

# Analysis of recent macrobenthic data from the Westerschelde - data and workflow

Peter M.J. Herman

2022-12-06

## Introduction

The monitoring of macrobenthos in the Westerschelde has been performed by Rijkswaterstaat since the early 1990s. Recently, the preliminary reporting on the data by Bureau Waardenburg, indicated a remarkable increase of total community biomass in the Westerschelde. In order to investigate this trend, the data were requested from Rijkswaterstaat in september 2022. As the dataset was not yet ready and available online, Rijkswaterstaat has sent a number of temporary files covering the period since about 1992 in the South-West Netherlands. These data were processed in a number of scripts documented in this workflow. Results of the analysis are included in the directory structure and have been used in a report on the Westerschelde (Herman, 2022) and will be used in a forthcoming paper.

## Received macrobenthos data and data parsing script

All data received from Rijkswaterstaat have been included in the directory **./data/delivered/**. Older data, obtained between 1992 and 2017, are provided in the file

**./data/delivered/rwsbenthos\_zwdelta\_tijdelijk.csv.**

Data for the years 2018, 2019 and 2020 were delivered as three files per year. The file

**./data/delivered/densbiomyyyy.csv**

contains data on density and biomass of the species in the samples for the year yyyy. In addition, per year a file

**./data/delivered/minfoyyyy.csv**

contains information on the sample points (geography, time of sampling) and on the details of the sampling method. For the year 2020, a third file

**./data/delivered/substr20.csv**

contains information on the substrate (sediment) found in the samples.

The script **./scripts/parse\_data.R** reads the older data and cleans up the file by deselecting empty columns and some other small manipulations. Subsequently, for the years 2018-2020 the different input files are merged into a single data frame per year. Per year, the details of the manipulations differ because the structure of the files also differed per year. Details can be derived from the subtle differences between the scripts of the three years. After these manipulations, all data frames are merged into a single data frame containing all observations between 1992 and 2000. Only Westerschelde data are retained. Similar operations could, in principle, also be performed for the other water bodies of S.W. Netherlands, which are all represented in the delivered files.

The results of the data parsing are stored as a binary file: **./data/processed/wsnew.Rdata** The results are also stored as a .csv file: **./data/processed/wsnew.csv**

## Ecotope maps

Ecotope maps are used in the data exploration to draw maps of biomass with the ecotope as background. In addition, ecotope maps provided information on the surface of the different ecotypes. This information can be used to calculate system-wide averages, by weighing the different ecotypes by their surface area.

Ecotope maps were downloaded from Rijkswaterstaat using `wfsrequests`. The resulting zip files with the shapefiles, were unzipped and stored in directories under `./ecotopenkaart/`. The directories given are the only ecotope maps that exist. For years in between, we have always chosen the map that was closest in time to the given year. The directories contain a file `wfsrequest.txt` that specifies the request needed to download the files. All the other files are needed in order to work with the shapefile.

The downloaded ecotope maps do not contain information on the salinity in the different zones. The script

`./scripts/add_salt_to_shp.R`

reads the shapefiles, adds salt information to them, expresses the surface of the different ecotopes in km<sup>2</sup> rather than in m<sup>2</sup>, and rewrites them. Once the surfaces have been stored in km<sup>2</sup>, this operation will not be repeated if the script is run again. In addition, the script also sums the surfaces of all polygons belonging to the same ecotype, and stores this as

`./data/processes/shpsum_all.Rdata` in binary format, or

`./data/processes/shpsum_all.csv` in .csv format.

## Plot total macrobenthic biomass on ecotope maps (and save the biomass information)

The script in file

`./scripts/plot_biomass_on_ecotope_map.R`

calculates the total biomass of all species per sample, and also log-transforms these biomass data. The script restricts the samples to the Autumn (“Najaar”) because in later years this is the only season sampled. Per year, the total biomass of all samples is plotted on top of the ecotope map, giving an idea of how biomass is distributed in space and in relation to the ecotopes. Output of this script is written to

`./output/kaartjes.pdf`.

As a by-product, the script also produces a file with the total biomass per sample, and the basic ecotype information and geographical position of that sample. This data frame is stored as:

- binary file: `./data/processed/biom_pe.Rdata`
- csv file: `./data/processed/biom_pe.csv`

## Trends by ecotope

The script in the file

`./scripts/trends_by_ecotope.R`

reads in the data file with tot biomass per station, and groups these per (simplified) ecotype. It then plots time series of means  $\pm$  standard error of the total biomass, together with a loess smoother and the standard error of the prediction by this smoother. The y-axis has equal units for all littoral ecotypes, but a different scale for the sublittoral ones, where sometimes very high biomass has been found.

Output with these plots is written to the file `./output/trends_by_ecotope.pdf`.

## Trends by species and species group

The script in the file

**./scripts/trends\_by\_species.R**

reads in the basic data file **./data/processed/wsnew.Rdata**.

It first analyses all the taxa present, and determines the frequency of finding each of these taxa. This information is written to the file

**./data/processed/freq.Csv**.

I have manually manipulated the values in this file, by adding a third column indicating whether the taxon should be dropped or changed into another taxon name. The edited file was written as

**./data/processed/freq\_edit.csv**

and is read by the script. The changes in taxa are performed. For the most frequent taxa, together representing >95% of the total biomass, mean occurrence in the main ecotypes is presented as time series with loess smoothers. Taxa are also grouped into higher-level groups on taxonomic basis (Mollusca, Crustacea, Echinodermata, Polychaeta, Other) and time series as well as stacked graphs are produced for these groups. All results are written to the file

**./output/trends\_by\_species.pdf**.

## References

Herman, P.M.J. 2022. De Schelde, op weg naar een mooie toekomst? Deltares report, Deltares, Delft.

Kruijt, D. B., O. Duijts, M. Japink, and R. P. Middelveld. 2020. Macrozoöbenthosbemonstering in de Zoute Rijkswateren, Hoofdrapport MWTL 2019. Waterlichamen: Westerschelde, Veerse Meer en Grevelingenmeer. Bureau Waardenburg, Culemborg.