

Bioreactor time-course analysis

Olivier Chapleur

Contents

1	Data	1
2	Data preprocessing	1
3	Spline smoothing	2
4	Filtering of the obtained profiles	3
4.1	OTUs	3
4.2	Metabolites	6
5	sPLS	8
6	block sPLS	11

1 Data

In this data set, **three bioreactors** with similar performances were considered as replicates. Different parameters were measured accross time in the three bioreactors.

Performance data: Based on chemical measurement, the time course evolution of a set of parameters was measured (CH₄, CO₂, acetate, propionate).

Metabolites data: The time course evolution of 20 selected metabolites was measured with GCMS.

Microbial data: DNA from samples taken across time was extracted and sequenced. (16S metabarcoding).

2 Data preprocessing

Metabolites (GCMS) data are log transformed.

Microbial data

- 1) are filtered (only OTUs with at least 1% of abundance in at least 1 sample are kept = 51 OTUs).
- 2) a count of 1 sequence is added to each sample/OTU (to avoid 0 in the datamatrix)
- 3) relative abundance is calculated
- 4) obtain data is clr transformed

Performance data is not transformed.

Ther are 51 OTUs after 0.01 % filter

3 Spline smoothing

All the data are modelled with spline smoothing.

```
## Data-driven Linear Mixed-Effect Model Splines
## Profiles were modelled for 20 features with 48 time points.
##
## Basis:
## [1] "p-spline"
##
## Knots:
##
## [1] 17.57143 26.14286 32.71429 38.28571 44.14286 50.57143
##
## Time points:
##
## [1] 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
## [24] 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55
## [47] 56 57
##
## Table of models used to model profiles:
## 0 2 3
## 10 4 6
##
## Profiles not modelled:
## [1] "All features were modelled"

## Data-driven Linear Mixed-Effect Model Splines
## Profiles were modelled for 51 features with 48 time points.
##
## Basis:
## [1] "p-spline"
##
## Knots:
##
## [1] 17.57143 26.14286 32.71429 38.28571 44.14286 50.57143
##
## Time points:
##
## [1] 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
## [24] 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55
## [47] 56 57
##
## Table of models used to model profiles:
## 0 1 2
## 30 19 2
##
## Profiles not modelled:
## [1] "All features were modelled"

## Data-driven Linear Mixed-Effect Model Splines
## Profiles were modelled for 2 features with 48 time points.
##
## Basis:
## [1] "p-spline"
```

```

##
## Knots:
##
## [1] 5.428571 13.000000 27.142857 44.142857 66.000000 113.000000
##
## Time points:
##
## [1] 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
## [24] 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55
## [47] 56 57
##
## Table of models used to model profiles:
## 1
## 2
##
## Profiles not modelled:
## [1] "All features were modelled"

## Data-driven Linear Mixed-Effect Model Splines
## Profiles were modelled for 2 features with 48 time points.
##
## Basis:
## [1] "p-spline"
##
## Knots:
##
## [1] 3.5 11.0 19.5 29.0 40.5 57.0 92.5
##
## Time points:
##
## [1] 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
## [24] 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55
## [47] 56 57
##
## Table of models used to model profiles:
## 3
## 2
##
## Profiles not modelled:
## [1] "All features were modelled"

```

4 Filtering of the obtained profiles

4.1 OTUs

```

## # A tibble: 51 x 3
## # Groups:   molecule [51]
##   molecule model_used MSE
##   <chr>      <fct>      <dbl>
## 1 OTU_1      0          0.0750
## 2 OTU_10     1          0.0235
## 3 OTU_107    0          3.42
## 4 OTU_11     0          0.0553

```

```

## 5 OTU_13 1 0.0534
## 6 OTU_130 1 0.377
## 7 OTU_14 0 0.0114
## 8 OTU_15 0 0.267
## 9 OTU_16 0 0.0820
## 10 OTU_169 1 0.162
## # ... with 41 more rows

## [1] "OTU_1" "OTU_10" "OTU_11" "OTU_13" "OTU_130" "OTU_14" "OTU_15"
## [8] "OTU_16" "OTU_169" "OTU_17" "OTU_18" "OTU_19" "OTU_2" "OTU_20"
## [15] "OTU_21" "OTU_24" "OTU_25" "OTU_26" "OTU_28" "OTU_29" "OTU_30"
## [22] "OTU_304" "OTU_31" "OTU_35" "OTU_38" "OTU_4" "OTU_41" "OTU_44"
## [29] "OTU_45" "OTU_46" "OTU_5" "OTU_50" "OTU_59" "OTU_6" "OTU_60"
## [36] "OTU_61" "OTU_65" "OTU_68" "OTU_7" "OTU_74" "OTU_75" "OTU_8"
## [43] "OTU_82" "OTU_97"

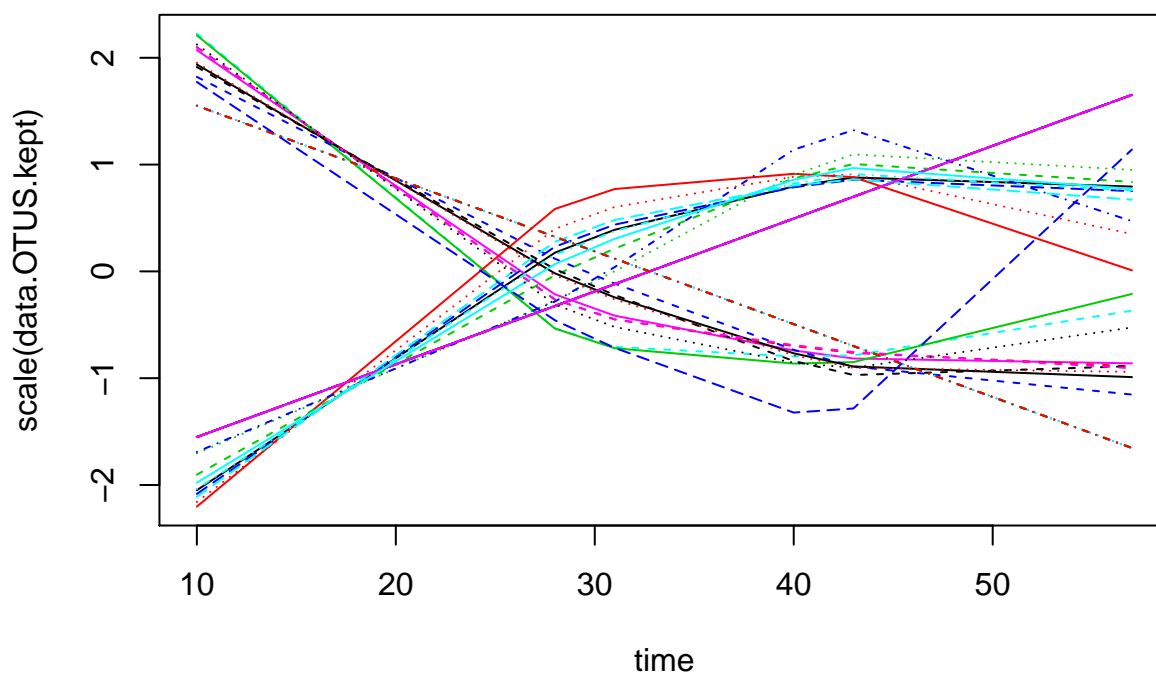
```

	molecule	MSE.filter	modelsUsed	BP.test
	OTU_2	TRUE	1	TRUE
	OTU_1	TRUE	0	TRUE
	OTU_4	TRUE	0	TRUE
	OTU_5	TRUE	0	TRUE
	OTU_6	TRUE	0	TRUE
	OTU_7	TRUE	2	TRUE
	OTU_8	TRUE	0	TRUE
	OTU_10	TRUE	1	TRUE
	OTU_11	TRUE	0	TRUE
	OTU_13	TRUE	1	TRUE
	OTU_14	TRUE	0	TRUE
	OTU_15	TRUE	0	TRUE
	OTU_16	TRUE	0	TRUE
	OTU_17	TRUE	0	TRUE
	OTU_18	TRUE	0	TRUE
	OTU_19	TRUE	0	TRUE
	OTU_20	TRUE	1	TRUE
	OTU_21	TRUE	0	TRUE
	OTU_22	FALSE	0	TRUE
	OTU_24	TRUE	0	TRUE
	OTU_25	TRUE	0	TRUE
	OTU_26	TRUE	1	TRUE
	OTU_28	TRUE	1	TRUE
	OTU_29	TRUE	0	TRUE
	OTU_30	TRUE	0	TRUE
	OTU_31	TRUE	0	TRUE
	OTU_33	TRUE	0	FALSE
	OTU_34	FALSE	0	TRUE
	OTU_35	TRUE	1	TRUE
	OTU_38	TRUE	1	TRUE
	OTU_41	TRUE	0	TRUE
	OTU_44	TRUE	1	TRUE
	OTU_45	TRUE	0	TRUE
	OTU_46	TRUE	1	TRUE
	OTU_50	TRUE	1	TRUE
	OTU_51	FALSE	0	TRUE
	OTU_59	TRUE	1	TRUE

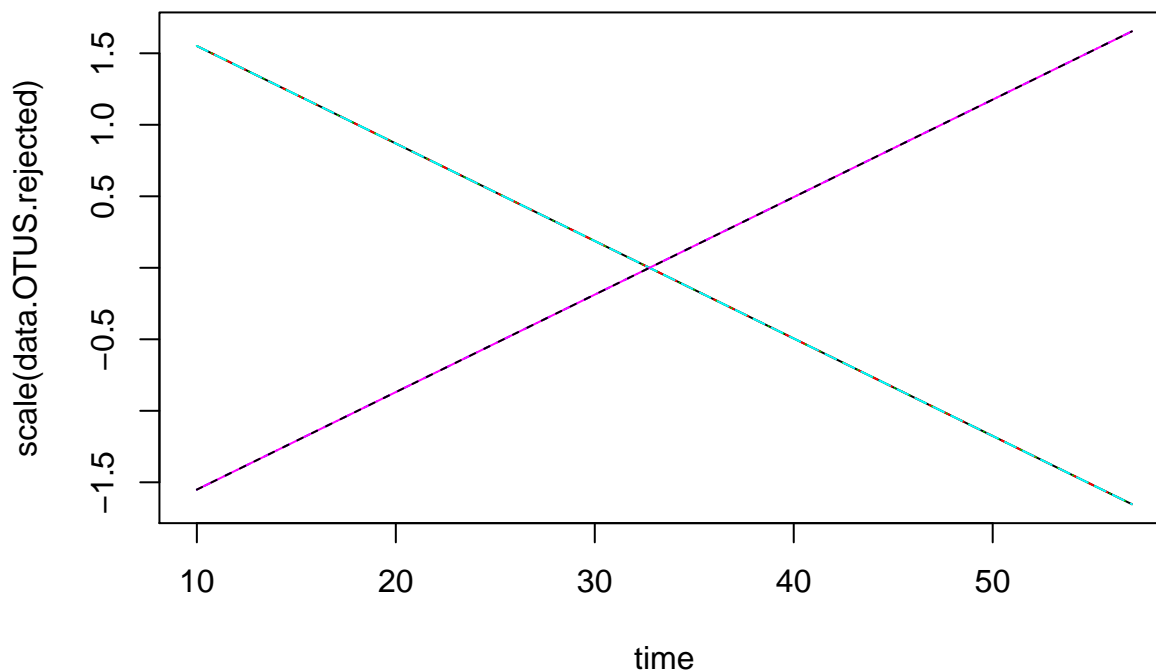
	molecule	MSE.filter	modelsUsed	BP.test
	OTU_60	TRUE	1	TRUE
	OTU_61	TRUE	2	TRUE
	OTU_65	TRUE	0	TRUE
	OTU_68	TRUE	1	TRUE
	OTU_74	TRUE	1	TRUE
	OTU_75	TRUE	0	TRUE
	OTU_82	TRUE	1	TRUE
	OTU_84	FALSE	0	TRUE
	OTU_92	FALSE	0	TRUE
	OTU_97	TRUE	1	TRUE
	OTU_107	FALSE	0	TRUE
	OTU_130	TRUE	1	TRUE
	OTU_169	TRUE	1	TRUE
	OTU_304	TRUE	0	TRUE

```
## MSE.filter      BP.test
## Mode :logical   Mode :logical
## FALSE:6         FALSE:1
## TRUE :45        TRUE :50
```

plot of scaled kept OTUs



plot of scaled rejected OTUs



4.2 Metabolites

```
## [1] "M106T894" "M179T1018" "M205T1473" "M207T1196" "M229T1227"
## [6] "M271T1466" "M285T1569" "M290T1524" "M291T1584" "M292T1383"
## [11] "M308T1437" "M310T1500" "M357T2099" "M379T1799" "M398T1643"
## [16] "M415T2220"
```

##	molecule	modelsUsed	BP.test
## 1	M266T1372	0	FALSE
## 2	M271T1466	2	TRUE
## 3	M179T1018	2	TRUE
## 4	M129T1196	0	FALSE
## 5	M207T1196	3	TRUE
## 6	M106T894	0	TRUE
## 7	M308T1437	0	TRUE
## 8	M310T1500	2	TRUE
## 9	M290T1524	3	TRUE
## 10	M285T1569	2	TRUE
## 11	M379T1799	3	TRUE
## 12	M369T1850	0	FALSE
## 13	M357T2099	0	TRUE
## 14	M415T2220	0	TRUE
## 15	M229T1227	3	TRUE
## 16	M205T1473	3	TRUE
## 17	M292T1383	3	TRUE

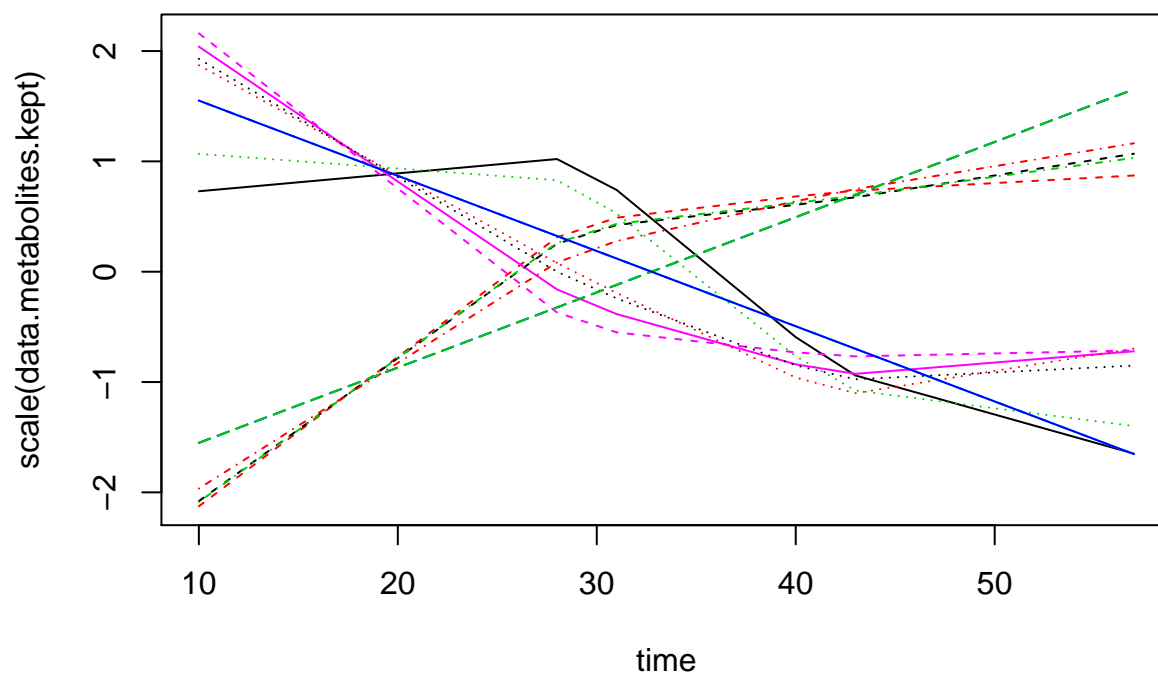
```

## 18 M299T1033      0  FALSE
## 19 M291T1584      0   TRUE
## 20 M398T1643      0   TRUE

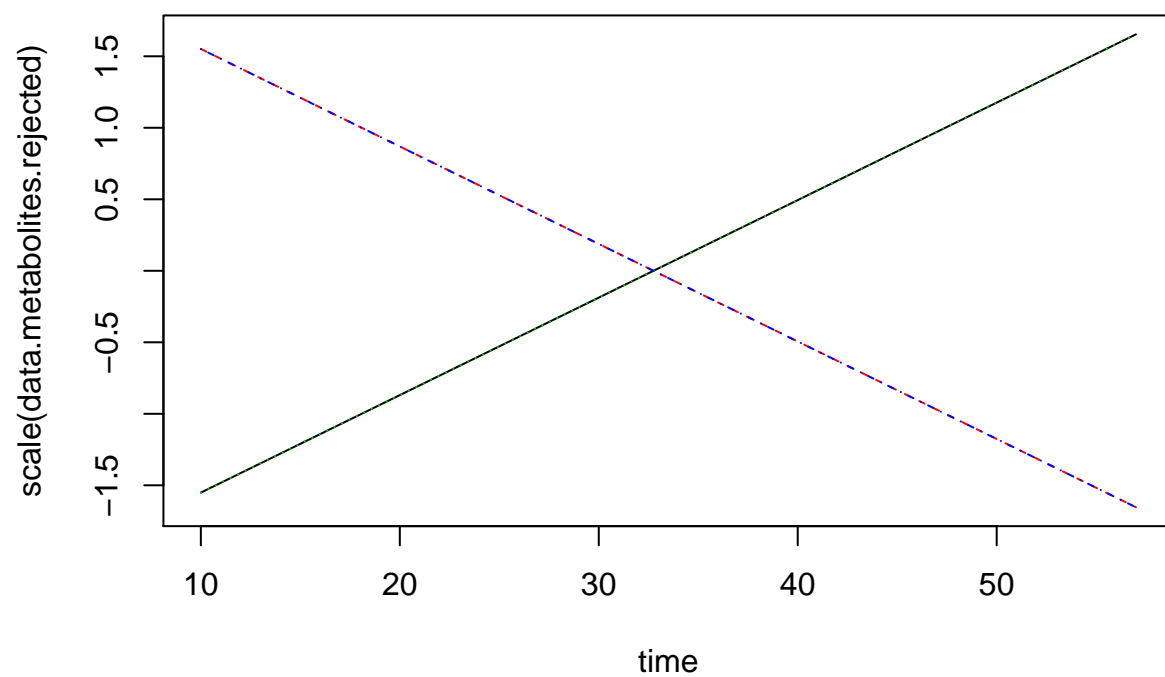
##  molecule      modelsUsed  BP.test
## Length:20      Min.   :0.0  Mode :logical
## Class :character 1st Qu.:0.0  FALSE:4
## Mode  :character Median  :1.0  TRUE :16
##                  Mean    :1.3
##                  3rd Qu.:3.0
##                  Max.    :3.0

```

plot of scaled kept metabolites

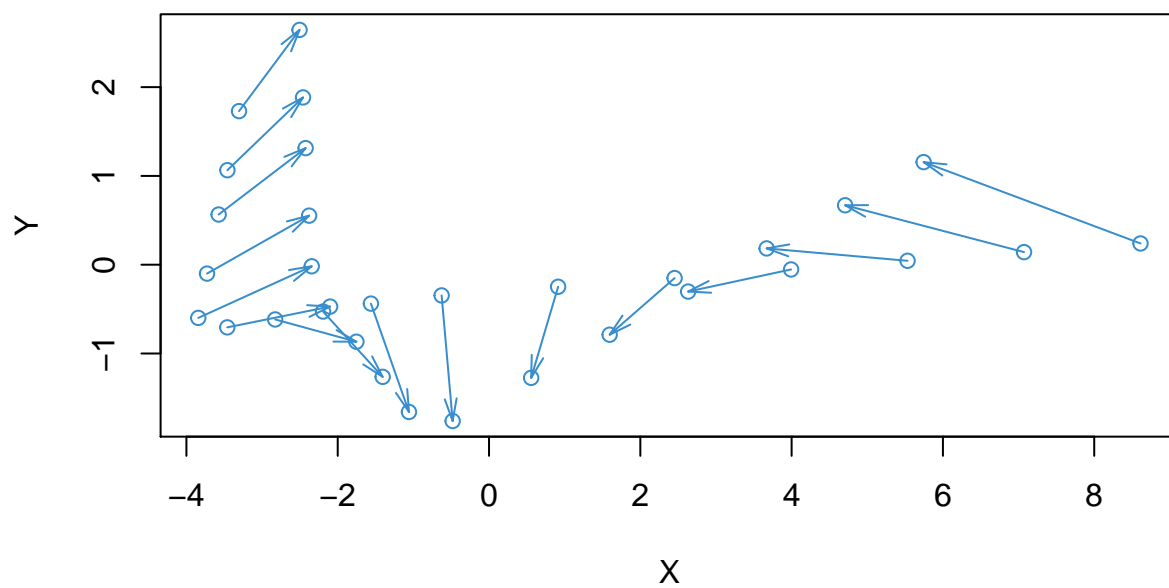
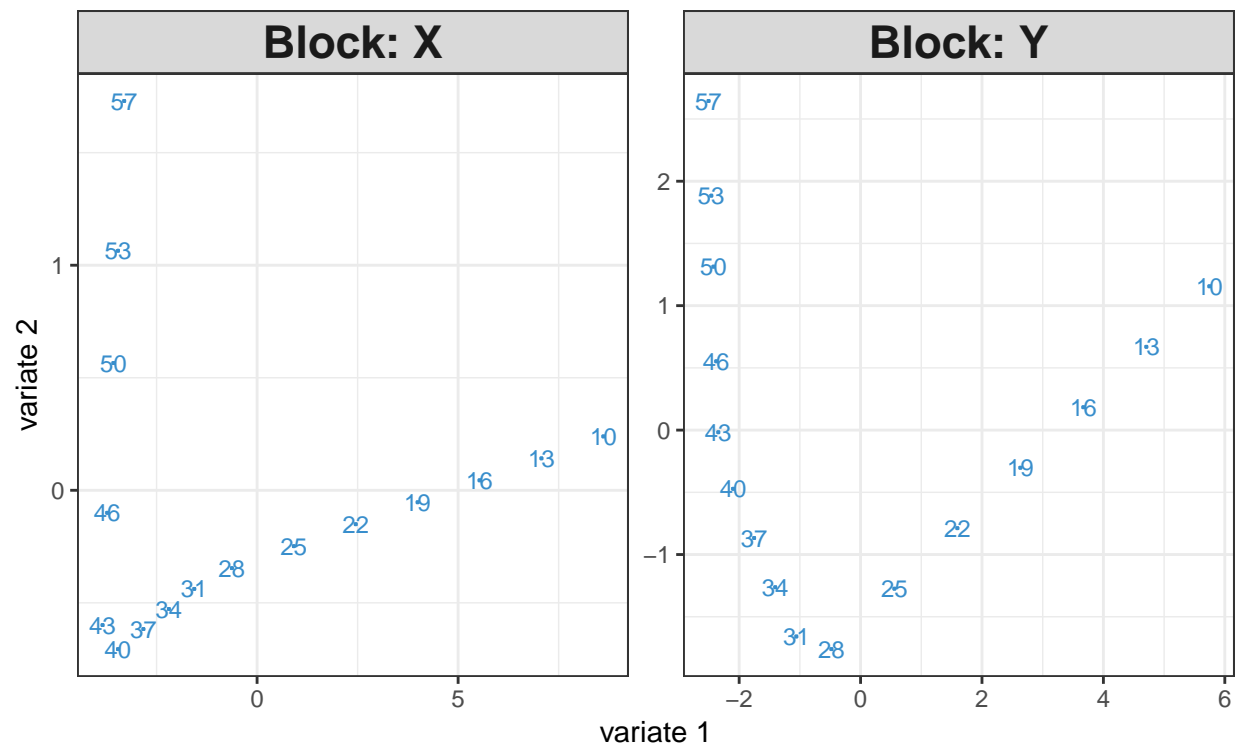


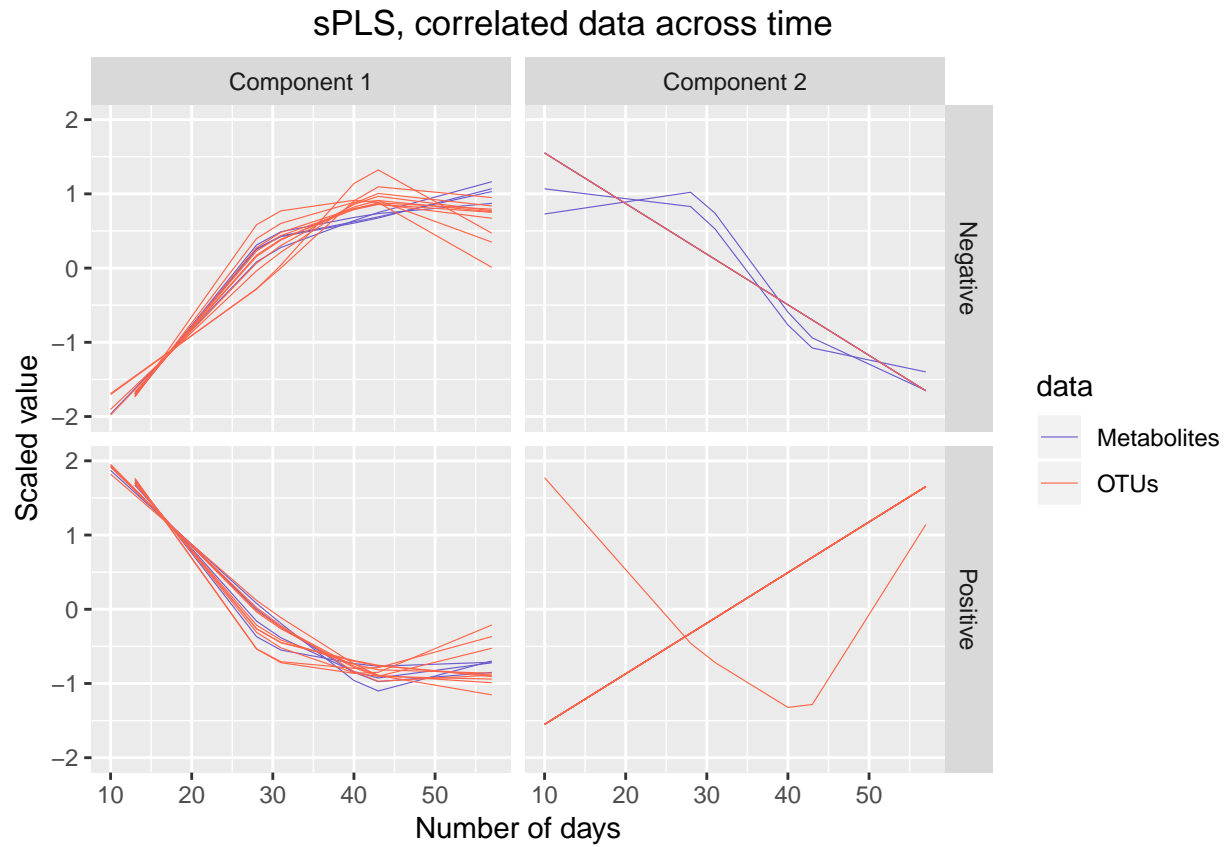
plot of scaled rejected metabolites



5 sPLS

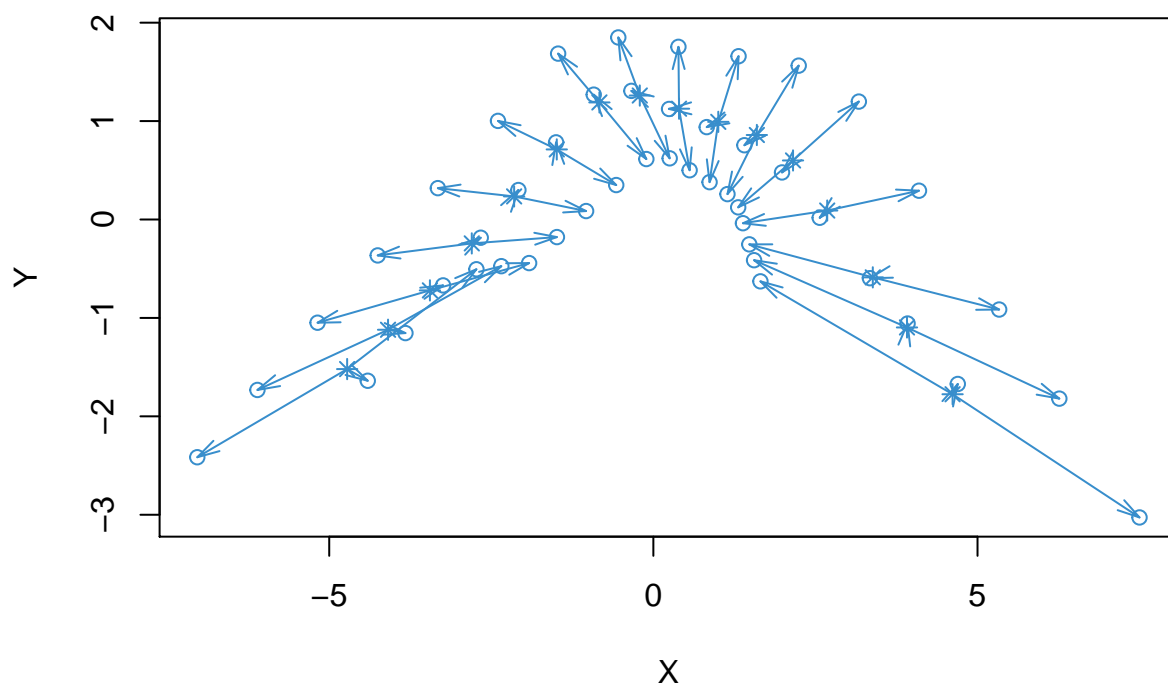
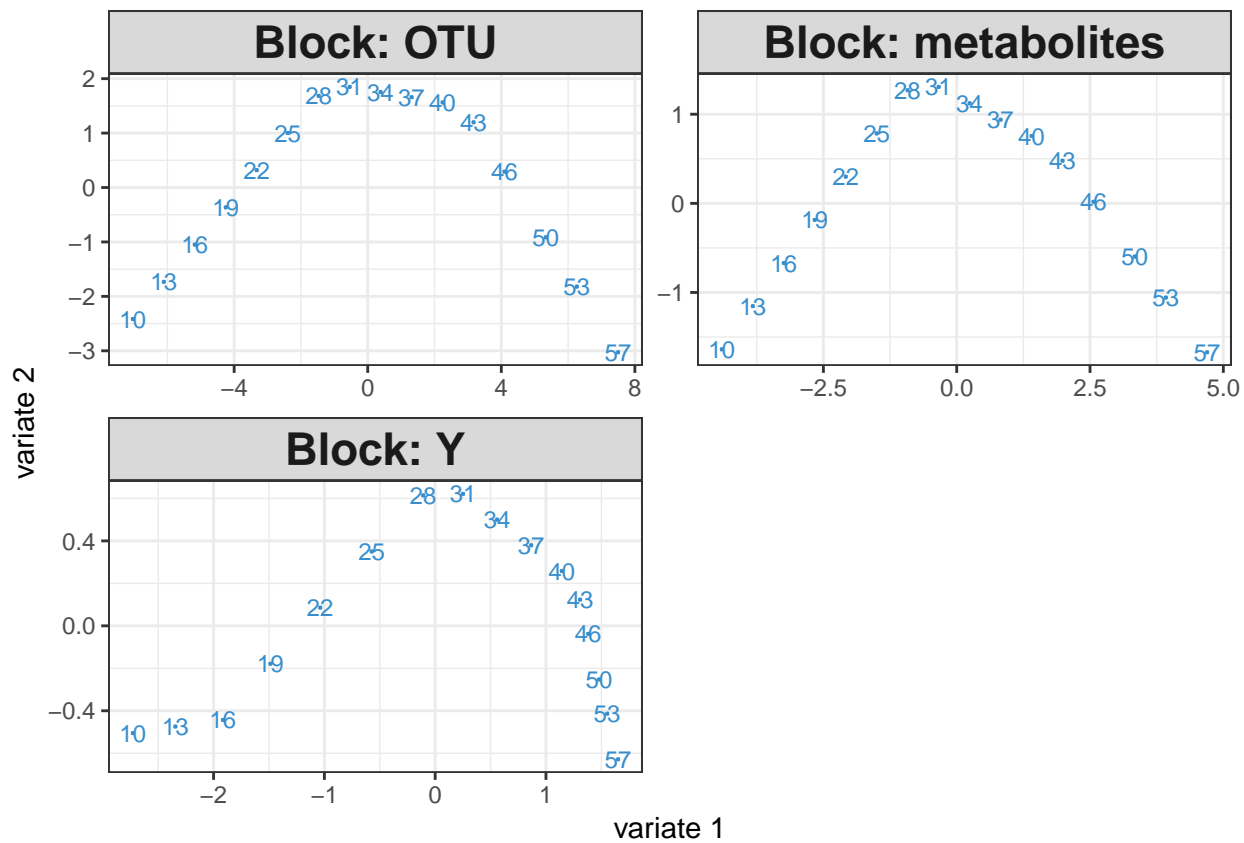
16S data and metabolites are analysed.



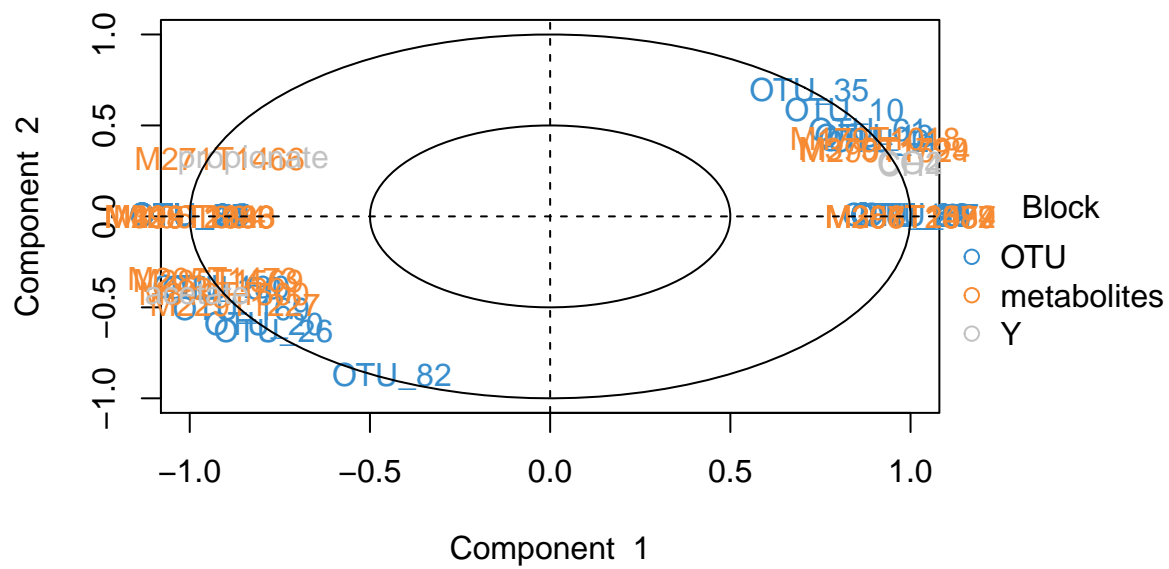


6 block sPLS

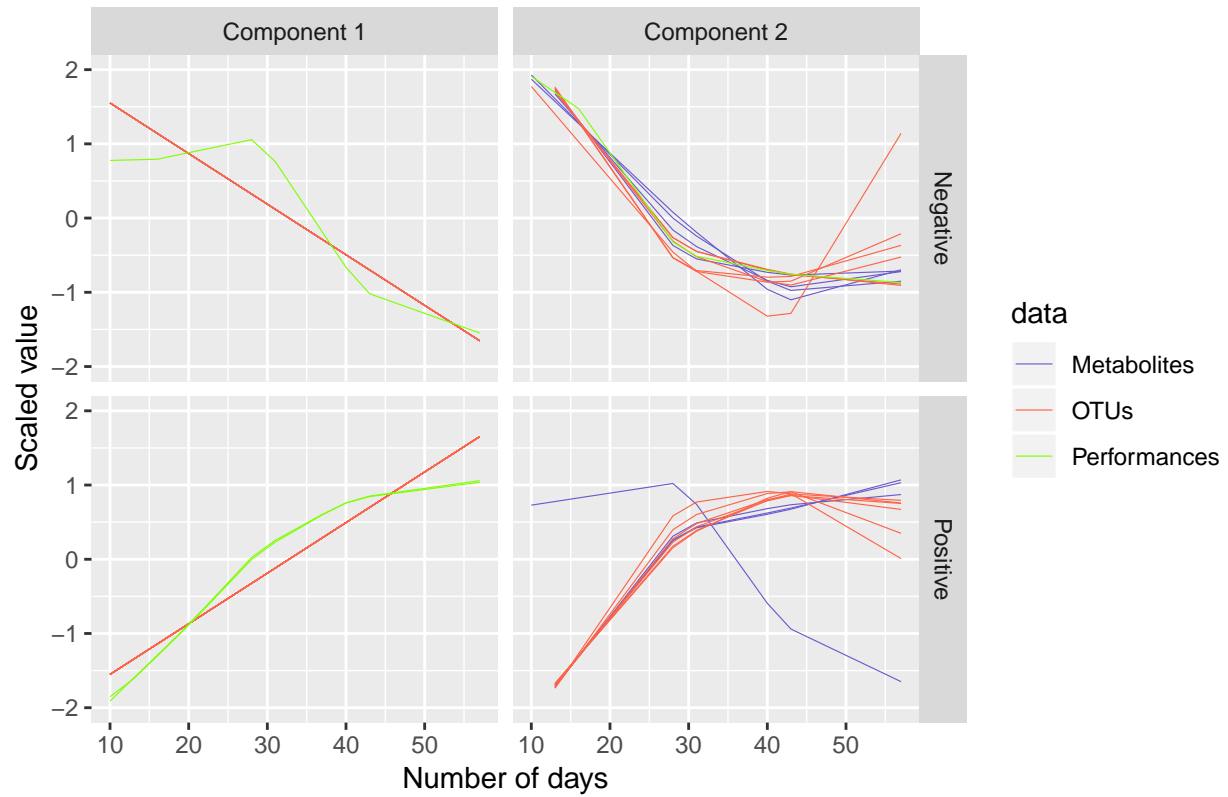
The three datasets (OTUs, metabolites and performances) are analysed together.



Correlation Circle Plots



block sPLS, correlated data across time



```
## variables selected with block spls
selected.variables=levels(melt.rgccca$Var1)
non.selected.OTUs=colnames(data.OTUS)[!colnames(data.OTUS)%in%selected.variables]
non.selected.metabolites=colnames(data.metabolites)[!colnames(data.metabolites)%in%selected.variables]
```