

Trait-based products of EMODnet benthic biology

Karline Soetaert, Olivier Beauchard, Peter Herman

9 October 2018

Abstract

Aim of this analysis.

Throughout their activity, benthic animals play an important role in marine ecosystem functioning. More specifically, they mix the sediment by their movement and feeding activity, a process called “bioturbation”. They also create water movements, enhancing the exchange of dissolved constituents such as oxygen, and dissolved inorganic nutrients, a process called “bio-irrigation”.

Both these activities have a large impact on the biogeochemical cycles in the environment, and they are commonly parameterised as single parameters in biogeochemical models.

Biologists have tried to categorise the bioturbation or bio-irrigation activity based on the identity of the organisms. They do not derive the rate parameters as used in the biogeochemical models but rather derive a potential of the organisms to perform these tasks.

Here we use Benthic abundance data from EMODnet to estimate the bioturbation potential and the bioirrigation potential index.

To derive these potentials we need information on:

- The weights of the species and their total biomass
- The species mobility and sediment reworking mode

Reading the data

Species taxonomic tree

The names of all species encountered in the dataset were checked against the WORMS database, and their taxonomic tree added (this was done via <http://www.marinespecies.org/>, menu item *tools/Match taxa*). These data are read first. The dimensions of the data set and the first two entries are printed.

```
##  
## dimension and first part of the data set :  
## [1] 7650     8  
  
##      id      phy      clas          ord          fam  
## 1 7500 Hemichordata Enteropneusta [unassigned] Enteropneusta  
## 2     1   Cnidaria    Hydrozoa          Leptothecata Sertulariidae  
##           gen            tx            txa  
## 1 [unassigned] [unassigned] Enteropneusta [unassigned] Enteropneusta  
## 2 Abietinaria       Abietinaria       Abietinaria
```

the MWTL data

The MWTL data is the only data set that contains both species *biomass* and *densities* (the other data sets comprise only densities or just presence/absence). This data is used to estimate mean individual weights of the various species.

Biomass is in *g AFDW/m²*, density in *number /m²*

```
##  
## dimension and first part of the data set :  
## [1] 10369      6  
##   monsterpunkt.id      tx      biomass      density      lon      lat  
## 1       BREEVTN02     Abra 8.333333e-05  0.8111111 37.05556 95.05556  
## 2       BREEVTN02 Abra alba 7.369833e-02 58.7722222 37.05556 95.05556
```

Species body mass

There are two biomass data sets. The first data set was prepared by Olivier Beauchard; it contains the body mass in ash-free dry weight.

```
##  
## dimension and first part of the data set :  
## [1] 385      2  
##                  tx BodyMassAFWD  
## 1 Abludomelita obtusata    0.000314  
## 2             Abra alba    0.010100
```

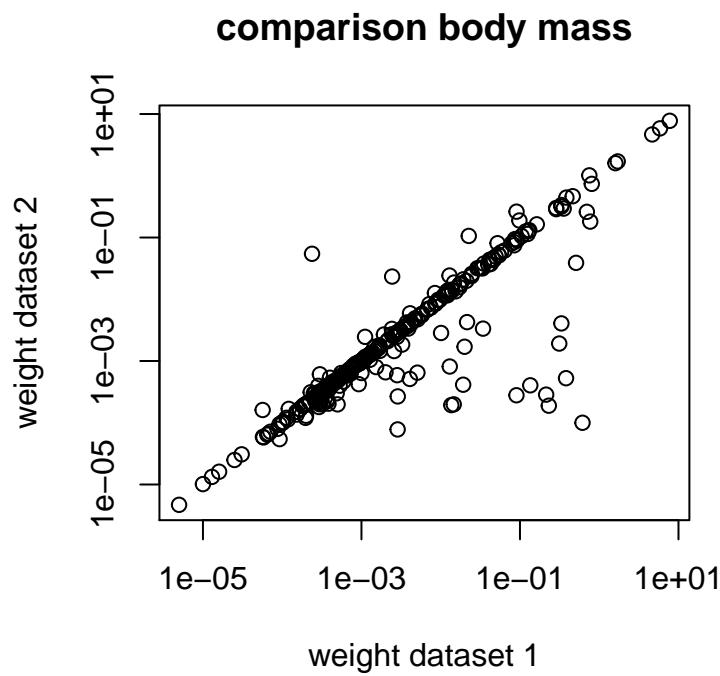
Another dataset was derived from the MWTL data series by Peter Herman; it contains both the mean weight (*mw*) and geometric mean weight (*gmw*), in ash-free dry weight, and the number of instances on which this was determined (*n*).

```
##  
## dimension and first part of the data set :  
## [1] 605      5  
##      X          species          mw          gmw  n  
## 1 1 Abludomelita obtusata 0.0003007284 0.0003007279 12  
## 2 2                 Abra 0.0003744496 0.0001958083  7
```

Checking consistency in weight data

Both data sets contain species weights that are not present in the other data set, so they are merged. First it is checked if the data are compatible.

```
## number of species in first dataset absent in second: 37  
## number of species in second dataset absent in first: 257  
## number of species in common: 348
```



Combining the two weight datasets.

We combine both data sets and add the taxonomic tree by merging the data with the taxon data.

Expanding the weight dataset

Next we estimate mean weights on species level, on genus level and on the family level, and bind all in one data.frame(mws).

The total number of weight values thus obtained is 1136, of which 642 are estimated at species level, 318 at genus level and 176 at family level.

FeedingTypes

Feedingtypes are known from a subset of the species. The following types are distinguished:

- CaSc = carnivore/scavenger
- De = depositfeeder
- He = herbivore
- Om = omnivore
- Pa = parasite
- Su = suspension feeder
- SuDe = suspension/deposit feeder

```
##  
## dimension and first part of the data set :
```

```

##          n.acc Trophy
## 1 Abludomelita obtusata     De
## 2             Abra alba   SuDe

##
## CaSc   De   He   Om   Pa   Su SuDe
## 78    100   20    56    3    79   37

```

Feeding types on higher taxonomic levels

We now assign feeding types to the genera and families, for which we take the most commonly encountered feedingtype at the lower level.

```

##          taxon   ft   level n
## 1 Abludomelita obtusata   De species 1
## 2             Abra alba SuDe species 1

```

The total number of feeding types thus obtained is 810, of which 373 are known at species level, 270 at genus level and 167 at family level.

life history traits

Life history traits are also assigned on genus and family level, where we take the most commonly encountered trait at the lower level.

```

##
## dimension and first part of the data set :
## [1] 273 21

##      phy      cla      ord      fam      gen
## 1 Annelida Polychaeta Sabellida   Serpulidae Spirobranchus
## 2 Annelida Polychaeta Sabellida Sabellariidae   Sabellaria
##          taxon Motility Body.size Burrowing.depth Morphology
## 1 Spirobranchus triqueter Tubicolous    1-3cm        0cm Cylindrical
## 2   Sabellaria spinulosa Tubicolous    3-10cm        0cm Cylindrical
##      Mobility Mixing.type Mixing.rate Mi Mi2 Ri Morphology.2
## 1 Very low Surface mixing   Very low  1   1  2           1
## 2 Very low Surface mixing   Very low  1   1  1           1
##      Morphology.2.1 Body.size.2 Burrowing.depth.2 Mixing.rate.2
## 1           1           2           1           2
## 2           1           3           1           3

```

The total number of traits thus obtained is 493, of which 273 are estimated at species level, 161 at genus level and 93 at family level.

We tabulate the number of occurrences of each trait in the resulting data.frame.

```

## motility:
##
##      Attached      Crawler Crawler-Swimmer      Tubicolous
##                 13         268          128            84
##      Body size:
##
##      <1cm   >20cm   1-3cm 10-20cm 3-10cm

```

```

##      98      34      97      54     210
## Burrowing depth:
##
##      >30cm   0-5cm    0cm 10-15cm 15-30cm 5-10cm
##      27      204      70      59      79      54
## Morphology:
##
##      Articulated   Bivalved Cylindrical      Flat   Globular   Stellar
##      112            99          226           1       31        24
## Mobility:
##
##      High Intermediate      Low   Very low
##      37            230          145          81
## Mixing type:
##
##      D/U conveying      Diffusion Downward conveying
##      33                  118          20
##      Regeneration      Surface mixing Upward conveying
##      9                  282          31
## Mixing rate:
##
##      High Intermediate      Low   Very high   Very low
##      32            57          133          41        230
## Mi:
##
##      1    2    3    4
##      81  145  230  37
## Ri:
##
##      1    2    3    4    5
##      25  257  82  120   9

```

The density data

Density data are read.

```

##
## dimension and first part of the data set :
## [1] 1128549      6
##      data      sta      x      y      tx      dens
## 1 HELCOM HELCOM1 8.267167 56.72267  Abra alba 139.8601
## 2 HELCOM HELCOM1 8.267167 56.72267  Abra nitida 139.8601
##
## total number of data points per provider :

```

```

##  

##    HELCOM MACROBEL MAREANO      MWTL      NSBS      ODAM      PMP      PORT  

##    10399     12701    21591      8475     14622    94544      333      6796  

##    REBENT     RSMP     SHARK      SMHI  

##    24516     875244   56268      3060

```

Some data providers did not record density, but just presence/absence; these data are removed.

The total number of species in this data set is 5433

A look at the data

Species densities are summed per station to give the total densities

```

## total number of stations per provider :  

##  

##    HELCOM MACROBEL MAREANO      MWTL      NSBS      ODAM      PMP      REBENT  

##    914       768     370       103      235     3057      23       216  

##    RSMP     SHARK      SMHI  

##    23915     6925     587

```

The total number of stations in this data set is 37113

The positions of the stations are plotted, colored according to the data provider

```

## Loading required package: rworldmap  

## Warning: package 'rworldmap' was built under R version 3.5.1  

## Loading required package: sp  

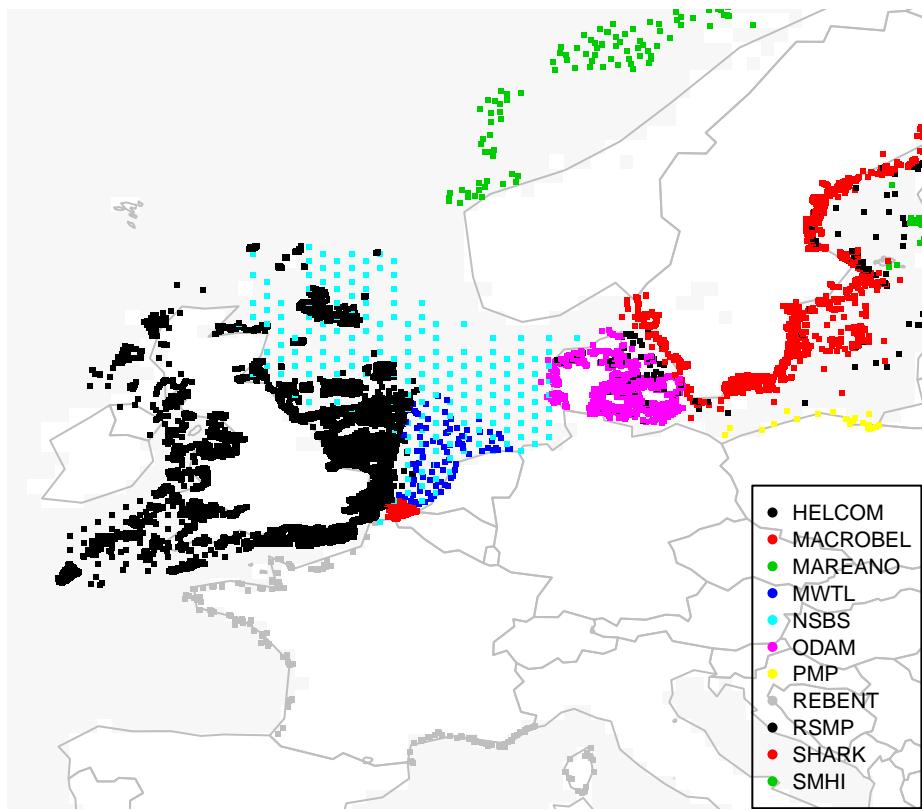
## Warning: package 'sp' was built under R version 3.5.1  

## #### Welcome to rworldmap ###  

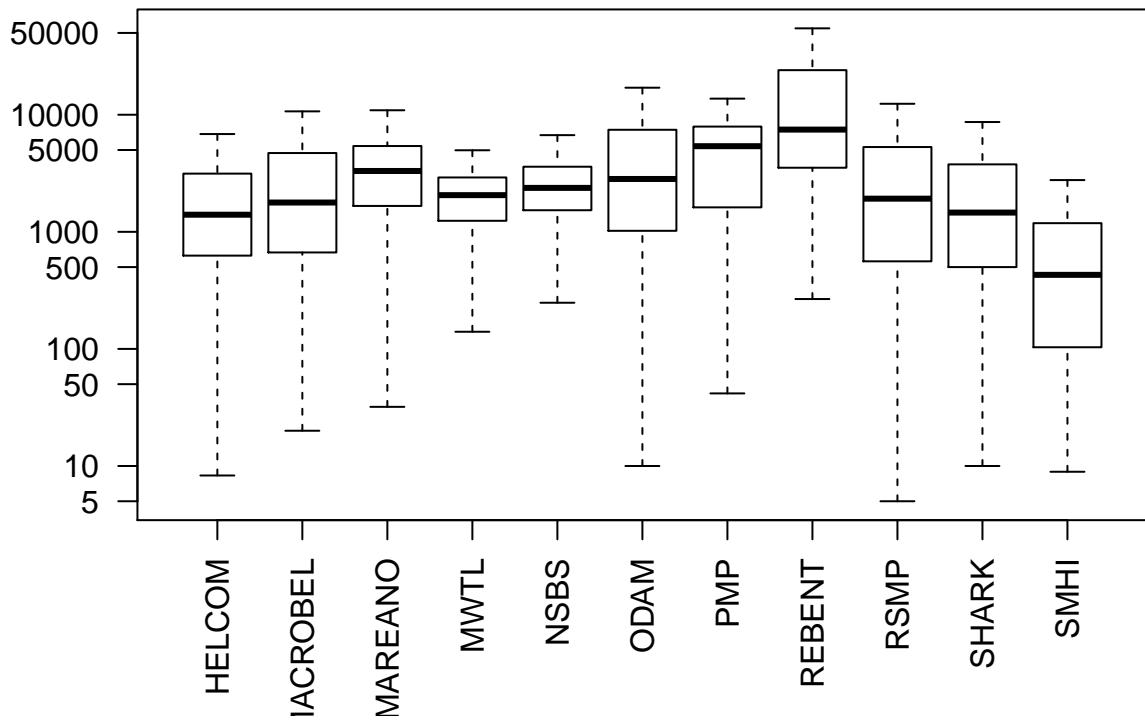
## For a short introduction type : vignette('rworldmap')

```

data providers



Mean Density



Adding individual weights and traits to density data

Weights and trait information is not available for all species, even after determining this information on genus and family level. We create data sets that remove the species that do not have all information. We then calculate total density on this reduced data set and see what fraction of total density we obtain.

First the required taxonomic information is added to the density list:

```
## integer(0)
```

To merge density data with the other data sets, a function is created

Merging density and weights

Merging density and traits

Merging density and feeding types

Data that have all necessary information for estimating bioturbation potential

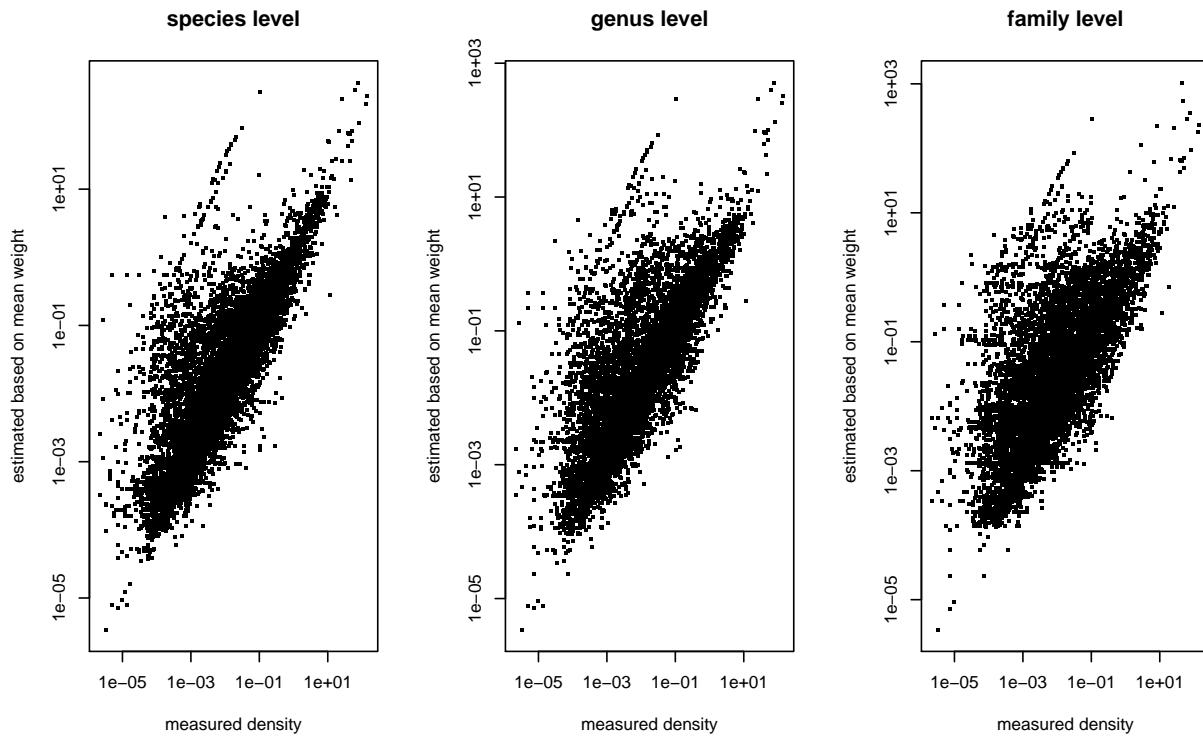
```
## [1] 758588      23  
## [1] 1103726     14
```

The fraction of data that has all information to estimate bioturbation potential = 0.6872974

Check on the representativeness of the reduced data set

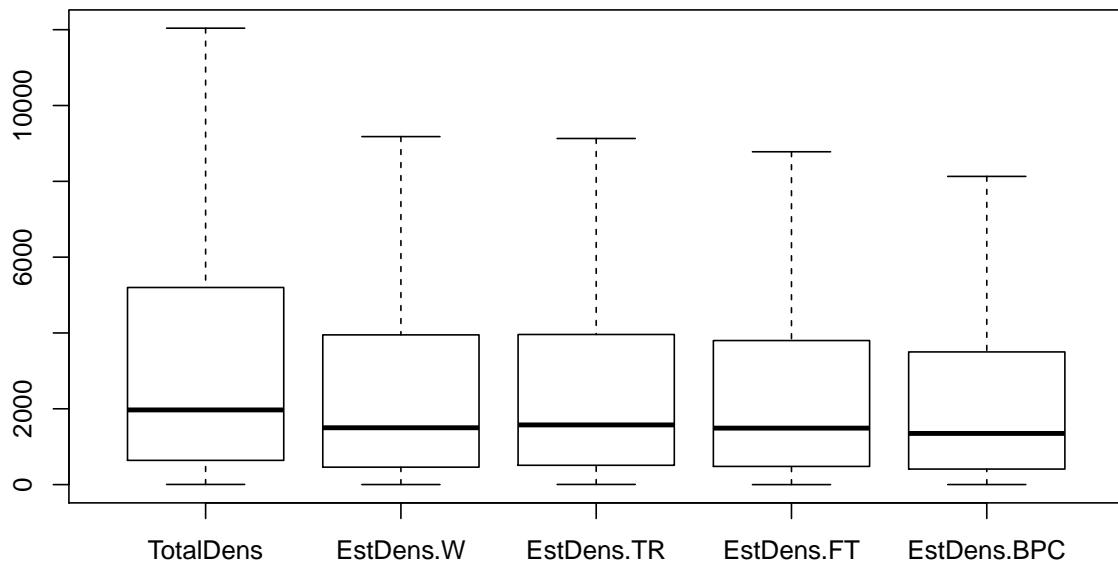
For the mwlt data we now compare the measured biomass with the biomass that we estimate, based on mean weights.

```
##      Min. 1st Qu. Median   Mean 3rd Qu. Max.
## -33.90963 -0.00006 0.00062 0.23759 0.01563 288.33301
##      Min. 1st Qu. Median   Mean 3rd Qu. Max.
## -25.4000 -0.0002 0.0008 0.3563 0.0179 434.2944
```



```
##      Min. 1st Qu. Median   Mean 3rd Qu. Max.
## -18.1332 -0.0001 0.0025 0.5481 0.0427 995.4458
```

We also check the error we would make in the density data if we would estimate them on the merged data sets.



```

## fraction of density based on species for which weight is known:
##      Min. 1st Qu. Median     Mean 3rd Qu.      Max.
## 0.0003466 0.7106343 0.8709677 0.8025307 0.9811321 1.0000000

## fraction of density based on species for which traits are known:
##      Min. 1st Qu. Median     Mean 3rd Qu.      Max.
## 0.001121 0.699346 0.845454 0.807406 0.962963 1.0000000

## fraction of density based on species for which feeding type is known:
##      Min. 1st Qu. Median     Mean 3rd Qu.      Max.
## 0.000999 0.654321 0.800000 0.774501 0.945946 1.0000000

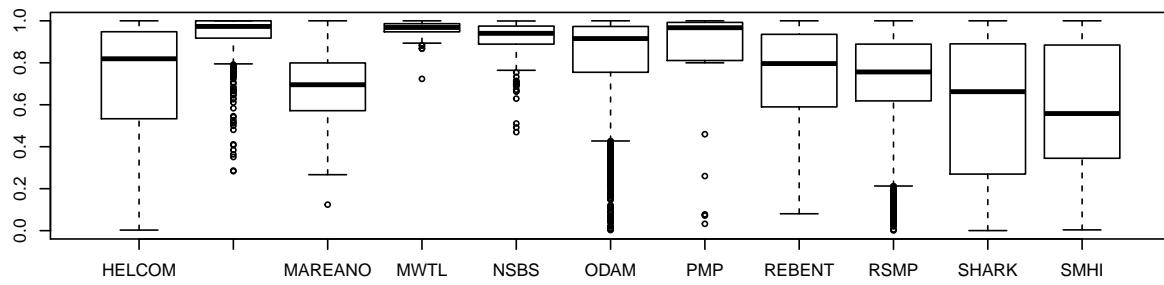
## fraction of density based on species for which all is known:
##      Min. 1st Qu. Median     Mean 3rd Qu.      Max.
## 0.0003466 0.5959077 0.7692308 0.7257031 0.9166667 1.0000000

```

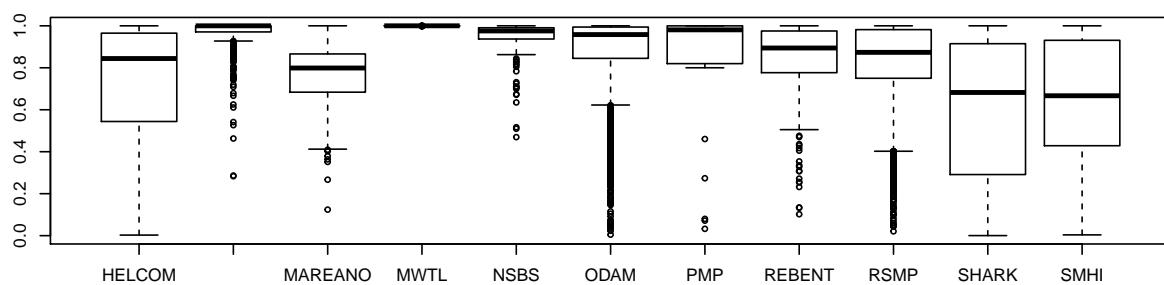
The estimated density if we only use species for which have all data required for BPc estimation is thus on average 0.7257031 of the true value.

As we want to estimate bioturbation potential, we now split this number for the different data providers

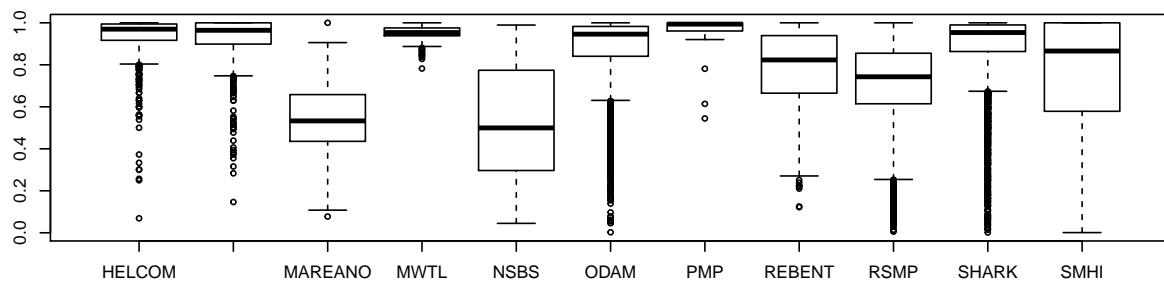
estimated fraction of density for BPC calculations



estimated fraction of density for Weights



estimated fraction of density for Feeding types



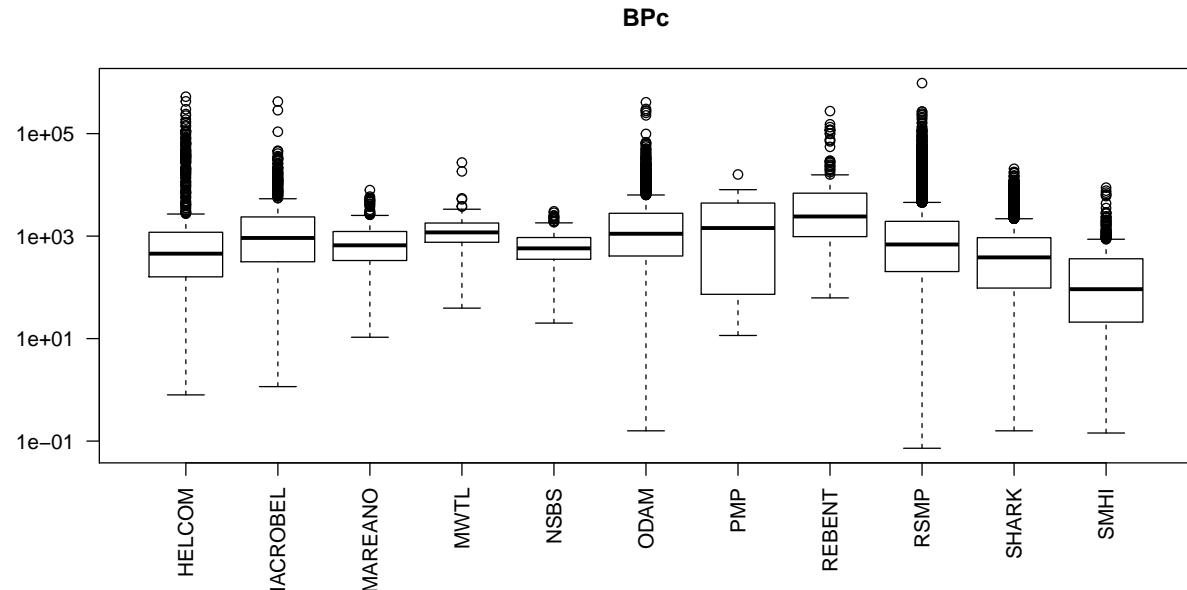
The recovered fraction of density:

```
##          BPC        W        FT
## HELCOM  0.6986901 0.7111337 0.9363019
## MACROBEL 0.9363593 0.9692866 0.9220701
## MAREANO  0.6832806 0.7699789 0.5438608
## MWTL    0.9597666 0.9999305 0.9458738
## NSBS    0.9121893 0.9454297 0.5267406
## ODAM    0.8314929 0.8823917 0.8770125
## PMP     0.7782387 0.7908448 0.9409039
## REBENT   0.7400029 0.8250691 0.7712088
## RSMP    0.7413248 0.8395028 0.7296203
## SHARK   0.5797466 0.5996586 0.8885059
## SMHI    0.5836764 0.6335552 0.7576398
```

Estimating BPc, the bioturbation potential

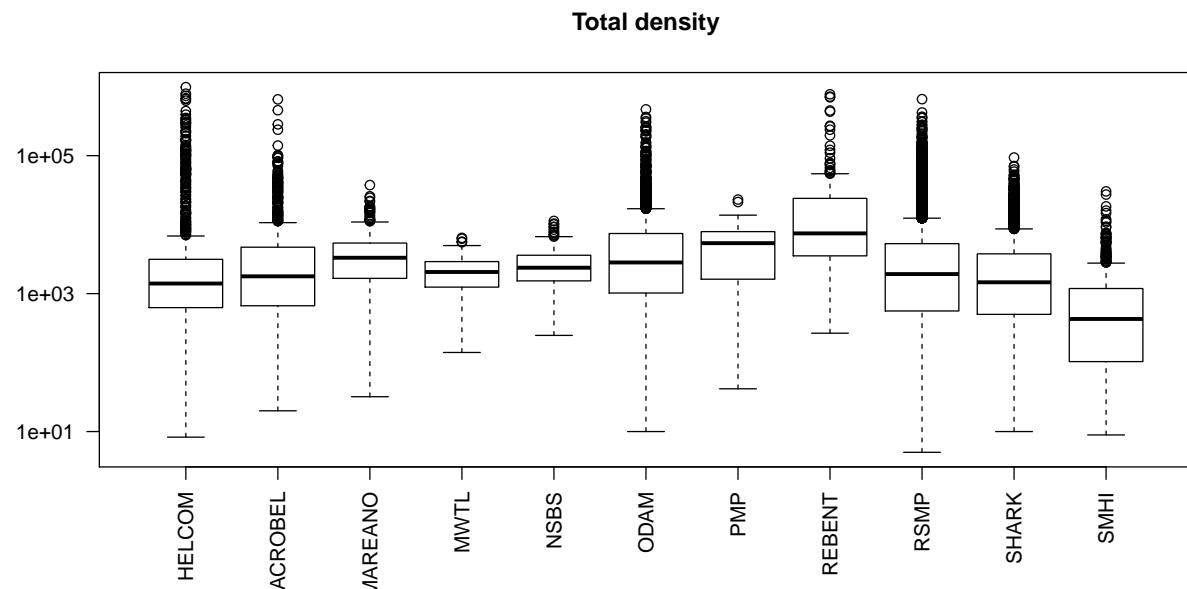
The bioturbation potential is now estimated on the reduced data set. First the contribution of each species to BPc is estimated.

Then the BPcs of all species per station are added.



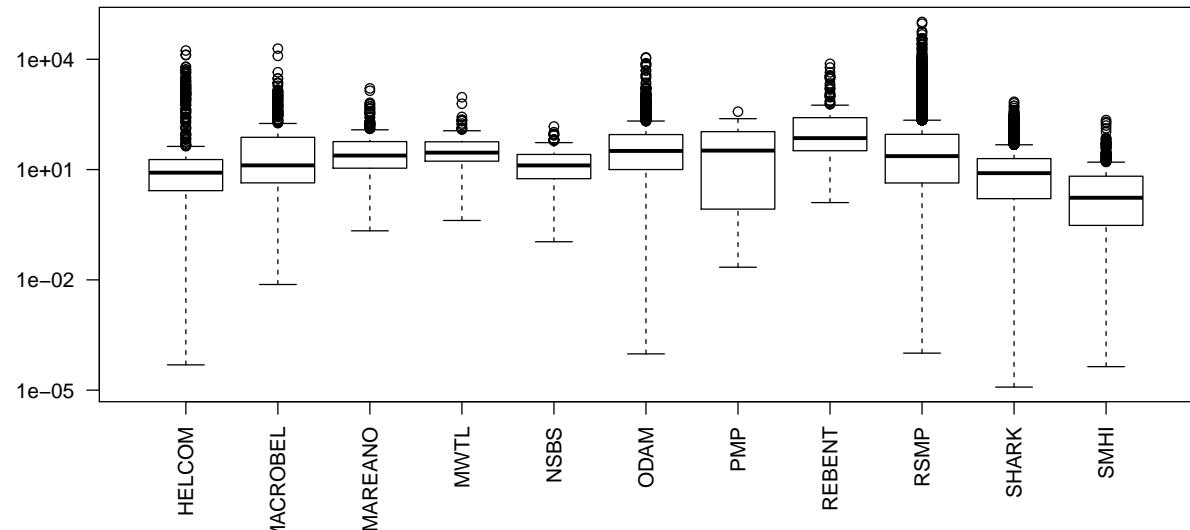
Biomass and total density

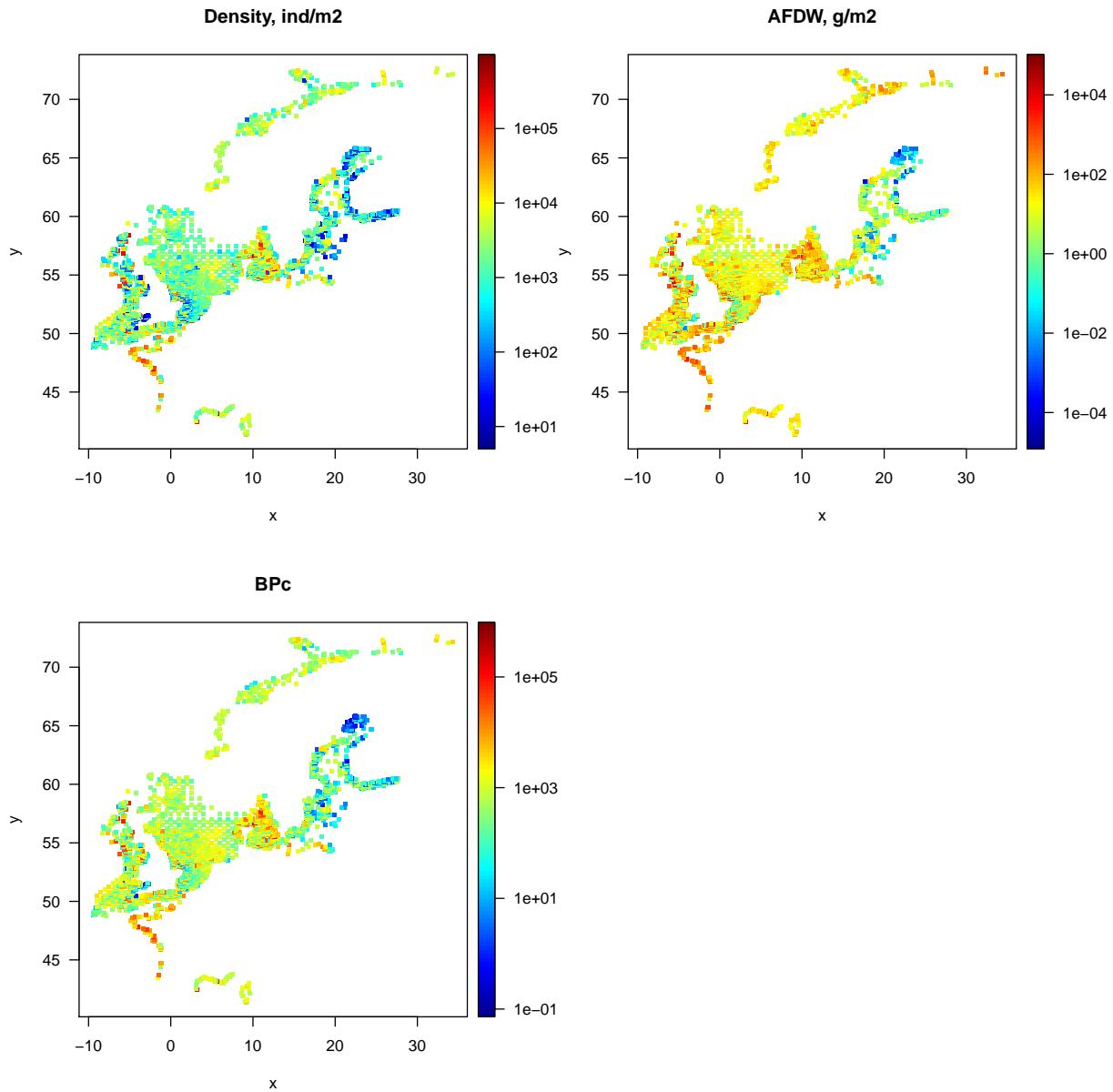
TotalDensity was already calculated



The total biomass per station:

Total biomass





Biomass per feeding type

```
##          sta   data      x      y     AFDW     CaSc     De He
## 1  HELCOM1 HELCOM 8.267167 56.72267 1332.03566 23.313394 60.084310 NA
## 2  HELCOM10 HELCOM 9.238333 56.65717 2478.42383  2.219319 566.765465 NA
## 3 HELCOM100 HELCOM 11.485833 56.20500  89.55318 23.483159 18.779218 NA
## 4 HELCOM101 HELCOM 11.500000 57.83333  15.27973 5.225638 1.817793 NA
## 5 HELCOM102 HELCOM 11.516667 57.75000  12.58953 3.677745 2.728795 NA
## 6 HELCOM103 HELCOM 11.525000 57.54998 133.81228 59.643812 23.212889 NA
##          Om Pa      Su     SuDe    AFDW_FT
## 1 819.86727366 NA 335.6890991 80.890008 1319.84408
## 2 124.07348421 NA 1781.3173901 4.048168 2478.42383
## 3      NA NA 0.9817005 25.491745  68.73582
```

```

## 4   0.06344955 NA    2.3267914  5.763861  15.19753
## 5   0.09517433 NA    2.0635217  3.914807  12.48004
## 6   2.04354271 NA   16.4510060 31.065997 132.41725

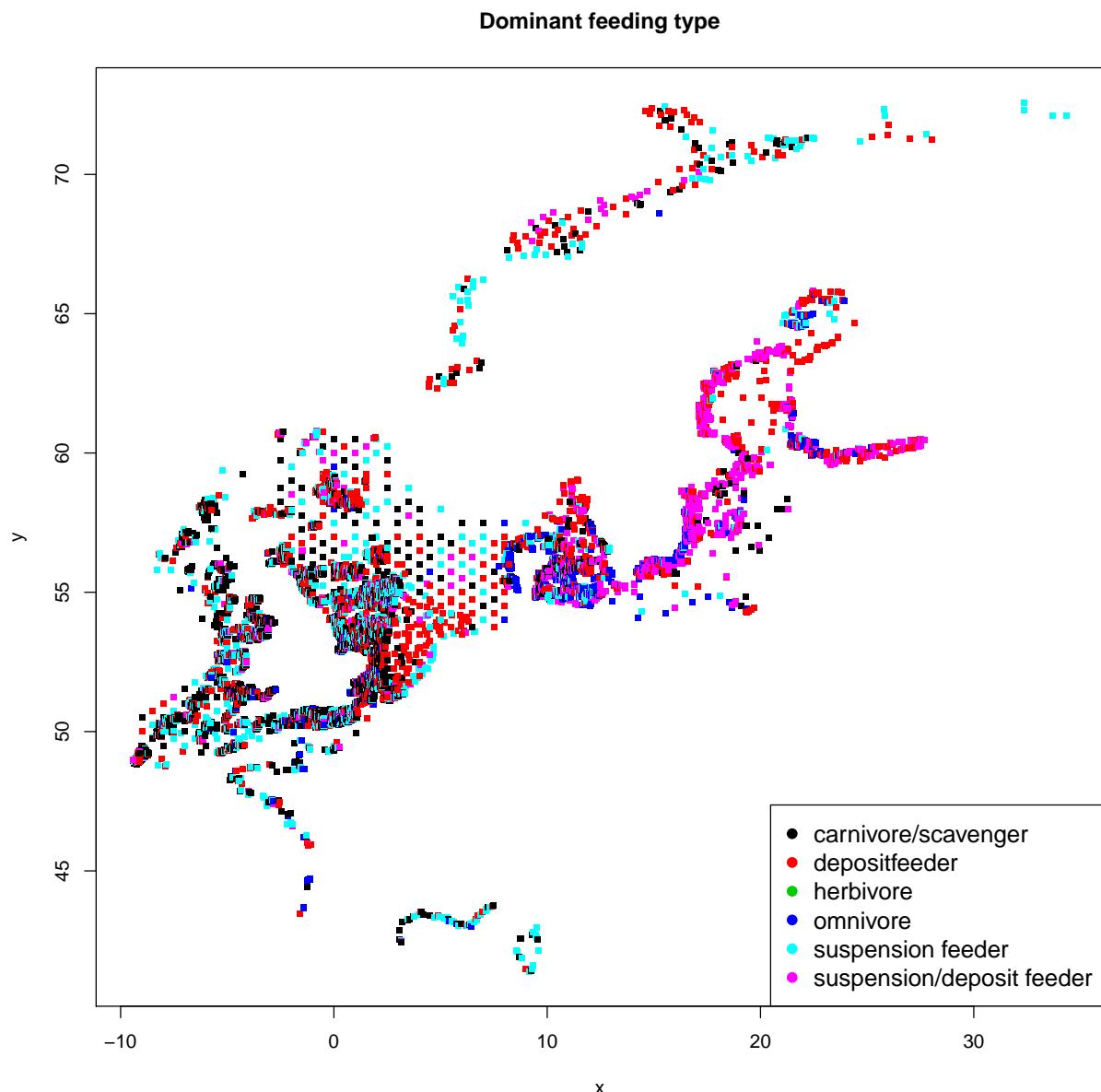
```

The stations where none of the feeding types are known are removed, and the most dominant feeding type selected.

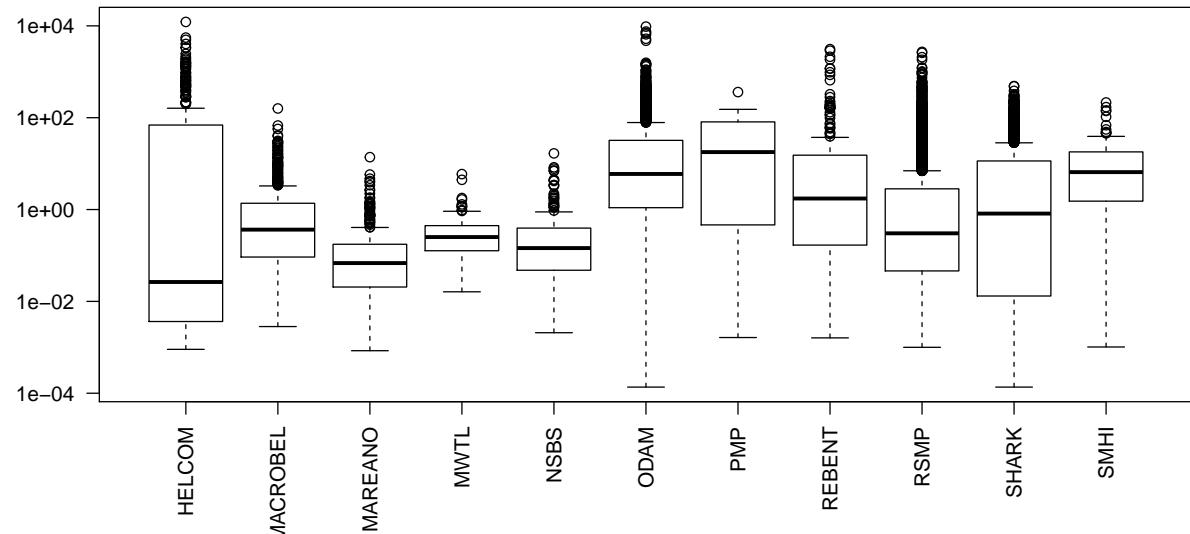
```

## ISNAS
##   1    2    3    4    5    6    7
## 89 17198 7692 4520 3758 2272 212

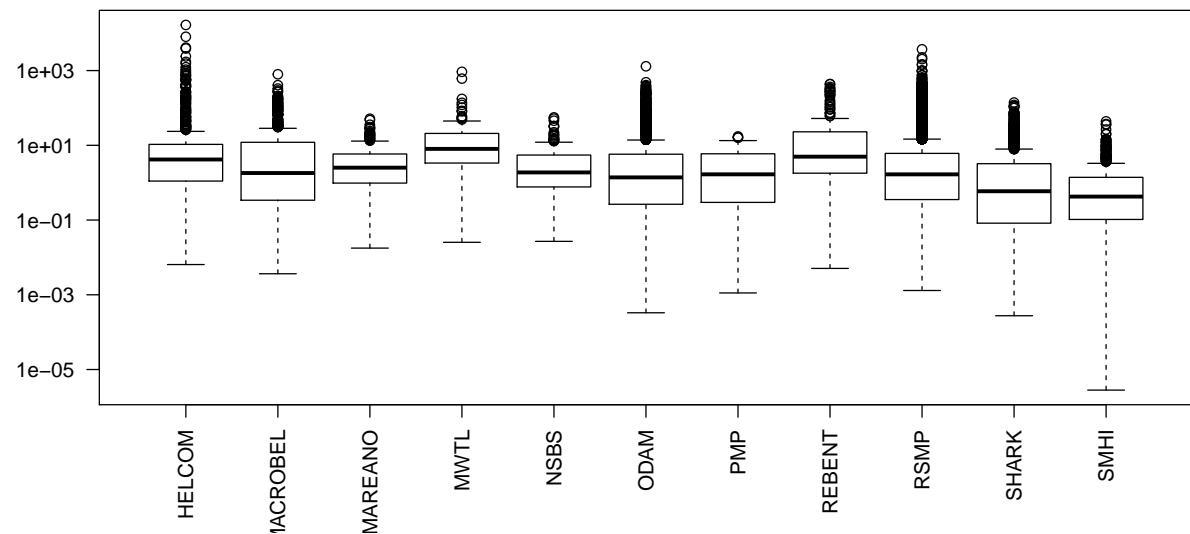
```



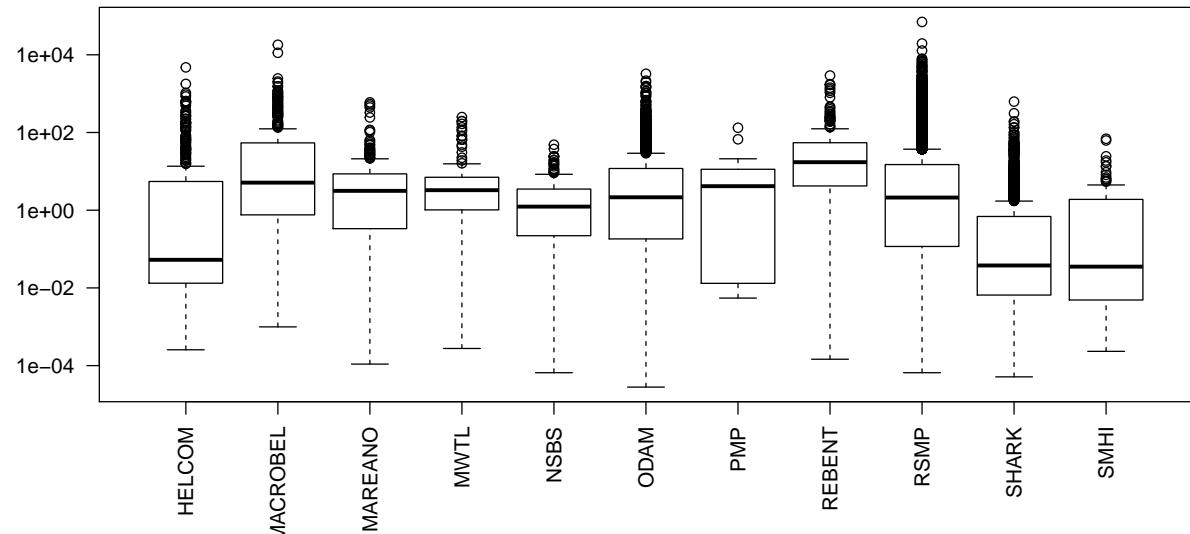
Total biomass Omnivores



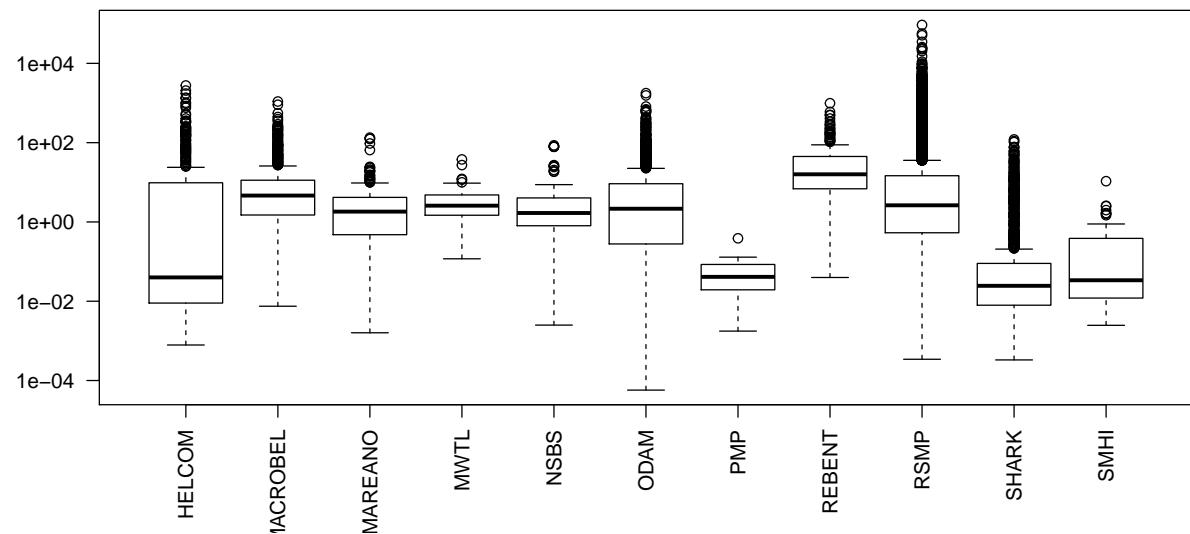
Total biomass Deposit feeders



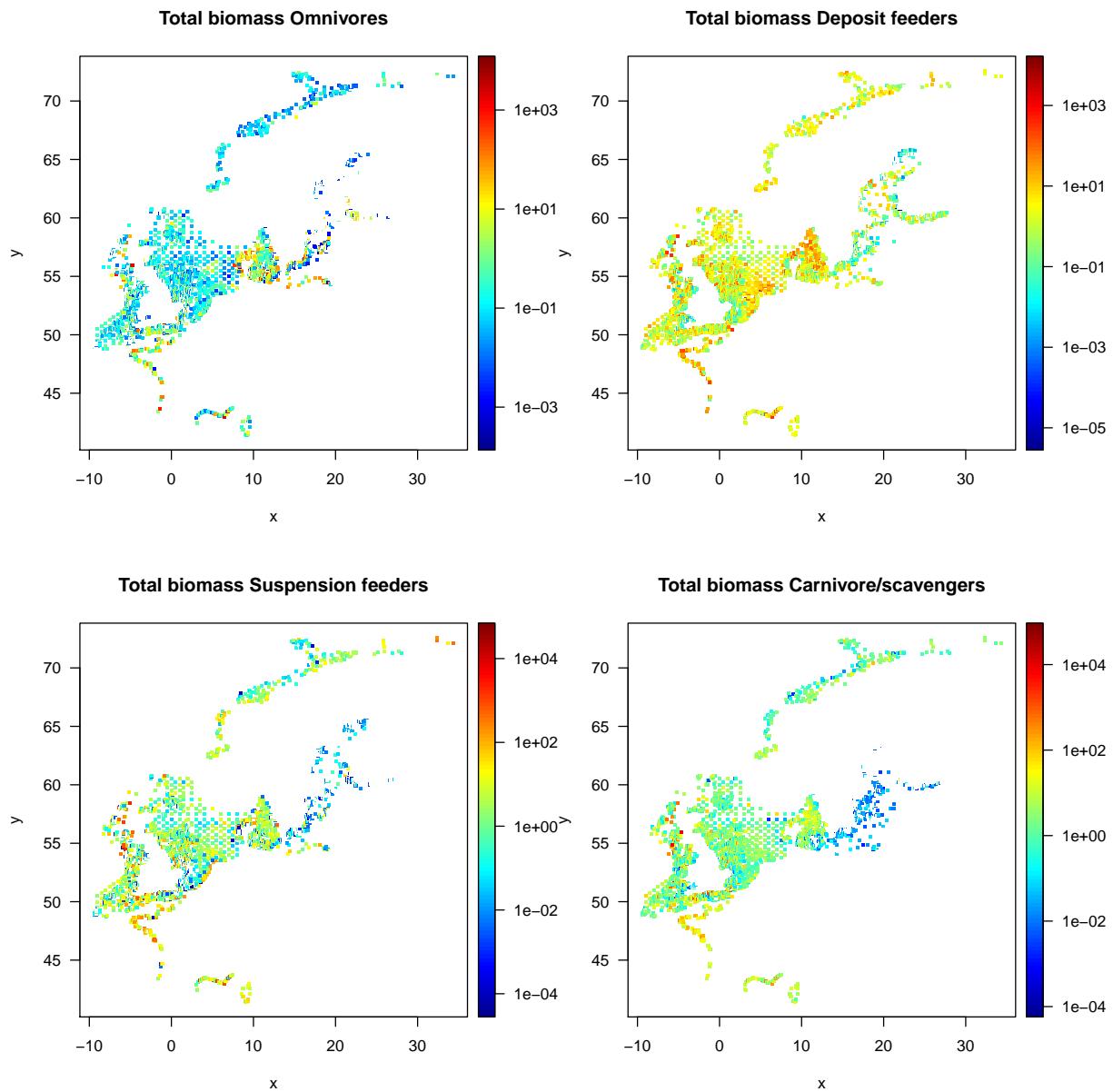
Total biomass Suspension feeders



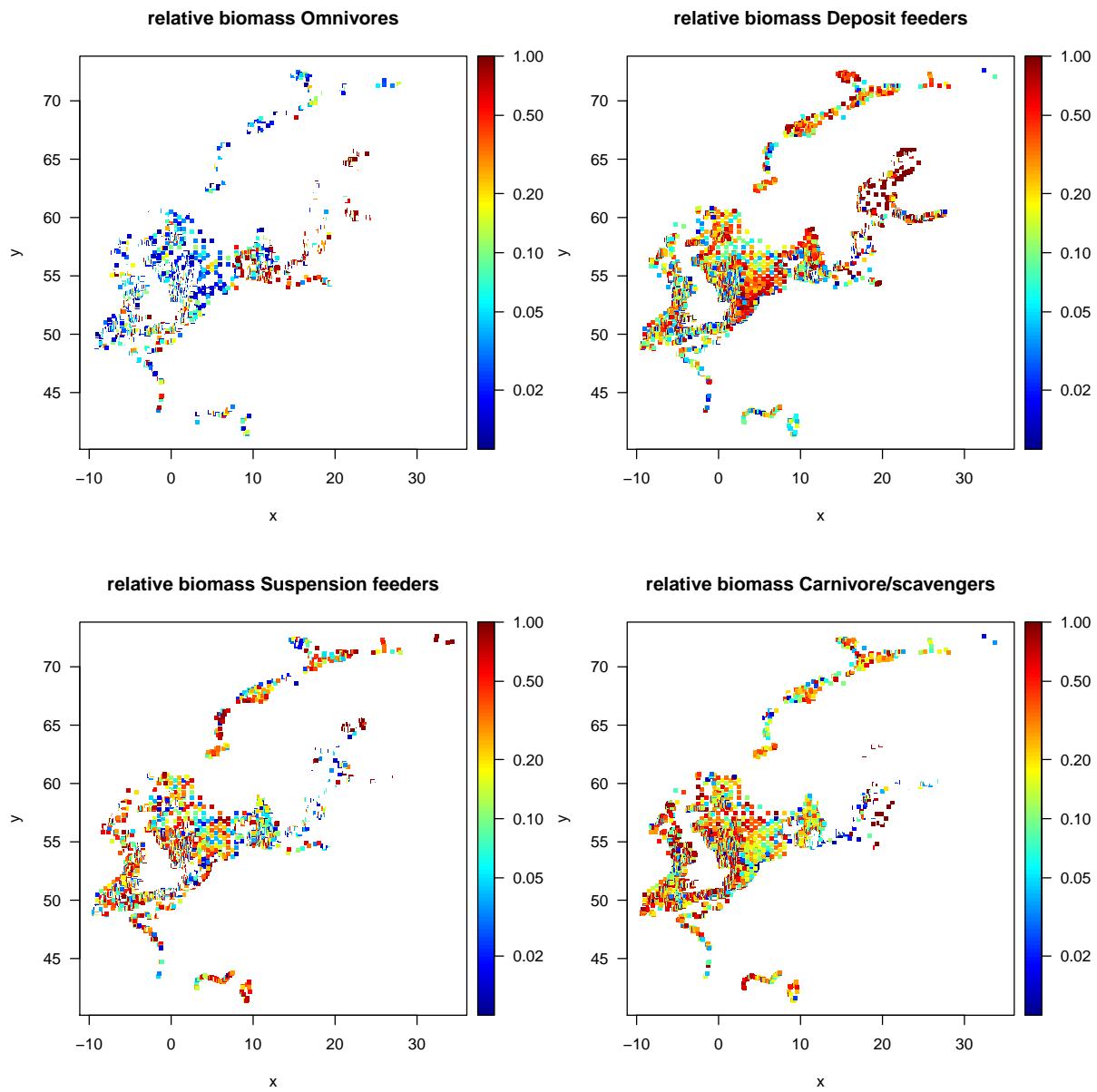
Total biomass Carnivore/scavengers



The total biomasses of feeding types



The relative biomasses of feeding types



Writing the results