879641e-03 2.140725e-02 2.011049e-02 1.914750e-02 3.759326e-02
##  [46] 2.755127e-03 7.659415e-02 1.769349e-02 3.424897e-04 6.931307e-03
##  [51] 1.206275e-02 5.042841e-03 2.291199e-03 2.856523e-04 1.549799e-02
##  [56] 3.277952e-03 1.386491e-02 1.186245e-02 3.494765e-03 9.188642e-03
##  [61] 4.823215e-03 5.237910e-03 2.660112e-02 1.447483e-02 9.277759e-04
##  [66] 2.276873e-02 6.988778e-03 9.913467e-03 3.174575e-03 2.386670e-02
##  [71] 3.343269e-02 3.035349e-02 4.063832e-02 5.655219e-03 8.408960e-03
##  [76] 2.504330e-03 1.301433e-02 5.220641e-03 1.403214e-02 5.071359e-03
##  [81] 3.096151e-04 1.607717e-02 3.321956e-03 4.727998e-03 3.304449e-03
##  [86] 1.952084e-02 1.154781e-02 3.666981e-03 1.250026e-03 3.874486e-03
##  [91] 3.616871e-01 4.741068e-05 1.859228e-02 3.200797e-01 9.077070e-02
##  [96] 3.738005e-02 1.916901e-01 1.912776e-01 1.974625e-01 3.447302e-01
##
## $ID_outliers
## [1]  47  91  94  95  97  98  99 100
```

```
fbplot(fdata_mag,adjust = list(VERBOSE=T))
```



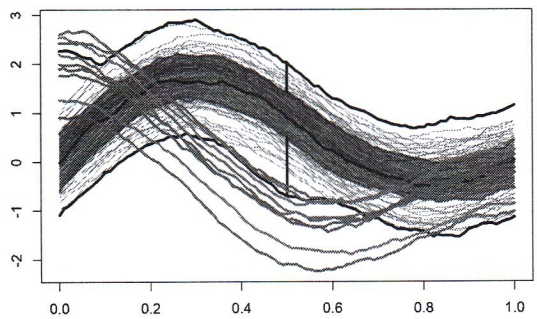

```
## $Depth
##   [1] 0.15155716 0.49304730 0.17154715 0.44284228 0.11557356 0.50774077
##   [7] 0.22760076 0.30161216 0.50522052 0.40827083 0.48219822 0.48753475
##  [13] 0.13126313 0.37234723 0.40059806 0.37719972 0.20371037 0.21928793
##  [19] 0.49156716 0.44595060 0.46281628 0.42413641 0.46461246 0.15279328
##  [25] 0.42972897 0.40783878 0.11814981 0.46549655 0.37640364 0.39155316
##  [31] 0.44165017 0.47504150 0.39858586 0.30673667 0.31079308 0.49347135
##  [37] 0.33600360 0.49284328 0.41968797 0.46092809 0.50079208 0.48013001
##  [43] 0.42024002 0.37634763 0.45936394 0.32662666 0.43114511 0.38872887
##  [49] 0.15804180 0.37934393 0.42494449 0.24604660 0.37410741 0.04435044
##  [55] 0.49008701 0.26066407 0.50009201 0.50280028 0.32625063 0.50112411
##  [61] 0.46213621 0.36167817 0.45923592 0.30545255 0.22938094 0.41154315
##  [67] 0.48901090 0.31372537 0.27781378 0.44008601 0.45616762 0.45042704
##  [73] 0.45787179 0.46000000 0.48888689 0.24042604 0.45388339 0.36939094
##  [79] 0.32421442 0.49241124 0.05545555 0.48170217 0.23132513 0.15851385
##  [85] 0.27316932 0.27725773 0.50342434 0.18416842 0.08526253 0.27820182
##  [91] 0.12800280 0.02744874 0.10026403 0.15428143 0.21408341 0.22867287
##  [97] 0.30816482 0.26544054 0.25510751 0.15580958
##
## $Fvalue
## [1] 1.09659
##
## $ID_outliers
##  [1]  54  91  92  93  94  95  96  97  98  99 100
```

```
outliergram(fdata_mag,adjust = list(VERBOSE=T))
```



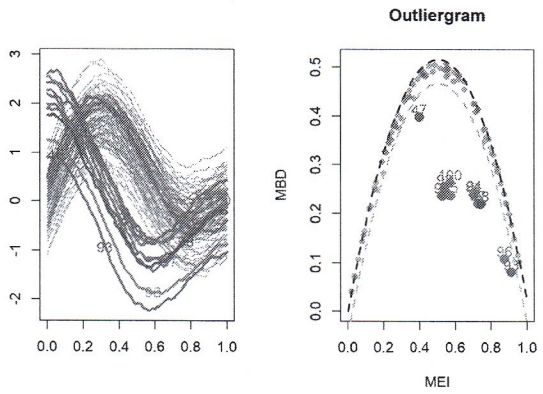
```
## $Fvalue
## [1] 4.786113
##
## $d
##   [1] 2.105478e-03 6.000481e-03 1.309874e-03 4.854109e-03 4.433315e-04
##   [6] 5.820859e-03 7.046328e-03 2.079139e-03 7.162736e-03 1.793205e-02
##  [11] 3.092919e-02 6.679005e-03 2.681298e-03 2.053354e-03 7.671223e-03
##  [16] 1.673359e-02 2.689338e-03 4.830384e-03 1.740942e-02 1.503505e-02
##  [21] 1.108970e-02 2.829285e-02 1.115478e-02 3.065415e-03 1.237973e-02
##  [26] 1.463548e-02 3.006043e-03 2.964003e-02 1.158254e-02 1.300526e-03
##  [31] 3.763663e-03 3.945939e-03 1.503952e-02 2.397658e-02 2.064440e-02
##  [36] 5.576439e-03 4.196230e-02 1.913680e-02 8.682136e-03 2.867221e-02
##  [41] 9.879641e-03 2.140725e-02 2.011049e-02 1.914750e-02 3.759326e-02
##  [46] 2.755127e-03 7.659415e-02 1.769349e-02 3.424897e-04 6.931307e-03
##  [51] 1.206275e-02 5.042841e-03 2.291199e-03 2.856523e-04 1.549799e-02
##  [56] 3.277952e-03 1.386491e-02 1.186245e-02 3.494765e-03 9.188642e-03
##  [61] 4.823215e-03 5.237910e-03 2.660112e-02 1.447483e-02 9.277759e-04
##  [66] 2.276873e-02 6.988778e-03 9.913467e-03 3.174575e-03 2.386670e-02
##  [71] 3.343269e-02 3.035349e-02 4.063832e-02 5.655219e-03 8.408960e-03
##  [76] 2.504330e-03 1.301433e-02 5.220641e-03 1.403214e-02 5.071359e-03
##  [81] 3.096151e-04 1.607717e-02 3.321956e-03 4.727998e-03 3.304449e-03
##  [86] 1.952084e-02 1.154781e-02 3.666981e-03 1.250026e-03 3.874486e-03
##  [91] 3.616871e-01 4.741068e-05 1.859228e-02 3.200797e-01 9.077070e-02
##  [96] 3.738005e-02 1.916901e-01 1.912776e-01 1.974625e-01 3.447302e-01
##
## $ID_outliers
## [1]  91  94  97  98  99 100
```

```
fbplot(fdata_shape)
```



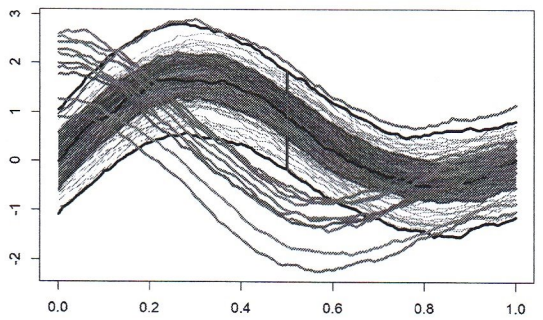

```
## $Depth
##   [1] 0.07050105 0.50606861 0.23906591 0.40226223 0.17753175 0.50388439
##   [7] 0.28343034 0.35293729 0.48890289 0.36656666 0.48005001 0.49738374
##  [13] 0.19776978 0.41149915 0.34385239 0.32451845 0.26013601 0.14443644
##  [19] 0.49661166 0.48224222 0.43004500 0.37604360 0.43358936 0.07294129
##  [25] 0.45358336 0.35114911 0.04039404 0.48383438 0.32731473 0.42448045
##  [31] 0.47070107 0.48886289 0.35546155 0.34881688 0.24731473 0.49417142
##  [37] 0.28117412 0.48345435 0.37953995 0.45227523 0.49288729 0.49447545
##  [43] 0.37776778 0.42460046 0.47997000 0.37322332 0.39789779 0.34440844
##  [49] 0.21943994 0.40831083 0.38662066 0.30008801 0.41280328 0.10208021
##  [55] 0.47447345 0.31988999 0.48220222 0.48980298 0.26327633 0.48444244
##  [61] 0.48232223 0.30547655 0.43806581 0.35717772 0.29081908 0.38113211
##  [67] 0.46573657 0.25105511 0.20935094 0.39526953 0.43893789 0.47083708
##  [73] 0.43519352 0.47448145 0.50108811 0.16509451 0.48456646 0.41343534
##  [79] 0.37215522 0.49712771 0.11564156 0.46379238 0.15675768 0.22115612
##  [85] 0.32416242 0.21472747 0.50144814 0.10674467 0.14790479 0.33596760
##  [91] 0.23880988 0.25293129 0.08082608 0.24095810 0.23640564 0.10812881
##  [97] 0.22018002 0.21910391 0.23538554 0.26285229
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## [1]  91  92  93  94  95  96  97  98 100
```

```
outliergram(fdata_shape)
```



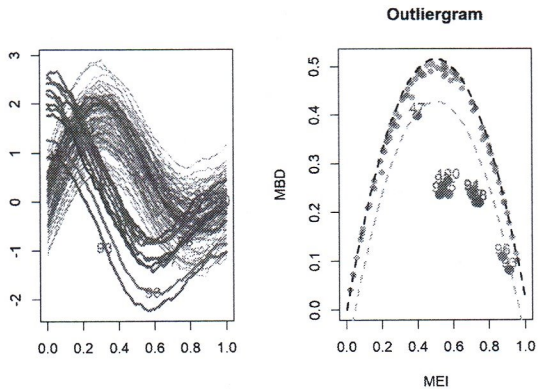

```
## $Fvalue
## [1] 1.5
##
## $d
##   [1] 0.001145857 0.004876408 0.001805289 0.008384363 0.004482468 0.010381751
##   [7] 0.008323446 0.003174416 0.012764682 0.014404054 0.023496171 0.010710814
##  [13] 0.001553977 0.004708352 0.017929476 0.019129319 0.010111981 0.004132809
##  [19] 0.018215208 0.004595430 0.013725848 0.040517438 0.011527173 0.001512230
##  [25] 0.017479372 0.026608324 0.001535758 0.025207550 0.007841893 0.007658390
##  [31] 0.002344631 0.009755471 0.010912497 0.027476926 0.022654701 0.016350902
##  [37] 0.042587823 0.017051527 0.007441180 0.012248195 0.004632542 0.018133100
##  [43] 0.023588775 0.010110674 0.030419517 0.005574815 0.095039009 0.014384409
##  [49] 0.005891480 0.018227209 0.010039182 0.011675583 0.005428424 0.003738869
##  [55] 0.014721551 0.004486865 0.023680784 0.023299993 0.006274964 0.012363217
##  [61] 0.007680174 0.007047635 0.021735995 0.011573870 0.002527659 0.014070952
##  [67] 0.007539407 0.008496137 0.004816878 0.033183398 0.025585529 0.029293662
##  [73] 0.043419075 0.014668318 0.008608267 0.005453813 0.006361705 0.002506587
##  [79] 0.012092496 0.012629976 0.003037571 0.013081902 0.003305400 0.005234702
##  [85] 0.010635955 0.016781361 0.008806623 0.002487972 0.004233572 0.003931680
##  [91] 0.197409365 0.259704703 0.101350531 0.193193814 0.269288275 0.131414528
##  [97] 0.196204333 0.180347104 0.278769976 0.243819115
##
## $ID_outliers
##  [1]  47  91  92  93  94  95  96  97  98  99 100
```

```
fbplot(fdata_shape,adjust = list(VERBOSE=T))
```



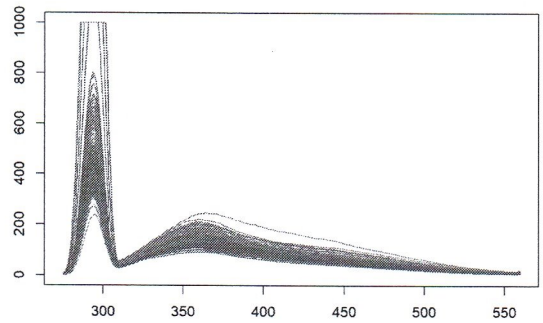
```
## $Depth
##   [1] 0.07050105 0.50606861 0.23906591 0.40226223 0.17753175 0.50388439
##   [7] 0.28343034 0.35293729 0.48890289 0.36656666 0.48005001 0.49738374
##  [13] 0.19776978 0.41149915 0.34385239 0.32451845 0.26013601 0.14443644
##  [19] 0.49661166 0.48224222 0.43004500 0.37604360 0.43358936 0.07294129
##  [25] 0.45358336 0.35114911 0.04039404 0.48383438 0.32731473 0.42448045
##  [31] 0.47070107 0.48886289 0.35546155 0.34881688 0.24731473 0.49417142
##  [37] 0.28117412 0.48345435 0.37953995 0.45227523 0.49288729 0.49447545
##  [43] 0.37776778 0.42460046 0.47997000 0.37322332 0.39789779 0.34440844
##  [49] 0.21943994 0.40831083 0.38662066 0.30008801 0.41280328 0.10208021
##  [55] 0.47447345 0.31988999 0.48220222 0.48980298 0.26327633 0.48444244
##  [61] 0.48232223 0.30547655 0.43806581 0.35717772 0.29081908 0.38113211
##  [67] 0.46573657 0.25105511 0.20935094 0.39526953 0.43893789 0.47083708
##  [73] 0.43519352 0.47448145 0.50108811 0.16509451 0.48456646 0.41343534
##  [79] 0.37215522 0.49712771 0.11564156 0.46379238 0.15675768 0.22115612
##  [85] 0.32416242 0.21472747 0.50144814 0.10674467 0.14790479 0.33596760
##  [91] 0.23880988 0.25293129 0.08082608 0.24095810 0.23640564 0.10812881
##  [97] 0.22018002 0.21910391 0.23538554 0.26285229
##
## $Fvalue
## [1] 1.042686
##
## $ID_outliers
##  [1]  27  91  92  93  94  95  96  97  98  99 100
```

```
outliergram(fdata_shape,adjust = list(VERBOSE=T))
```

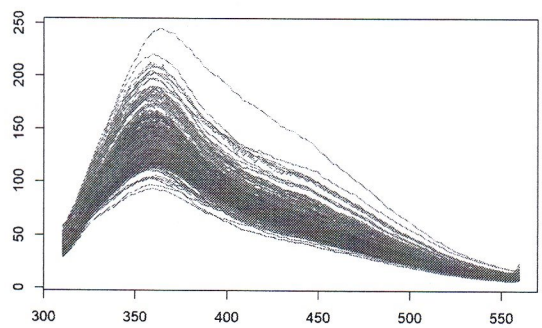



```
## $Fvalue
## [1] 3.765121
##
## $d
##   [1] 0.001145857 0.004876408 0.001805289 0.008384363 0.004482468 0.010381751
##   [7] 0.008323446 0.003174416 0.012764682 0.014404054 0.023496171 0.010710814
##  [13] 0.001553977 0.004708352 0.017929476 0.019129319 0.010111981 0.004132809
##  [19] 0.018215208 0.004595430 0.013725848 0.040517438 0.011527173 0.001512230
##  [25] 0.017479372 0.026608324 0.001535758 0.025207550 0.007841893 0.007658390
##  [31] 0.002344631 0.009755471 0.010912497 0.027476926 0.022654701 0.016350902
##  [37] 0.042587823 0.017051527 0.007441180 0.012248195 0.004632542 0.018133100
##  [43] 0.023588775 0.010110674 0.030419517 0.005574815 0.095039009 0.014384409
##  [49] 0.005891480 0.018227209 0.010039182 0.011675583 0.005428424 0.003738869
##  [55] 0.014721551 0.004486865 0.023680784 0.023299993 0.006274964 0.012363217
##  [61] 0.007680174 0.007047635 0.021735995 0.011573870 0.002527659 0.014070952
##  [67] 0.007539407 0.008496137 0.004816878 0.033183398 0.025585529 0.029293662
##  [73] 0.043419075 0.014668318 0.008608267 0.005453813 0.006361705 0.002506587
##  [79] 0.012092496 0.012629976 0.003037571 0.013081902 0.003305400 0.005234702
##  [85] 0.010635955 0.016781361 0.008806623 0.002487972 0.004233572 0.003931680
##  [91] 0.197409365 0.259704703 0.101350531 0.193193814 0.269288275 0.131414528
##  [97] 0.196204333 0.180347104 0.278769976 0.243819115
##
## $ID_outliers
##  [1]  47  91  92  93  94  95  96  97  98  99 100
```

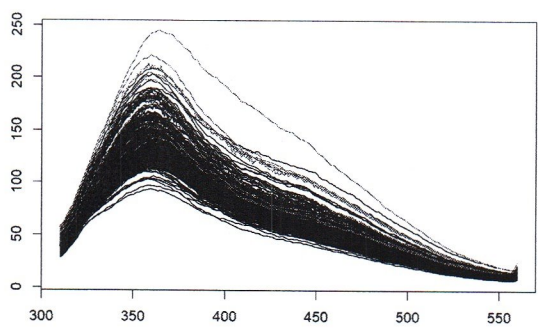
```
### ------------------------------------------------------------------------------
### ------------------------------------------------------------------------------
### EXAMPLE 3 - Sugar Emission Spectra
### ------------------------------------------------------------------------------
### ------------------------------------------------------------------------------
# In chemometrics there are often function-like absorbance or emission spectra given
# - in particular for food samples - that are used to determine the content of certain ingredients.
# Using the spectra is typically much cheaper than alternative chemical analysis.

# 268 samples of sugar were dissolved and the solution was measured spectrofluorometrically.
# For every sample the emission spectra from 275-560 nm were measured in 0.5 nm intervals
# (i.e., at 571 wavelengths) at seven excitation wavelengths: 230, 240, 255, 290, 305, 325, and 340 nm.
# In addition, there are laboratory determinations of the quality of the sugar given, such as ash
# content (in percentage). Ash content measures the amount of inorganic impurities in
# the refined sugar, cf. Bro (1999)

# let's focus our attention on the emission spectra at 290 nm
# they come in matrix form, let's convert them in fdata objects...
load("C:/Users/utente/Documents/Sugar.RDA")
grid      = seq(275,560,by=.5)
fdata_290 = fData(grid,wave_290)
plot(fdata_290)
```



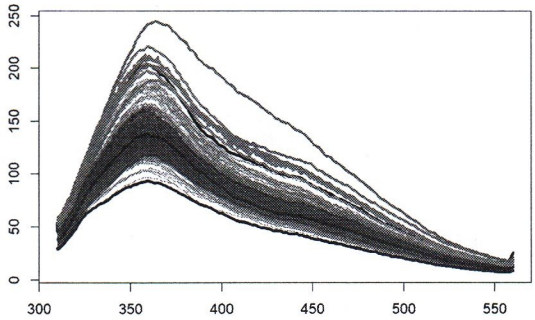

```
# it is evident how the spectrofluorimeter goes to saturation for wavelenghts < ~310 nm.
# let's cut the saturation area...
fdata_290_cut = fdata_290[,71:571]
plot(fdata_290_cut)
```




```
# let's see if we can identify a pattern in the curves, related to the ash content...
color = outcome_data$ash>20
plot(fdata_290_cut,col=color+1)
```

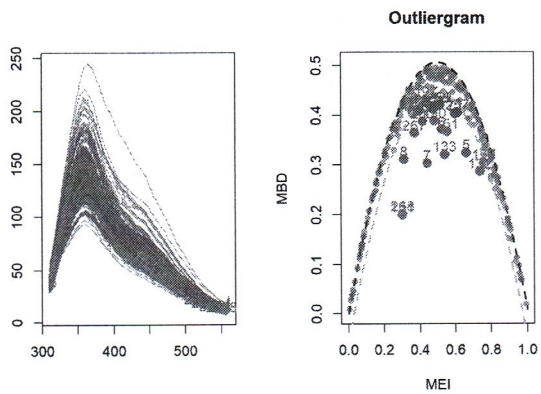


```
# somewhat...
# let's see boxplots and outliergrams
fbplot(fdata_290_cut)
```




```
## $Depth
##        2023        2031        2032        2033        2041        2042
## 0.347198052 0.172635834 0.066582805 0.251584315 0.323538233 0.370953939
##        2043        2051        2052        2053        2061        2062
##          ..
##        5011        5012        5013        5021        5022        5023
## 0.419325472 0.382354749 0.333122954 0.280311198 0.350782476 0.312747416
##        5031        5032        5033        5041
## 0.463267160 0.469151361 0.363525172 0.201229215
##
## $Fvalue
## [1] 1.5
##
## $ID_outliers
## 2053 2072 2081 2082 2091 2153 2271 3161
##   10   14   16   17   19   38   71  129
```

```
outliergram(fdata_290_cut)
```




```
## $Fvalue
## [1] 1.5
##
## $d
##        2023        2031        2032        2033        2041        2042
## 3.844280e-02 1.130550e-02 1.903999e-03 4.480398e-02 1.313059e-01 1.337221e-01
##        2043        2051        2052        2053        2061        2062
##          ..
##        5011        5012        5013        5021        5022        5023
## 8.594042e-02 5.477261e-02 3.372193e-02 2.714826e-02 2.479939e-02 6.857332e-03
##        5031        5032        5033        5041
## 3.314307e-02 2.401248e-02 1.077804e-01 2.253242e-01
##
## $ID_outliers
## 2041 2042 2043 2051 2272 3033 3072 3092 3183 3191 3192 3252 3253 4042 4043 4083
##    5    6    7    8   72   93  104  109  131  132  133  150  151  177  178  190
## 4111 4112 4121 4122 4312 4313 5011 5033 5041
##  197  198  200  201  257  258  259  267  268
```