Descriptif des traitements pour les évaluations DCE Phytoplancton

# Objectif du document

Description de l'extraction et des traitements destinés à la production du document Evaluation DCE Phytoplancton

**Un document similaire existe pour l'évaluation des indicateurs physico-chimiques.**

Traitement des résultats de la métropole et des DOMS.

## Perspectives : discutées lors d’une visio VIGIES + Anne Daniel en juin 2017

Les évaluations phyto et hydro ne tournent pas tout à fait de la même façon, il faut les uniformiser.

A court terme, les modifications minimales à faire pour qu'elles sortent rapidement sont :

* sur les cartes : (i) chaque ME est évaluée en couleur même si pas assez de résultats, (ii) les ME apparaissant en gris ("non évaluée") regroupent des cas différents comme : ME sans aucun résultat ou ME n'ayant pas à être évaluée
* toutes les ME ont une fiche, même vide

A moyen terme, il faut prévoir de décrire des modifications plus substantielles, pour que ce soit plus lisible, comme par exemple :

* faire une liste de catégories de ME : évaluable mais pas évaluée pour telle ou telle raison, non évaluable, etc
* améliorer les commentaires mis en automatique en bas des fiches ME
* faire apparaître sur les cartes quelque chose qui ressemble à de la confiance dans le résultat, par ex un hachurage de la couleur si l'évaluation est faite sur un nombre de données en deçà d'un pourcentage de données attendues

# Rappel des documents et fichiers produits lors de l'évaluation

Pour chaque masse d'eau (ME), voir exemple en **annexe 1** :

* une fiche détaillant les résultats de l'évaluation
* une fiche descriptive avec tableaux et graphiques
* une fiche des taxons contributifs

Des cartes synthétisant les résultats par région (une par indice + une pour l'indicateur) : voir exemple en **annexe 2**

Des fichiers intermédiaires : voir contenu en **annexe 3**

Un fichier Zip pour mise à disposition sur INTRANET

# Description générale

en orange : les modifications / suppressions à faire ultérieurement

en bleu : les modifications apportées à la demande des Antilles (visio du 28/06).

*tracelog.txt Log*

Dix programmes sont exécutés : leur contenu est décrit ci-dessous dans l'ordre de leur exécution.

Des tables complémentaires sont utilisées par les programmes.

Le fichier support *ME et points DCE* : identifie les ME et les points à prendre en compte dans les évaluations, et fournit des éléments descriptifs pour les fiches ME. Ce fichier est commun à Phytoplancton et Physico-chimie. Il se trouve ici :

\\nantesdir\vigiestat\International\Rapports et etudes\Simulation DCE\Documents\00-General\2021\ME groupes ME et points DCE phyto hydro - AD et ML - 05-07-2021

## 1. Data\_Bio\_Phyto\_Flores

Concerne les flores phytoplancton (dénombrements).

### Extraction Q²

#### Critères de sélection

Programmes Q² :

* REPHY
* SRN
* RHLN
* ARCHYD
* RSLHYD
* LA\_REUNION\_ETUDES\_DIVERSES

*Info DOMs : les suivis DCE sont normalement dans le programme REPHY pour tous les DOMs. Il arrive que des résultats acquis dans le cadre de programmes spécifiques soient pertinents pour être ajoutés.*

Aucune sélection sur les lieux.

Période : six années pleines.

Paramètres :

* FLORTOT
* FLORIND

Aucune sélection sur l'état de validation ou de qualification.

#### Champs à extraire

* Lieu de surveillance : Identifiant
* Lieu de surveillance : Libellé
* Lieu de surveillance : Mnémonique
* Résultat : Code des programmes
* Passage : Date
* Prélèvement : Niveau
* Echantillon : Identifiant interne
* Résultat : Code paramètre
* Résultat : Nom du taxon
* Résultat : Libellé du groupe de taxon
* Résultat : Valeur de la mesure (champ qui permet également d’extraire « Résultat : Valeur qualitative »)
* Résultat : Niveau de qualité
* Résultat : Service saisisseur : Libellé
* Résultat : Libellé précision
* Résultat : Date de validation

### Importation du fichier d'extraction

Suppression des enregistrements sans résultat (étape de vérification, reste 206 408).

Suppression des enregistrements sans n° d'échantillon (pour éliminer résultats sur prélèvements).

Suppression des enregistrements pour lesquels le lieu n'est pas présent dans le fichier support *ME et points DCE* (à ce stade tous les lieux du fichier sont gardés, puis la sélection des lieux à évaluer pour le phyto est faite plus loin).

Suppression des enregistrements sans code masse d’eau.

Suppression des enregistrements autres que BON ou non qualifiés. Une exception : les enregistrements DOUTEUX pour les lieux décrits dans une table annexe DIREN-Bretagne[[1]](#footnote-1) sont également gardés.

### Traitements

Suppression des enregistrements sans code taxon dénombré.

Suppression des enregistrements non validés.

Garder une seule valeur (le max) par ME/échantillon/taxon.

Garder seulement les niveaux suivants :

* pour toutes les ME sauf FRI, FRJ, FRK, FRL et FRM :
  + « Surface (0-1m) »,
  + « Mi-profondeur »,
  + « Surface-Fond (profondeur <3 m) ;
* pour les ME FRI, FRJ et FRK (Guadeloupe, Martinique et Guyane) :
  + « Surface (0-1m) »,
  + « Surface-Fond (profondeur <3 m) ;
* pour les ME FRL (La Réunion) et FRM (Mayotte) :
  + « Surface (0-1m) ».

Garder une seule valeur (le max) par ME/id lieu/date/taxon.

Sélection des seuls lieux évalués pour le paramètre étudié.

## 2. Data\_Bio\_Phyto\_Chla

Concerne la chlorophylle-a.

### Extraction Q²

#### Critères de sélection

Programmes Q² :

* REPHY
* SRN
* RHLN
* ARCHYD
* RSLHYD
* LA\_REUNION\_ETUDES\_DIVERSES

*Info DOMs : les suivis DCE sont normalement dans le programme REPHY pour tous les DOMs. Il arrive que des résultats acquis dans le cadre de programmes spécifiques soient pertinents pour être ajoutés.*

Aucune sélection sur les lieux.

Période : six années pleines.

Paramètre :

* CHLOROA.

Aucune sélection sur l'état de validation ou de qualification.

#### Champs à extraire

* Lieu de surveillance : Identifiant
* Lieu de surveillance : Libellé
* Lieu de surveillance : Mnémonique
* Résultat : Code des programmes
* Passage : Date
* Passage : Mnémonique
* Prélèvement : Niveau
* Echantillon : Identifiant interne
* Résultat : Code paramètre
* Résultat : Valeur de la mesure (champ qui permet également d’extraire « Résultat : Valeur qualitative »)
* Résultat : Niveau de qualité
* Résultat : Libellé méthode
* Résultat : Service saisisseur : Libellé
* Résultat : Libellé précision
* Résultat : Date de validation

### Importation du fichier d'extraction

Suppression des enregistrements sans résultat (étape de vérification, reste 16 281).

Suppression des enregistrements sans n° d'échantillon (pour éliminer résultats sur prélèvements).

Suppression des enregistrements pour lesquels le lieu n'est pas présent dans le fichier support *ME et points DCE* (à ce stade tous les lieux du fichier sont gardés, puis la sélection des lieux à évaluer pour l’indicateur phyto est faite plus loin).

Suppression des enregistrements sans code masse d’eau.

Suppression des enregistrements autres que BON ou non qualifiés. Une exception : les enregistrements DOUTEUX pour les lieux décrits dans une table annexe DIREN-Bretagne1 sont également gardés.

### Traitements

Suppression des enregistrements non validés.

**Spécificité ARCHYD :** suppression de tous les enregistrements du programme ARCHYD dont le mnémo passage ne contient pas PM (c.-à-d. suppression des passages qui ne sont pas faits à pleine mer).

**Gestion des méthodologies**

Pour toutes les passages, on calcule le nombre de méthodes qui ont été utilisées pour mesurer la chloro.

On crée deux data frames, un pour lequel il y a des données en doublons (deux méthodes) et un autre pour lequel il n’y a pas de doublons.

**Pour les ME FRI (Guadeloupe) et FRJ (Martinique) :**

* **SI UN ECHANTILLON PORTE DEUX OU PLUSIEURS RESULTATS**, on garde les méthodes  :
  + « Chromatographie liquide - pigments phytoplanctoniques (Van Heukelem et Thomas 2001) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Zapata et al. 2000) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Wright et al. 1991) - µg/l » ;
* ne pas prendre les données avant août 2013 pour la Martinique et avant 2016 pour la Guadeloupe ;
* garder seulement les niveaux suivants :
  + « Surface (0-1m) » ;
* dans le cas de duplicats (ou +), prendre la valeur moyenne par échantillon.

**Pour les ME FRK (Guyane) :**

* **SI UN ECHANTILLON PORTE DEUX OU PLUSIEURS RESULTATS**, on garde les méthodes :
  + « Chromatographie liquide - pigments phytoplanctoniques (Van Heukelem et Thomas 2001) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Zapata et al. 2000) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Wright et al. 1991) - µg/l » ;
* garder seulement les niveaux suivants :
  + « Surface (0-1m) »,
  + « Surface-Fond (profondeur <3 m) ».

**Pour les ME FRM (Mayotte) :**

* **SI UN ECHANTILLON PORTE DEUX OU PLUSIEURS RESULTATS**, on garde les méthodes :
  + « Chromatographie liquide - pigments phytoplanctoniques (Van Heukelem et Thomas 2001) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Zapata et al. 2000) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Wright et al. 1991) - µg/l » ;
* garder seulement les niveaux suivants :
  + « Surface (0-1m) »,
  + « Surface-Fond (profondeur <3 m) ».

**Pour les ME FRL (La Réunion) :**

* **SI UN ECHANTILLON PORTE DEUX OU PLUSIEURS RESULTATS**, on garde les méthodes :
  + « Chromatographie liquide - pigments phytoplanctoniques (Van Heukelem et Thomas 2001) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Zapata et al. 2000) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Wright et al. 1991) - µg/l » ;
* garder seulement les niveaux suivants :
  + « Surface (0-1m) ».

**Pour les ME de Manche-Atlantique (FRA,FRF,FRG,FRH) :**

* **,** on **SUPPRIME** les méthodes
* garder seulement les niveaux suivants :
  + « Surface (0-1m) »,
  + « Surface-Fond (profondeur <3 m) ;

**Pour les ME de Méditerranée (FRD,FRE) :**

* **SI UN ECHANTILLON PORTE DEUX OU PLUSIEURS RESULTATS,** on garde les méthodes :
  + « Chromatographie liquide - pigments phytoplanctoniques (Van Heukelem et Thomas 2001) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Zapata et al. 2000) - µg/l »,
  + « Chromatographie liquide - pigments phytoplanctoniques (Wright et al. 1991) - µg/l ».
* garder seulement les niveaux suivants :
  + « Surface (0-1m) »,
  + « Surface-Fond (profondeur <3 m) ;

Garder une seule valeur (le max) par échantillon.

Garder une seule valeur (le max) par id lieu/date.

Sélection des seuls lieux évalués pour le paramètre étudié.

## 3. Data\_Bio\_Phyto\_CMF

Concerne les résultats de cytométrie en flux pour les lagunes de métropole strictement, pour lesquelles ces résultats sont traités globalement sur nano et pico, sans autres détails.

**Attention** : il faudra prévoir un autre module de traitement des données cytométrie en flux pour les résultats plus détaillés qui existent déjà dans les DOMs (Martinique, Réunion, Mayotte, etc.).

Mais pour prendre en compte ces résultats, il faut construire des grilles correspondant à ces nouveaux paramètres, ce n’est pas encore fait. Pour Martinique, Felipe Artigas y travaille ? Pour le reste, ce sera à dire d’expert dans un premier temps.

### Extraction Q²

#### Critères de sélection

Programme Q² :

* RSLHYD

Aucune sélection sur les lieux.

Période : six années pleines.

Paramètres :

* NANOSUP3
* PEUKINF3

Aucune sélection sur l'état de validation ou de qualification.

#### Champs à extraire

* Lieu de surveillance : Identifiant
* Lieu de surveillance : Libellé
* Lieu de surveillance : Mnémonique
* Résultat : Code des programmes
* Passage : Date
* Prélèvement : Niveau
* Echantillon : Identifiant interne
* Résultat : Code paramètre
* Résultat : Valeur de la mesure (champ qui permet également d’extraire « Résultat : Valeur qualitative »)
* Résultat : Niveau de qualité
* Résultat : Service saisisseur : Libellé
* Résultat : Libellé précision
* Résultat : Date de validation

### Importation du fichier d'extraction

Suppression des enregistrements sans résultat (étape de vérification, reste 1 731).

Suppression des enregistrements sans n° d'échantillon (pour supprimer les résultats sur prélèvement).

Suppression des enregistrements pour lesquels le lieu n'est pas présent dans le fichier support *ME et points DCE* (à ce stade tous les lieux du fichier sont gardés, puis la vraie sélection des lieux à évaluer pour cet indice est faite plus loin).

Suppression des enregistrements sans code masse d’eau.

Suppression des enregistrements autres que BON ou non qualifiés.

### Traitements

Suppression des enregistrements non validés.

Garder seulement les niveaux suivants :

* « Surface (0-1m) » ;
* « Surface-Fond (profondeur <3 m) ».

Garder une seule valeur (le max) par échantillon/paramètre.

Garder une seule valeur (le max) par id lieu/date/paramètre.

## 4. Data\_Bio\_Phyto\_Flores\_Abondance\_Taxons

### Seuils de blooms par taxon

* Pour les ME FREC (Corse) : 25 000 cellules par litre.
* Pour les ME FRLC (Réunion côtières) et FRMC (Mayotte côtières) : 10 000 cellules par litre.
* Pour les ME FRIC (Guadeloupe côtières), FRJC et FRJT (Martinique côtières et transition) : 25 000 cellules par litre pour les grosses cellules, les petites cellules étant évaluées à part (indicateur en cours de développement).
* Pour les ME FRKC (Guyane côtières) : 1 000 000 cellules par litre (info Luis Lampert).
* Pour toutes les autres ME : 100 000 cellules par litre pour les espèces de taille ≥ 20 μm ou 250 000 cellules par litre pour les espèces de taille 5 μm ≤ x ≤ 20 μm (déterminée dans la table *Bio-Phyto-Flores-Taxons-Q2[[2]](#footnote-2)*).

### Détermination de l’état de bloom par taxon

Reprise du fichier de sortie du programme 1 « Data\_Bio\_Phyto\_Flores » : comparaison des taxons avec la table des tailles par taxon *Bio-Phyto-Flores-Taxons-Q²* (sauf pour les taxons dont la taille est indiquée à 0).

Warning **non bloquant** si des taxons n’existent pas : reboucle si c’est le cas, après mise à jour de la table des tailles.

Pour les Antilles (FRI et FRJ) : retrait des petites cellules (taille 1).

Sélection de données avec le fichier GroupeME, duquel les lagunes mesohalines et oligohalines ont été retirées.Warning **non bloquant** si des ME n’existent pas : reboucle si c’est le cas, après mise à jour de la table des ME.

Comparaison des valeurs avec les seuils pour définir si le taxon est à l’état de bloom ou pas (1 = bloom ; 0 = non bloom).

## 5. Data\_Bio\_Phyto\_Flores\_Composition\_Taxons

Ce programme est débranché

## 6. Data\_Evaluation\_Bio\_Phyto\_Biomasse

Reprise fichier du programme 2 « Data\_Bio\_Phyto\_Chla ».

Jointure fichier avec le fichier support *ME*, duquel les lagunes mesohalines et oligohalines ont été retirées.

### Premier lot : ME méditerranée

Filtrage sur groupes MEC et MET Méditerranée : 1, 2A, 3W, Corse (pour les MEC) + Delta, Lagunes (pour les MET).

Sélection du seul groupe MET Lagunes.

Puis filtrage sur les seuls mois de juin à août.

Sélection de toutes les autres ME (hors lagunes).

### Deuxième lot : DOMs

Filtrage sur groupe « EC Réunion » (ne pas inclure ER Réunion).

Filtrage sur groupes « EC Martinique » et « ET Martinique ».

Filtrage sur groupe « EC Guadeloupe ».

Filtrage sur groupe « EC Mayotte ».

Filtrage sur groupe « ET Mayotte type delta ».

Filtrage sur groupe « EC Guyane ».

Assemblage de tous les DOMs.

### Troisième lot : ME non méditerranéennes et non DOM

Filtrage sur les mois de mars à octobre.

### Assemblage des trois lots

Table des grilles pour la biomasse : *Bio-Phyto-Chla.txt*

Filtrage sur point.phytoplancton = OUI (détermine quels sont les points à évaluer pour l’indicateur phyto) s’il s’agit de l’évaluation nationale. Sur demande et pour certaines régions, il est possible de faire une évaluation dite « régionale », utilisant la colonne point.phytoplancton.supplémentaire en plus de la colonne point.phytoplancton. Dans ce cas, la valeur « OUI » est attribuée à la colonne point.phytoplancton pour les points supplémentaires dans le script « Data\_Bio\_Phyto\_Chla ».

Agrégation temporelle : sélection d’une seule valeur par lieu/année/mois (la première du mois).

Séparer les ME lagunes de toutes les autres.

Sur ME lagunes - agrégation spatiale : sélection d’une seule valeur par ME/année/mois (moyenne).

ME autres que lagunes.

Sur toutes ME autres que lagunes - agrégation spatiale : sélection d’une seule valeur par ME/année/mois (max).

Assemblage de toutes les ME.

## 7. Data\_Evaluation\_Bio\_Phyto\_Abondance

Reprise fichiers des programmes 3 « Data\_Bio\_Phyto\_CMF » et 4 « Data\_Bio\_Phyto\_Flores\_Abondance\_Taxons »

Jointure fichier avec le fichier support *ME*, duquel les lagunes mesohalines et oligohalines ont été retirées.

### Sur fichier CMF

Filtrage sur les lagunes méditerranéennes.

Filtrage sur les seuls mois de juin à août.

Ajout des deux tables décrivant les grilles : *Bio-Phyto-CMF-NANOSUP3* et *Bio-Phyto-CMF-PEUKINF3* (voir annexe 5).

Filtrage sur point.phytoplancton = OUI (détermine quels sont les points à évaluer pour l’indicateur phyto) s’il s’agit de l’évaluation nationale. Sur demande et pour certaines régions, il est possible de faire une évaluation dite « régionale », utilisant la colonne point.phytoplancton.supplémentaire en plus de la colonne point.phytoplancton. Dans ce cas, la valeur « OUI » est attribuée à la colonne point.phytoplancton pour les points supplémentaires dans le script « Data\_Bio\_Phyto\_CMF ».

Agrégation temporelle : sélection d’une seule valeur par param/lieu/année/mois (la première du mois).

Agrégation spatiale : sélection d’une seule valeur par param/ME/année/mois, tous lieux confondus (la moyenne).

### Sur fichier flores

Ajout de la table décrivant les grilles : *Bio-Phyto-Flores-Abondance* (voir annexe 6).

Suppression des flores des lagunes

Codage de l'état bloom par échantillon : 0 (pas de taxon en état bloom) ou 1 (au moins un taxon en état bloom).

Filtrage sur point.phytoplancton = OUI (détermine quels sont les points à évaluer pour l’indicateur phyto) s’il s’agit de l’évaluation nationale. Sur demande et pour certaines régions, il est possible de faire une évaluation dite « régionale », utilisant la colonne point.phytoplancton.supplémentaire en plus de la colonne point.phytoplancton. Dans ce cas, la valeur « OUI » est attribuée à la colonne point.phytoplancton pour les points supplémentaires dans le script « Data\_Bio\_Phyto\_Flores ».

Sélection d'une seule valeur pour un même jour, cas de plusieurs échantillons (valeur max).

Agrégation temporelle : sélection d’une seule valeur par ME/lieu/année/mois (la première du mois).

Agrégation spatiale : sélection d’une seule valeur par ME/année/mois (max).

### Assemblage des fichiers flores et CMF

## 8. Data\_Evaluation\_Bio\_Phyto\_Composition

Ce programme est débranché

## 9. Data\_Evaluation\_Bio\_Phyto

Rassembler fichiers biomasse (programme 6 « Data\_Evaluation\_Bio\_Phyto\_Biomasse ») et abondance (programme 7 « Data\_Evaluation\_Bio\_Phyto\_Abondance »).

Exécuter fonction EQR phytoplancton par ME

### Biomasse

calcul P90 par ME

calcul EQR par ME

calcul indices confiance, etc

### Abondance

#### Lagunes

calcul P90 par ME/paramètre

calcul EQR par ME/paramètre, puis seul le minimum des EQRs par ME est retenu (le plus déclassant)

calcul indices confiance, etc

#### ME autres que lagunes

calcul % d'échantillons avec blooms par ME

calcul EQR par ME

calcul indices confiance, etc

### Assemblage Biomasse et Abondance

Comparaison aux grilles

Attribution d'une classe

Mise en forme des tableaux de la fiche ME

Dans le cadre de l'Abondance, pour les lagunes (Code Typologie « T10 ») les valeurs de l'indice et de la grille de l'indice sont mises à blanc.

**N.B. des ajouts sont faits spécifiquement pour les lagunes, avec production de graphiques et de tendances : voir annexe 7 (Antonin Gimard).**

## 10. Table\_Evaluation\_Bio\_Phyto\_Fiche\_ME

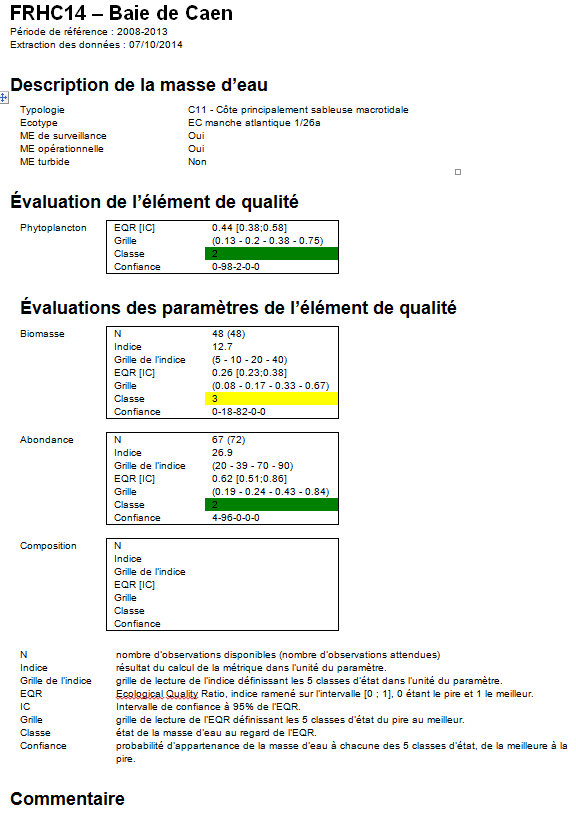
Fourniture de la table : Il s'agit du dernier fichier du programme 9, dont la structure est détaillée dans l'annexe 3.

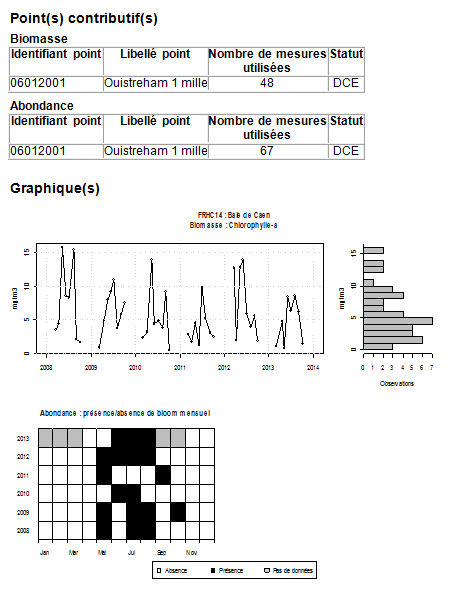
Sont aussi décrits dans ce programme : un patron de fiche ME par type de fiche.

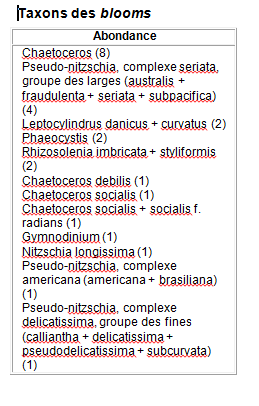
## 11. Figure\_Evaluation\_Bio\_Phyto

Sont décrits dans ce programme : un patron de fiche contribution par type de fiche.

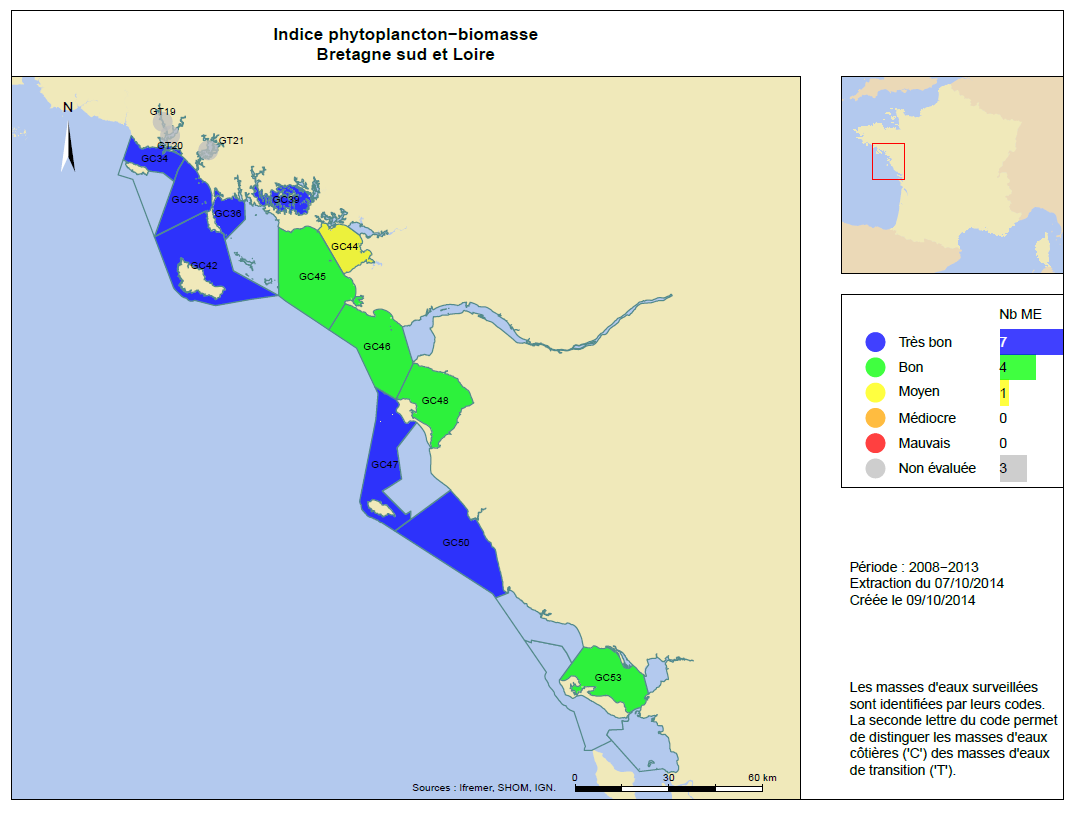
Annexe 1 : Exemple de fiche ME, avec ses annexes

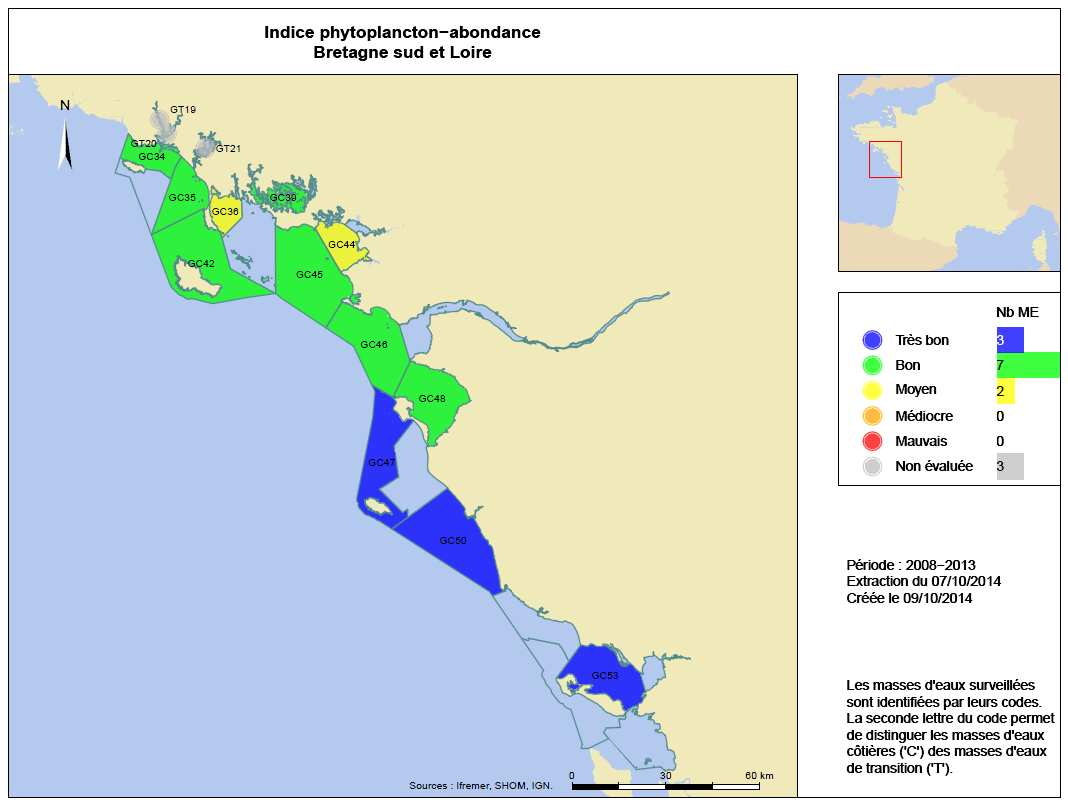






Annexe 2 : Exemple de cartes synthétisant les résultats par région





Annexe 3 : Contenu des fichiers fournis avec l'évaluation

## Evaluation\_Bio\_Phyto

Cette table récapitulative est fournie avec les fiches ME. Le tableau ci-dessous récapitule les en-têtes des colonnes.

|  |
| --- |
| GroupeME |
| CodeMasseEau |
| LibelléMasseEau |
| NumOrdre |
| MESurv |
| MEInterc |
| METurbide |
| MEControleOpérationnelPhytoSelonLER |
| MEControleSurveillanceOfficielleLER |
| BiomasseN |
| BiomasseIndiceValeur |
| BiomasseIndiceGrille |
| BiomasseEQRValeur |
| BiomasseEQRIC |
| BiomasseEQRClasse |
| BiomasseEQRGrille |
| BiomasseEQRConfiance |
| AbondanceN |
| AbondanceIndiceValeur |
| AbondanceIndiceGrille |
| AbondanceEQRValeur |
| AbondanceEQRIC |
| AbondanceEQRClasse |
| AbondanceEQRGrille |
| AbondanceEQRConfiance |
| CompositionN |
| CompositionIndiceValeur |
| CompositionIndiceGrille |
| CompositionEQRValeur |
| CompositionEQRIC |
| CompositionEQRClasse |
| CompositionEQRGrille |
| CompositionEQRConfiance |
| PhytoplanctonEQRValeur |
| PhytoplanctonEQRIC |
| PhytoplanctonEQRClasse |
| PhytoplanctonEQRGrille |
| PhytoplanctonEQRConfiance |
| ExtractionDesDonnées |
| PeriodeDeReference |

## Compilation\_Taxon\_Abondance

Pour chaque ME (en ligne) : la fréquence de chacun des taxons (en colonnes) présents dans l'ensemble du fichier d'extraction.

Un extrait :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Var1 | Freq.Akashiwo sanguinea | Freq.Alexandrium | Freq.Amphora | Freq.Asterionella + Asterionellopsis + Asteroplanus | Freq.Asterionellopsis glacialis | Freq.Attheya armatus | etc |
| FRAC01 | 0 | 0 | 0 | 0 | 2 | 0 |  |
| FRAC02 | 0 | 0 | 0 | 0 | 2 | 0 |  |
| FRAC03 | 0 | 0 | 0 | 0 | 3 | 0 |  |
| FRAC05 | 0 | 0 | 0 | 0 | 6 | 0 |  |
| FRAT01 | 0 | 0 | 0 | 0 | 32 | 5 |  |
| FRDC01 | 0 | 0 | 0 | 0 | 0 | 0 |  |
| FRDC02a | 0 | 0 | 0 | 0 | 1 | 0 |  |
| FRDC02c | 0 | 0 | 0 | 0 | 0 | 0 |  |
| etc |  |  |  |  |  |  |  |

Annexe 4 : Table des grilles pour la biomasse

Cette table est : *Grilles-Bio-Phyto-Chla.txt*, elle est dans :

\\nantesdir\vigiestat\International\Rapports et etudes\Simulation DCE\Programs\01-Phytoplancton

GroupeME Grille ValeurReference

EC Mer du Nord 1/26b 0,10,15,22.5,45,1000 6.67

EC Manche Atlantique 1/26a 0,5,10,20,40,1000 3.33

ET Mer du Nord 0,10,15,22.5,45,1000 6.67

ET Manche Atlantique 0,5,10,20,40,1000 3.33

EC Méditerranée type 1 0,5,10,20,40,1000 3.33

EC Méditerranée type 2A 0,2.4,3.6,7.2,14.4,1000 1.9

EC Méditerranée type 3W 0,1.1,1.8,3.6,7.2,1000 0.9

ET Méditerranée type delta 0,5,10,20,40,1000 3.33

ET Méditerranée type lagune 0,5,7,10,20,1000 3.33

EC Méditerranée type corse 0,0.75,1.22,2.44,4.88,1000 0.6

EC Réunion 0,0.6,0.9,1.8,3.7,1000 0.4

EC Martinique 0,0.3,0.6,1.2,2.4,1000 0.2

ET Martinique 0,0.3,0.6,1.2,2.4,1000 0.2

EC Guadeloupe 0,0.3,0.6,1.2,2.4,1000 0.2

EC Mayotte 0,0.6,0.9,1.8,3.7,1000 0.4

ET Mayotte type delta 0,5,10,20,40,1000 3.33

EC Guyane 0,10,15,22.5,45,1000 6.67

*Pour Martinique et Guadeloupe : infos JP Allenou, seuils en cours de révision*

*Pour Réunion et Mayotte : infos Magali Duval*

*Pour Guyane : infos Luis Lampert*

Annexe 5 : Table des grilles pour la cytométrie en flux pour les lagunes

Les tables sont :

*Grilles-Bio-Phyto-CMF-NANOSUP3*

GroupeME Grille ValeurReference

ET Méditerranée type lagune 0,4,10,20,100,100000 3

*Grilles-Bio-Phyto-CMF-PEUKINF3*

GroupeME Grille ValeurReference

ET Méditerranée type lagune 0,20,50,100,500,100000 15

Elles sont dans :

\\nantesdir\vigiestat\International\Rapports et etudes\Simulation DCE\Programs\01-Phytoplancton

Annexe 6 : table des grilles pour l’indice abondance (blooms)

Cette table est : *Grilles-Bio-Phyto-Flores-Abondance.* Elle est dans :

\\nantesdir\vigiestat\International\Rapports et etudes\Simulation DCE\Programs\01-Phytoplancton

GroupeME Grille ValeurReference

EC Mer du Nord 1/26b 0,20,39,70,90,100 16.7

EC Manche Atlantique 1/26a 0,20,39,70,90,100 16.7

ET Mer du Nord 0,20,39,70,90,100 16.7

ET Manche Atlantique 0,20,39,70,90,100 16.7

EC Méditerranée type 1 0,20,39,70,90,100 16.7

EC Méditerranée type 2A 0,20,39,70,90,100 16.7

EC Méditerranée type 3W 0,20,39,70,90,100 16.7

ET Méditerranée type delta 0,20,39,70,90,100 16.7

ET Méditerranée type lagune 0,20,39,70,90,100 16.7

EC Méditerranée type corse 0,20,39,70,90,100 16.7

EC Réunion 0,20,39,70,90,100 16.7

EC Martinique 0,20,39,70,90,100 16.7

ET Martinique 0,20,39,70,90,100 16.7

EC Guadeloupe 0,20,39,70,90,100 16.7

EC Mayotte 0,20,39,70,90,100 16.7

ET Mayotte type delta 0,20,39,70,90,100 16.7

EC Guyane 0,20,39,70,90,100 16.7

*N.B. : les grilles sont identiques pour tous les types de ME, c’est le seuil de bloom qui change (voir chapitre 4)*

Annexe 7 : Ajouts de traitements spécifiques aux lagunes (comprend les ajouts pour le phyto et pour les nutriments).

Document Antonin Gimard

## Dans Figures\_Evaluation\_BioPhyto\_Lagunes

### Nouveau nuage de point des valeurs

Pour chaque paramètre de chaque indicateur, un nuage de point est dessiné selon le modèle suivant :



Les seuils apparaissent ainsi en pointillés colorés, et les stations de mesure en points de nuances de gris différents. Ces graphiques sont rangés par lagunes dans le dossier « Figure Lagunes » du .OUT.

Ils ont pour objectif de permettre une visualisation des mesures directement issues des prélèvements, sans les transformations formant un indicateur. Ils permettent donc d’identifier les années extrêmes, et de donner un indice visuel de tendances.

## Dans Data\_Evaluation\_Bio\_Phyto et Data\_Evaluation\_PC\_GEN \_Nutriments

### Confiance et précision

La méthode déjà utilisée pour estimer la confiance et la précision des indices et des indicateurs est la méthode de ré-échantillonnage par bootstrap non paramétrique. Pour un indice donné, l’ensemble des mesures relevées forme un échantillon de taille n. Dans cet échantillon, un tirage aléatoire avec remise, stratifié par mois, de n réplicats est fait. Sur ce nouvel échantillon, ou certaines valeurs peuvent être tirées en doublons, le calcul de l’indice est réalisé. Cette procédure est répétée 1000 fois ce qui va donc nous donner 1000 valeurs possibles pour l’indice. La méthode du bootstrap permet d’obtenir un échantillon sur lequel peut être estimée la distribution de probabilité de l’indice. Les valeurs simulées de l’indicateur peuvent être calculées grâce aux valeurs obtenues pour les indices le composant, permettant alors d’estimer la distribution de probabilité de l’indicateur.

A partir de cette distribution, on calcule son intervalle de confiance à 95%, en utilisant les quantiles 0,025 et 0,975 ainsi que la probabilité d’appartenance à chacune des cinq classes d’état. Ces deux outils renseignent respectivement sur la précision et la confiance de l’indicateur. Si les données sont, de base, très différentes, les 1000 valeurs simulées de l’indice le seront: l’intervalle de confiance sera donc grand, et la probabilité d’appartenance à une seule classe sera faible. L’intervalle estimé par bootstrap traduit l’ensemble des variabilités s’appliquant sur l’indice étudié. La technique du bootstrap a été mise en place en 2010 sur le compartiment phytoplancton (Soudant et Belin, 2010), puis en 2014 sur les nutriments (Derolez et al., 2015) mais sans faire l’objet d’illustration et sans figurer dans les rapports de surveillance.

Afin de retracer l’évolution de ces intervalles de confiance au cours du temps et de décrire plus précisément la dispersion des données et de l’indicateur, de nouveaux sont créés. Pour représenter graphiquement la dispersion des indicateurs obtenus par bootstrap, une représentation en histogrammes de la fréquence d’apparition d’une valeur est produite. Visuellement la classe dominante ou médiane est ainsi identifiée et la représentation permet de visualiser l’homogénéité des données et son impact sur le diagnostic.

Suite à ce chapitre, sont intégrés au tableau de résultat :

* les intervalles de confiances,
* la répartition en classes,
* est produit le graphique suivant :



Ces graphiques sont rangés par lagunes dans le dossier « Figure Lagunes » du .OUT.

### Tendances.

Les tendances interannuelles ont été identifiées de deux façons :

* Le test de Mann Kendall (Mann, 1945; Kendall, 1975; Gilbert, 1987) est un test non paramétrique permet tant d’identifier si une tendance monotone, c’est -à-dire continument croissante ou décroissante, est présente au sein d’un jeu de donnée. Il s’agit d’un test sur les rangs, c'est-à-dire que les données (ici les données brutes) vont être comparées deux à deux, un signe positif étant attribué si la kième valeur est supérieure à la nième valeur, avec k>n. La somme de ces signes, ainsi que la variance associée sont utilisées pour produire la statistique du test. Le signe de la statistique Z indique le sens de la tendance des données. Cette valeur est ensuite comparée à la statistique de la distribution normale afin de déterminer la p-value associée au test. Il est considéré que la tendance déterminée par le test de Mann-Kendall n’est prise en compte que si sa p-value associée est inférieure à 0,05 ce qui signifie que la statistique Z Mann-Kendall est significativement différente de 0. Ce test a été réalisé et intégré aux scripts à l’aide du package «rkt». Sur le tableau récapitulant les résultats, cette tendance (augmentation, diminution ou pas de tendance) est indiquée, avec le coefficient de la courbe de régression.
* L’analyse de Kruskal-Wallis (Kruskal et Wallis, 1952) et les tests post-hoc de Dunn (Dunn, 1964). Cette analyse non paramétrique permettra de comparer les données des différentes années afin d’identifier celles pour lesquelles les valeurs des métriques sont significativement différentes des autres. Ce test a été réalisé et intégré aux scripts à l’aide du package «agricolae».Les résultats de l’intégration de cet outil statistique sont fournis sous la forme de lettres, représentant la catégorie de chaque année où une campagne de récolte de données a lieu. Si deux années possèdent une lettre différente, cela signifie que les valeurs de l’indicateur pour ces deux années sont significativement différentes.

Les packages «dplyr», «reshape2», «gdata» et «boot» ont été utilisés pour produire les différentes illustrations.

Suite à ce chapitre, sont intégrés dans le tableau de résultat

* La tendance de Mann Kendall
* Les codes en lettres des années (Test de Kruskal Wallis et tests post-hoc).

1. \\nantesdir\vigiestat\International\Rapports et etudes\Simulation DCE\Macros\points avec donnees DIREN Bretagne 1994-2007 en douteux.txt [↑](#footnote-ref-1)
2. Grandes cellules >20μm = taille 2 ; Petites cellules entre 2 et 20μm = taille 1 [↑](#footnote-ref-2)