Description of the pelagic habitat indicators PH2 and ph3

Operating manual of the R toolbox “MSFD”

01 janvier 2021

Wacquet, G., Aubert, A., Budria, A., Duflos, M., Johansen, M., Louchart, A., Mialet, B., Rombouts, I., & Artigas. L.F.

1. Pelagic Habitats 2 (PH2): Changes in phytoplankton biomass and zooplankton abundance
   1. About the indicator

Contributors :

* Development of the indicator in the frame of OSPAR and the MSFD: Anaïs Aubert (FR), Alexandre Budria (FR), Marie Duflos (FR), Marie Johansen (SW), Benoit Mialet (FR), Isabelle Rombouts (FR), Felipe Artigas (FR).
* Development and improvement of the R script: Guillaume Wacquet (FR).
* Improvement of the methodology and R-script: Arnaud Louchart (FR)

The PH2 indicator is a state and survey indicator, as defined by Bedford *et al.* (2018). It allows the detection of changes in phytoplankton biomass or zooplankton abundances (taxonomic group of the copepods) occurring in a given pelagic assessment unit.

The PH2 indicator computation car be done through an R package using a graphical interface. The package can be provided to user after being requested. The package includes the computation of the PH2 indicator and the plots of the results.

* 1. Data presentation and preparation
     1. Parameters and sampling strategy

Data format and monitoring strategy are detailed in OSPAR CEMP (OSPAR Agreement 2019‐06):

* Zooplankton abundance: only copepods are considered for the computation of the indicator (total abundance of copepods). Total abundance of copepods is considered as proxy of zooplankton abundance.
* Phytoplankton biomass: phytoplankton biomass can be measured as biovolume, carbon content or assessed with indirect indicator using the chlorophyll-*a* which is present in phytoplankton. Semi-quantitative measure of phytoplankton biomass is also possible by using satellite derived products or by using the “Phytoplankton Color Index” (PCI) which is a method applied on the “Continuous Plankton Recorder” (CPR) data.

Monthly sampling frequency is mandatory to obtain a representative assessment. For offshore samples, remote sampling techniques can be used to estimate phytoplankton biomass (chlorophyll-*a* modelling products, Gohin *et al.* 2002, remote-sensing of water color). For coastal samples, *in situ* samples are preferred as remote sampling data.

* + 1. Data preparation

Both phytoplankton and zooplankton data has to be save in a .csv file. The separator must be the semicolon (“;”). The decimal must be the comma (“,”). Each file must contain only three columns: “station” (character format), “date” (format yyyy-mm-dd) and “values” (Fig. 1). The file can contain one or several stations, the script is built to compute indicator per station.

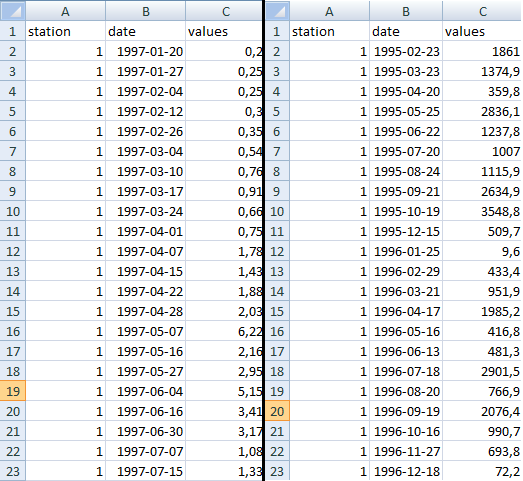


Figure 1 : example of phytoplankton (left) and zooplankton (right) data used for the PH2 indicator calculation.

* 1. Computation of the indicator.

Details on the formula and computation of the PH2 indicator are explained in the OSPAR Indicator PH2 Intermediation Assessment 2017 (IA 2017, section “assessment methods”). However, methodological supplements have been developed and implemented in the PelHabMSFD R package, an evolution of the MSFD R package (Duflos *et al.,* 2018):

* Possibility to select a reference period to calculate the reference seasonal cycle. This can be done to remove possible trends on the assessment period.
* Trends calculation for anomalies. It is a supplementary tool to identify and visualize the periods in which the values of the indicator are generally higher or lower than those of the reference seasonal cycle.

The R package is still under development. It will become available on the CRAN later. Therefore, only one installing procedure of the package is available. Prior to installation, you need to run R studio.

1. Manual installation

Once you have downloaded the package (tar.gz file extension), go in the **Tools** menu of R studio, click on **Install packages** and choose **Package archive file.** Browse your files to reach the **PelHabMSFD** package. Then, click on **install.**

1. Automatic installation

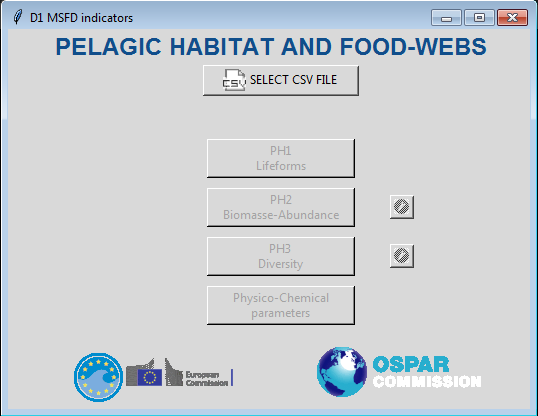
Once the package will be delivered on the CRAN, the automatic installation will be available following the procedure below:

In the **Tools** menu, click on **Install packages.** In the list **“Install from”**, choose **repository (CRAN)** and enter the name package: **PelHabMSFD.** Then, click on **install**

Nb: The installation may fail. It might be possible that all the dependency are not installed. If so, install them manually through the Tools menu and install packages link.

When the package has been installed, load the package with the following command: **require(PelHabMSFD)**

When the package is loaded, launch the graphical interface with the command: **D1\_MSFD()**

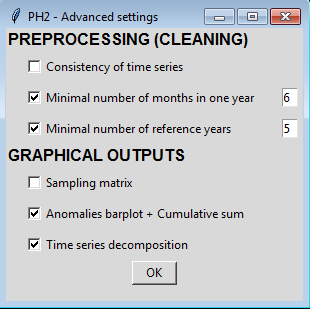


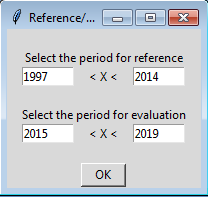
Click on **Select CSV** file to load the file .csv on which you want the calculation of the indicator (see the section “data preparation”).

Note that until any file has been selected, the settings button and the indicator computation are unavailable.

Before running the script for an indicator, click on the Settings button on the right (if available) to open the advanced settings window:

It is recommended to set the different options as shown in the Figure opposite:

* Consistency of time series: if checked, this option only keeps the time series without interruptions. If one or more years are missing for a station, all the years preceding the missing one are not taken into account.
* Minimal number of months in one year: if checked, this option sets a minimum number of months for each year for a given station (6 by default). Each year that does not include this minimum number of months is not taken into account for the calculation.
* Minimal number of reference years: if checked, this option sets a minimum number of years for the reference period (5 by default).
* For each station, several graphic outputs can be generated:
  + Sampling matrix: graphic preview of missing samples.
  + Barplot anomalies + cumulative sums.
  + Time series decomposition: plots with the raw data, the average annual cycle and the anomalies.

After selecting the .csv file and adjusting the advanced parameters, click on the **PH2 indicator** button to start the calculation of the PH2 indicator. The following window appears:

Starting and ending years (included in periods) should be entered for the reference and evaluation periods. It is recommended that the reference and evaluation periods have a duration of 5 years. The calculation of the PH2 indicator for each station begins after clicking **OK**. The results are then stored in the directory containing the .csv file, in a **PH2 Indicator** subdirectory.

An example of files for the PH2 calculation can be found in the repository of the code. The files are labelled “PH2\_phyto\_example.csv” for phytoplankton biomass and “PH2\_zoo\_example.csv” for zooplankton abundance.

1. Pelagic Habitats 3 (PH3): Changes in phytoplankton diversity
   1. About the indicator

Contributors

* Development of the indicator within the framework of OSPAR and MSFD: Isabelle Rombouts (FR), Anaïs Aubert (FR), Marie Duflos (FR), Alexandre Budria (FR) Felipe Artigas (FR).
* Development and improvement of the R script: Guillaume Wacquet (FR).
* Improvement of the methodology and R-script: Arnaud Louchart (FR)

PH3 is a state and monitoring indicator, as defined by Bedford et al. (2018), which can detect changes in phytoplankton or zooplankton diversity that occur in a given pelagic assessment unit. The methodology for the calculation and implementation of PH3 for phytoplankton in particular is detailed in Rombouts et al. (2019).

For the calculation of the PH3 indicator, an R package containing an ergonomic graphical interface has been developed and can be provided to all potential users on request. The package includes the calculation of the PH3 indicator, and the graphical representations of the results.

* 1. Data presentation and preparation
     1. Parameters and sampling strategy

The data format, as well as the monitoring strategy are detailed in the OSPAR CEMP guide (OSPAR, 2019). For the calculation of PH3, taxonomic composition data of zooplankton and phytoplankton with a minimum monthly sampling frequency are required. Taxonomic resolution is at the species level at best, however since taxonomic data is often at different taxonomic levels, the genus level is often used to reduce bias (Rombouts et al., 2019).

* + 1. Data preparation

Phytoplankton and zooplankton data should be saved in .csv files, using the semicolon as the field separator and the comma as the decimal character. For the R script, two types of formats are accepted. First, the files can be organized so that the lines correspond to the taxa (Figure 2). In this case, the file for calculating the indicator must contain up to 8 columns: “station” (character format), “day” (numeric format), “month” (numeric format), “year” (format numeric), “longitude” (numeric format), “latitude” (numeric format), “taxon” (character format), “abundance” (numeric format).

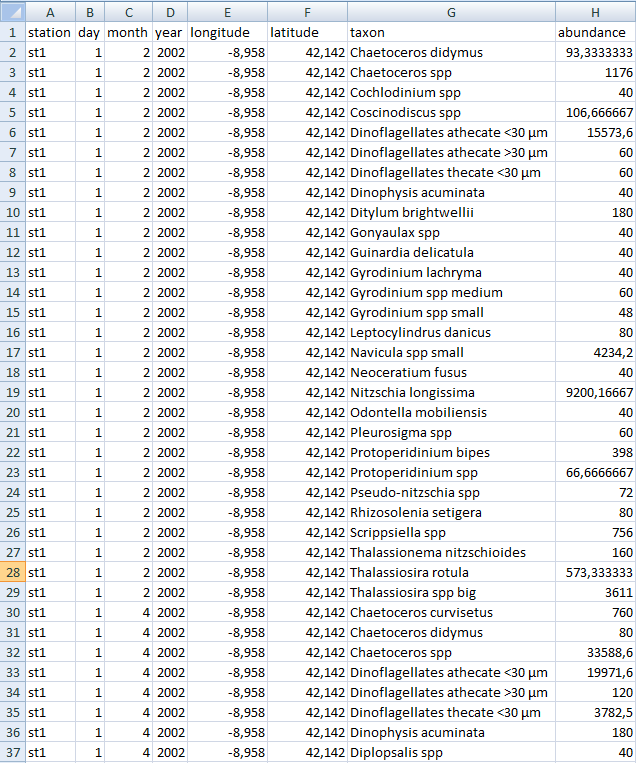


Figure 2 : Example of phytoplankton data used (taxonomic level : genus) for the PH3 indicator calculation (setting: taxa in rows').

Alternatively, files can be organized such that each row corresponds to a sample and the columns correspond to taxa (Figure 3). In this case, the file for calculating the indicator must contain up to 6 columns for metadata: “station” (character format), “day” (numeric format), “month” (numeric format), “year "(numeric format)," longitude "(numeric format)," latitude "(numeric format). The following columns then correspond to the abundances by species.

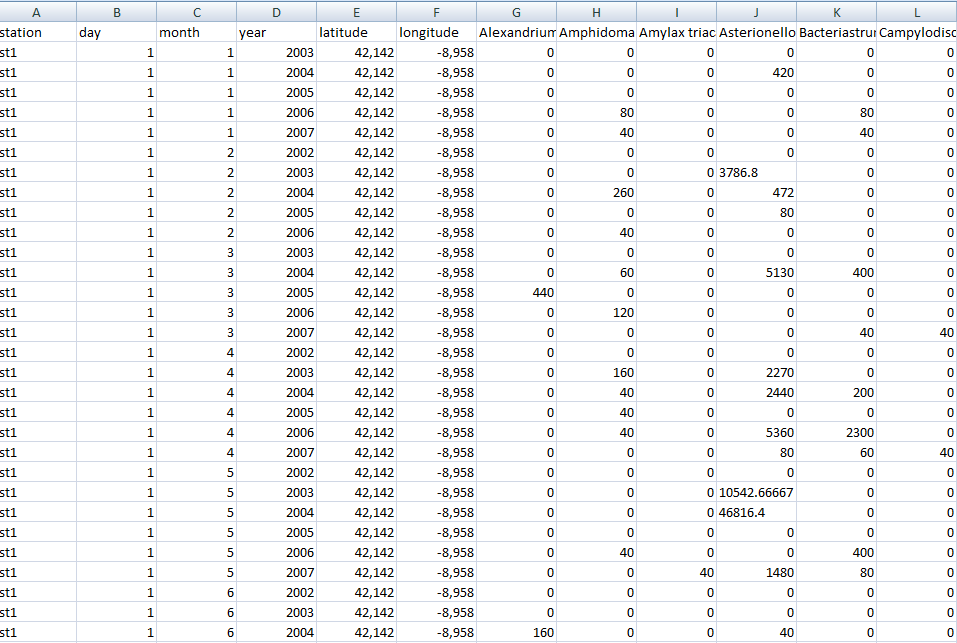


Figure 3: Example of phytoplankton data used for PH3 indicator calculation (setting: 'taxa in columns').

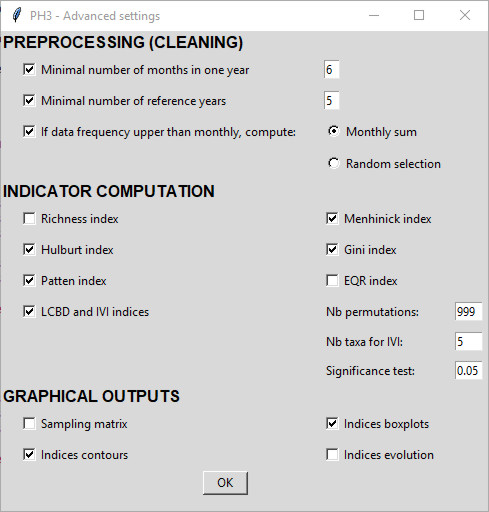
The package does not currently allow to merge files with different formats (taxa in columns and taxa in rows). However, software developments could eventually make it possible to perform this task and thus perform the analyzes and the indicator calculation in batch (in an automated manner) on several files.

* 1. Computation of the indicator

Details on the calculation of the PH3 indicator are explained in OSPAR Indicator PH2 Intermediation Assessment 2017 (IA 2017, section “assessment methods”). A detailed description can also be found in Rombouts et al., 2019.

The alpha diversity indices are calculated on the monthly abundances, while the Local Contribution to Beta Diversity (LCBD) is calculated on annual averages because its purpose is to compare the years of the time series. Only in the event that a year has a statistically significant and significant LCBD value, the Important Value Index (IVI) is calculated. These alpha and beta diversity indices must be considered together to correctly interpret potential changes in the diversity of plankton communities.

The procedure for installing the R package is described in the PH2 indicator section. Before running the script for an indicator, click on the Settings button on the right (if available) to open the advanced settings window:



It is recommended to set the different options as shown in Figure above:

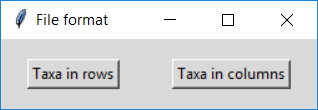
* Minimal number of months in one year: if checked, this option sets a minimum number of months for each year for a given station (6 by default). Each year that does not include this minimum number of months is not taken into account for the calculation. It is recommended to use monthly data. If not available, comparisons between years with different sampling frequencies should be made with caution.
* Minimal number of reference years: if checked, this option sets a minimum number of years for the reference period (5 by default). For the calculation of the LCBD, a minimum of 10 years is recommended to obtain relevant results.
* If data frequency upper than monthly: if checked, either the abundances obtained in the same month are summed (Monthly sum), or a date in the month is randomly selected (Random selection).

Several graphic outputs can be selected and saved automatically:

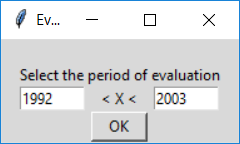
* Sampling matrix provides a graphical overview of missing samples.
* Contour Indices provides the “contourplots” of the diversity indices per year and per month. This output can be very useful for detecting seasonal variations and long term trends.
* “Boxplots” indices (per month and per year) for each calculated index.
* "Evolution" indices provide a graph of the evolution of indices calculated as a function of time. All the indices are displayed on a single graph.

After selecting the .csv file and adjusting the advanced parameters, click on the **PH3 indicator** button to start the calculation of the PH3 indicator.

The following window appears:

Two file formats are allowed: "Taxa in rows" or "Taxa in columns". In this package, all the functions have been developed to deal with columnar taxa. However, if the taxa are in rows, a simple transformation of the data is automatically performed before any processing.

After this step, a new window appears:

Starting and ending years (included in periods) must be entered for the evaluation period. The calculation of the PH3 indicator for each station begins after clicking **OK**. The results are then stored in the directory containing the .csv file, in a **PH3 Indicator** subdirectory

An example of files for the PH3 calculation can be found in the repository of the code. The files are labelled “PH3\_phyto\_long\_example.csv” for phytoplankton abundance with taxa in rows and “PH3\_phyto\_wide\_example.csv” for phytoplankton abundance with taxa in columns. Zooplankton abundance datasets can be found under the name ”PH3\_zoo\_long\_example.csv” for zooplankton abundance with taxa in rows and “PH3\_zoo\_wide\_example.csv” for zooplankton abundance with taxa in columns.

1. References

Bedford, J., Johns, D., Greenstreet, S., & McQuatters-Gollop, A. (2018). Plankton as prevailing conditions: A surveillance role for plankton indicators within the Marine Strategy Framework Directive. *Marine Policy*, *89*, 109-115.

Duflos, M., Wacquet, G., Aubert, A., Rombouts, I., Mialet, B., Devreker, D., Lefebvre, A., & Artigas, L. F. (2018). DCSMM2018 - HP - D1C6 - Ratio de la concentration de surface en chlorophylle a contenue dans les dinoflagellés sur celle contenue dans les diatomées (2012-2016). Ifremer.

Gohin, F. (2012). Répartition spatio-temporelle de la chlorophylle a. Sous-région marine Mers celtiques. Evaluation initiale DCSMM.

Jouandet, M. P., Petit, L., Duflos, M., Wacquet, G., Dédécker, C., Aubert, A., Rombouts, I., Mialet, B., Lebourg, E., Bruaut, M., Gómez, F., Nowaszyk, A., Delarbre, J., Cornille, V., Lécuyer, E., Didry, M., Lizon, F., & Artigas, L. F. (2019) Mise en œuvre du porgramme de surveillance « Habitats Pélagiques » : test de mise en place des campagnes saisonnières et calcul d’indicateurs sur le site atelier Manche Mer du Nord et Mers Celtiques. 127 pp.

OSPAR (2019). CEMP Guideline. PH3 Changes in plankton diversity. 7 pp.

Rombouts, I., Simon, N., Aubert, A., Cariou, T., Feunteun, E., Guérin, L., Hoebeke, M., McQuatters-Gollop, A., Rigaut-Jalabert, F., & Artigas, L. F. (2019). Changes in marine phytoplankton diversity: Assessment under the Marine Strategy Framework Directive. *Ecological Indicators*, *102*, 265-277.