
Plan de gestion des données du SNO PHYTOBS - Réseau d'observation du phytoplancton

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe : modèle structuré" fourni par Science Europe.

Renseignements sur le plan

Titre du plan Plan de gestion des données du SNO PHYTOBS - Réseau d'observation du phytoplancton

Version Version initiale

Objet/périmètre du plan Pour toute question sur le plan de gestion de données, veuillez contacter : data@ir-ilico.fr

Année de labellisation du réseau : 2018, puis relabellisé en 2020

Année d'intégration dans ILICO : 2016

Problématique scientifique :

Les questions scientifiques examinées dans le cadre de PHYTOBS concernent la structuration et les changements de la biodiversité phytoplanctonique. L'intérêt des questions d'écologie marine côtière centrées sur l'évolution à long terme des communautés et de certaines espèces a été souligné par des groupes de recherche internationaux sur le phytoplancton. L'étude de ces questions s'appuie sur d'importantes séries temporelles (certaines supérieures à 20 ans) en plusieurs sites littoraux du REPHY et/ou du SOMLIT. Ces sites sont soumis à des pressions environnementales et anthropiques contrastées. Les interactions entre recherche et observation s'inscrivent dans un cercle vertueux : les méthodes innovantes augmentent la quantité et la qualité des informations disponibles pour tester des hypothèses sur la structuration et le rôle fonctionnel de la biodiversité phytoplanctonique.

Étendue géographique d'action du réseau : Littoral métropolitain

Domaines de recherche (selon classification de l'OCDE) Earth and related environmental sciences, Biological sciences (Natural sciences)

Langue fra

Date de création 2022-09-02

Date de dernière modification 2023-05-04

Licence Creative Commons Attribution Share Alike 4.0 International

Documents (publications, rapports, brevets, plan expérimental....), sites web associés

- Site web du SNO PHYTOBS : <https://www.phytobs.fr/>
- DOI du jeu de données : <https://doi.org/10.17882/85178>

Renseignements sur le projet

Titre du projet SNO PHYTOBS : Réseau d'observation du phytoplancton

Acronyme SNO_PHYTOBS

Résumé **PHYTOBS** est un Service National d'Observation faisant partie de l'infrastructure de recherche littorale et côtière (**ILICO**). Son objectif principal est de constituer un réseau national de dispositifs d'observation du **microphytoplancton** dont les principales caractéristiques sont des fréquences d'échantillonnages élevées, l'identification et le dénombrement du microphytoplancton associé à d'autres paramètres environnementaux.

Sources de financement

- Ifremer :
- CNRS-INSU :
- Ministère de l'enseignement supérieur, de la recherche et de l'innovation - Mesri :

Date de début 1984-01-01

Partenaires

- Institut français de recherche pour l'exploitation de la mer ()
- Centre national de la recherche scientifique ()
- Université de Bordeaux ()
- Observatoire Océanologique de Banyuls sur Mer (201522729V)
- université de Bretagne-Occidentale ()
- Université de La Rochelle ()
- Université de Caen ()
- Institut Méditerranéen d'océanologie (201220349W)
- Station biologique de Roscoff (200119846X)
- Institut de la Mer de Villefranche (201622338R)
- université de Lille ()

Produits de recherche :

1. Données d'observation du SNO PHYTOBS (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
Cellule d'administration Quadrige	Ifremer	<ul style="list-style-type: none">• Responsable de la documentation des données
CLAQUIN Pascal - https://orcid.org/0000-0002-5183-203X	Université de Caen Normandie (UNICAEN)	<ul style="list-style-type: none">• Responsable de la qualité des données• Responsable du dépôt et de la diffusion des données• Responsable du traitement et de l'analyse des données

Nom	Affiliation	Rôles
HOEBEKE Mark - https://orcid.org/0000-0001-6311-9752	Station Biologique de Roscoff	<ul style="list-style-type: none"> • Responsable de la conservation à long terme des données • Responsable du stockage des données
IR ILICO (data@ir-ilico.fr)		<ul style="list-style-type: none"> • Responsable du plan de gestion de données
LEMOINE Maud - https://orcid.org/0000-0002-1322-5765	Ifremer	<ul style="list-style-type: none"> • Responsable de la documentation des données • Responsable de la production ou de la collecte des données • Responsable de la qualité des données • Responsable du dépôt et de la diffusion des données • Responsable du traitement et de l'analyse des données
LEMOINE Maud (Ifremer) et CLAQUIN Pascal (Université de Caen Normandie)	Ifremer - Université de Caen Normandie	<ul style="list-style-type: none"> • Coordinateur du projet • Personne contact pour les données
NEAUD-MASSON Nadine - https://orcid.org/0000-0002-6716-8102	Ifremer	<ul style="list-style-type: none"> • Responsable de la documentation des données • Responsable de la production ou de la collecte des données • Responsable de la qualité des données
RIGAUT-JALABERT Fabienne - https://orcid.org/0000-0002-5311-658X	Station Biologique de Roscoff	<ul style="list-style-type: none"> • Responsable de la documentation des données • Responsable de la production ou de la collecte des données • Responsable de la qualité des données

Nom	Affiliation	Rôles
SIMON Nathalie - https://orcid.org/0000-0001-6434-7861	Station Biologique de Roscoff	<ul style="list-style-type: none"> • Responsable de la documentation des données • Responsable de la production ou de la collecte des données • Responsable de la qualité des données

Plan de gestion des données du SNO PHYTOBS - Réseau d'observation du phytoplancton

1. Description des données et collecte ou réutilisation de données existantes

1.1 Description générale du produit de recherche

Nom	Données d'observation du SNO PHYTOBS
Description	<p>Nombre de jeux de données : Un jeu de données pour l'ensemble des sites et variables.</p> <p>Nombre de stations étudiées : 25 sites :</p> <ul style="list-style-type: none">• 15 sites REPHY : At-So, Auger, Barcarès, Boulogne, Bouzigues, Cabourg, Concarneau Large, Kervel Large, Le Cornard, Les Hébihens, Loguivy, Men er Roué, Ouest Loscolo, Parc Leucate 2, Teychan bis ;• 8 sites SOMLIT : Arcachon-Bouée 13, Banyuls - Sola, Brest - Sainte-Anne, La Rochelle - Antioche, Luc sur Mer, Marseille - Solemio, Roscoff SOMLIT-Astan, Wimereux SOMLIT-C ;• 2 sites REPHY/SOMLIT : Villefranche Point B, Sète. <p>Variables mesurées : Flore totale (nom du taxon identifié) ; Dénombrement de chaque taxon identifié</p>
Type	Jeu de données
Workpackage	Action 5 - Data sciences
Mots clés	<ul style="list-style-type: none">• /Activités humaines/Réseaux d'observation et de surveillance du littoral (Thèmes Sextant)• Caractéristiques géographiques océanographiques (GEMET - INSPIRE themes, version 1.0)• Dispositifs d'observation pour la recherche (Thématiques ODATIS)• CDS-IS-SBR (Centre de données ODATIS)• /Observations in-situ/Réseaux (Type de jeux de donnée ODATIS)• /Biologie marine/Zooplancton (Variables ODATIS)• Recherche (Thématiques - SIMM)• /Etat du Milieu/Littoral (Thématiques - SIMM)• /Etat du Milieu/Espèces (Thématiques - SIMM)• /Etat du Milieu/Biogéochimie (Thématiques - SIMM)• /Métropole/Manche mer du Nord (Sous-regions marines)• /Métropole/Golfe de Gascogne (Sous-regions marines)• /Métropole/Mers Celtiques (Sous-regions marines)• /Métropole/Méditerranée occidentale (Sous-regions marines)• /Métropole (Sous-regions marines)
Mots clés (texte libre)	PHYTOBS, IR-ILICO, Biogéochimie marine, Biologie marine, Climatologie, Hydrologie, Qualité des eaux et des écosystèmes
Langue	fra
Contient des données sensibles ?	Non

1.2 Est-ce que des données existantes seront réutilisées ?

Justification	Oui - Réutilisation des données du SOMLIT.
----------------------	--

1.3 Comment seront produites/collectées les nouvelles données ?

Description

Type de données (données collectées ou produites*) : Données produites

Nature des données : Données textuelles et numériques

Stabilité des données (données fixes, croissantes ou révisables**) : Données révisables

Méthode d'acquisition : Les données sont issues du traitement d'échantillons d'eau prélevés au travers d'un ensemble de sites de suivi.

Fréquence d'acquisition : Suivi bimensuelle

Collecte automatisée des données : Non

Délais moyens de récupération des données*** : Variable selon les sites, de bimensuelles à annuelle.

** Données collectées : données acquises et transmises par un instrument de mesure ; Données produites : données acquises manuellement ou résultant d'un processus d'analyse expérimental.*

*** Données fixes : qui ne subissent aucune modifications après leur collecte/production ; Données croissantes : de nouvelles données peuvent être ajoutées et/ou mises à jour sans que cela n'impacte les données d'origines ; Données révisables : de nouvelles données peuvent être ajoutées et/ou mises à jour et les données d'origines peuvent être modifiées et/ou supprimées.*

**** Délais entre la date de mesure et la date où les données sont récupérées pour ensuite être traitées/analysées par les laboratoires producteurs.*

Nature des données

Observation

Equipements, plateaux techniques utilisés

- Matériel d'analyse et de traitement : Microscope optique :

2. Documentation et qualité des données

2.1 Quelles métadonnées et quelle documentation (par exemple mode d'organisation des données) accompagneront les données ?

Description

- **Référentiel géographique (sites d'étude)** : chaque site d'études du SNO ont été documentés au standard ISO 19115 (associés à des tables d'attribut au standard ISO 19110)
 - Catalogue : Sextant
 - URL : <https://sextant.ifremer.fr/documentation/pages/ILICO.html#/metadata/905b35f5-46ac-4c74-ae6b-259f159d844e>
- **Autres métadonnées** : Pour un export en ODV, les métadonnées sont présentes dans l'en-tête du fichier. Ce n'est pas le cas pour un export en CSV.

Standards de métadonnées/données

- ISO 19115 : <https://rdamsc.bath.ac.uk/msc/m22>
- ISO 19110 : <https://standards.iso.org/iso/19110/>
- ODV : https://www.bodc.ac.uk/resources/delivery_formats/odv_format/
- Référentiel taxonomique WoRMS (World Register of Marine Species) : <https://doi.org/10.14284/170>
- Référentiel du SANDRE : <https://www.sandre.eaufrance.fr/>
- Standards de fichiers SANDRE (QELI) : <https://www.sandre.eaufrance.fr/urn.php?urn=urn:sandre:scenario:qeli:FRA:::ressource:1.2:::xsd>

Code langue des métadonnées

fra

Logiciel de documentation

GeoNetwork

2.2 Quelles seront les méthodes utilisées pour assurer la qualité scientifique des données ?

Description

Validation :

- **REPHY** : validation manuelle de manière différée par une expertise scientifique dans les 2 mois suivant le prélèvement ;
- **PELAGOS** : la validation a lieu lors de l'intégration des données dans la base. Le délai d'intégration des données dépend d'un laboratoire à un autre.

Qualification :

- **REPHY** : la qualification des données est pour le moment manuelle et réalisée par des experts. Une qualification automatique est en cours de mise en place ;
- **PELAGOS** : pas de qualification à proprement parler pour le moment.

Références associées

- Qualification des données d'hydrologie dans Quadrige : https://quadrige.ifremer.fr/support/content/download/124393/file/01_fiche_qualif_Hydrologie_20190206.pdf

3. Exigences légales et éthiques, code de conduite

4. Traitement et analyse des données

4.1 Comment et avec quels moyens seront traitées les données ?

Description

A. Traitement

A.1. Identification du phytoplancton

- **Automatisation** : Manuelle : Préparation des prélèvements d'eau avant identification du phytoplancton par un technicien analyste.
- **Méthodes/outils/protocoles utilisés** : Utermöhl - Méthode analytique au microscope
- **Temps moyen nécessaire** : N/A
- **Niveaux des données - pré-traitement** : Données brutes

A.2. Formatage des données

- **Automatisation** : Automatique
- **Méthodes/outils/protocoles utilisés** : Formatage des données en ODV pour permettre la fusion des données REPHY et PELAGOS sous la forme d'un unique jeu de données PHYTOBS.
- **Temps moyen nécessaire** : Traitement réalisé une fois par an
- **Niveaux des données - pré-traitement** : Données validées

B. Analyse

B.1. REPHY

- **Automatisation** : Partielle
- **Méthodes/outils/protocoles utilisés** : Analyse automatique de certaines données sous R (Aurige). L'analyse descriptive est réalisée avec Stip chart, Bubble chart et inférentielle via Seasonnal Mann-Kendall, Dynamic linear model.
- **Temps moyen nécessaire** : Variable en fonction des délais de traitement.
- **Niveaux des données - pré-analyse** : Données brutes et validées.
- **Niveaux des données - post-analyse** : Données analysées et interprétées.

B.2. PELAGOS

- **Automatisation** : N/A

- **Méthodes/outils/protocoles utilisés** : N/A
- **Temps moyen nécessaire** : N/A
- **Niveaux des données - pré-analyse** : N/A
- **Niveaux des données - post-analyse** : N/A

B.3. PHYTOBS

- **Automatisation** : Oui
- **Méthodes/outils/protocoles utilisés** : Les graphes sont générés automatiquement (voir <https://www.phytobs.fr/Stations#/map>) à partir de la base de données PHYTOBS pour chaque paramètre et espèce selon les filtres effectués.
- **Temps moyen nécessaire** : Quelques secondes.
- **Niveaux des données - pré-analyse** : Données validées
- **Niveaux des données - post-analyse** : N/A

C. Produits

C.1. REPHY

- Bulletins de surveillance ;
- Paramaps ;
- Synoptiques ;
- Évaluation DCE ;
- DLM.

C.2. PELAGOS

N/A

C.3. PHYTOBS

- Graphiques avec des filtres temporels et sur les espèces ainsi qu'une possibilité d'export sous format CSV ou image ;
- Fichiers avec les données par station téléchargeables aux formats CSV ou ODV à partir de <https://www.phytobs.fr/Donnees> ;
- Fichier avec un export total + la table des (groupes) de taxons labellisés déposés sur SEANOE ([10.17882/85178](https://doi.org/10.17882/85178)).

Références associées

- Méthode analytique au microscope : Utermöhl

5. Stockage et sauvegarde des données pendant le processus de recherche

5.1 Comment les données seront-elles stockées et sauvegardées tout au long du projet ?

Besoins de stockage

Est ce que les données sont stockées dans une base de données ?

Oui – Deux bases de données primaires (Quadrigé et PELAGOS) génèrent des exports vers la base de données PHYTOBS :

- **Quadrigé** : Le système d'information Quadrigé a été développé par l'Ifremer pour gérer les données d'observation et de surveillance du littoral : https://www.ifremer.fr/quadrigé2_support/. Cette base de données contient notamment les données du REPHY.
- **PELAGOS** : Base dans laquelle sont stockées les données des **CNRS-INSU/Universités**. La base de données PELAGOS résulte des travaux menés par le réseau des stations et des observations marines (RESOMAR), et regroupe l'ensemble des données de biodiversité des écosystèmes pélagiques côtiers (dont les comptages taxinomiques phytoplanctoniques du SOMLIT). Cette base a été développée avec la plateforme bio-informatique ABiMS, elle est disponible à l'adresse : <https://pelagos.sb-roscoff.fr>. L'accès à PELAGOS se fait après signature de la charte RESOMAR et envoi à contact.pelagos@sb-roscoff.fr. L'ouverture d'un compte se fait par le remplissage d'un formulaire en ligne (en anglais).
- **PHYTOBS** : La base PHYTOBS a été mise en place par la Station Biologique de Roscoff pour **centraliser les données issues des bases primaires** Quadrigé et PELAGOS afin d'alimenter le portail en ligne.

■ Organismes porteurs :

- **Quadrigé** : Ifremer ;
- **PELAGOS et PHYTOBS** : Station Biologique de Roscoff.

■ **Modèle de données utilisé (relationnel, entité-association, objet, hiérarchique et réseau) :** Schéma relationnel pour les trois bases de données.

■ **Système de gestion de base de données (SGBD) utilisé :**

- **Quadrigé** : Oracle SQL ;
- **PELAGOS et PHYTOBS** : PostgreSQL

■ **Historique** : Historiques accessibles via chacun des exports annuel/semestriel

Volume estimé des données 55

Unité Mo

Equipements, plateaux techniques

- Quadrigé (Oracle SQL) :
- PELAGOS (PostgreSQL) :
- PHYTOBS-P (PostgreSQL) :
- ABiMS : <https://cat.opidor.fr/index.php/ABiMS>

Mesures prises pour la sécurité des données

Protection des données via un accès physique au centre de donnée sécurisé par des accréditations, ainsi que par les Access Control List (ACL).

6. Partage des données et conservation à long terme

6.1 Comment les données seront-elles partagées ?

Modalités de partage

Principe de diffusion : Accès libre

Période d'embargo : Non

Mécanisme d'accès : Il est recommandé de télécharger les données par l'intermédiaire du DOI du jeu de données PHYTOBS : <https://doi.org/10.17882/85178>. Des informations relatives au téléchargement des données sont disponibles dans la page : <https://www.phytobs.fr/Donnees>.

Moyen(s) mis en œuvre pour garantir l'intégrité des données :

- **Quadriga** : les données sont accessibles à la modification uniquement par les saisisseurs et par les administrateurs de la base de données ;
- **PELAGOS** : les données sont accessibles à la modification uniquement par les fournisseurs de données et les administrateurs de la base de données ;
- **PHYTOBS-P** : les données ne sont pas accessibles à la modification.

Document nécessaire à l'intelligibilité des données (code, abréviation, historique des versions et des logiciels de lecture, etc.) : Les différentes informations relatives aux données livrées sont disponibles sur la page de téléchargement des données (<https://www.phytobs.fr/Donnees>). La page d'accueil du DOI fournit également des informations sur le jeu de données. Le format ODV contient aussi des informations.

Format de livraison des données : CSV, ODV

Condition d'utilisation : Creative Commons Attribution Share Alike 2.0 Generic (**URL** : <http://spdx.org/licenses/CC-BY-SA-2.0.json>)

Identifiant associé aux données (ARK, DOI, Handle, LSID, NCBI, PURL, URL, URN, etc) : Un identifiant de type DOI.

Temps moyen nécessaire à la mise à disposition des données (période maximale allant de la collecte/production des données brutes à la récupération des données par un utilisateur) : Objectif à moyen terme : 1 an.

Traçabilité des données : Suivi des données via le DOI.

Nombre moyen de publications associées aux données : 63 publications recensées en 2017.

Caractéristiques des fichiers déposés

- phytobs SNO full 2021-01.odv.txt (odv)
- phytobs SNO full 2021-01.csv (text/csv)

6.2 Comment les données seront-elles conservées à long terme ?

Justification

Procédure : Sauvegarde à l'aide d'un logiciel dédié et sur une infrastructure de stockage dédiée. Archivage au CeDoNA par réplication/synchronisation hebdomadaire (en cours).

Type de sauvegarde : Uniquement les données validées sont accessibles dans le jeu de données PHYTOBS.

Convention de nomenclature des fichiers de données : Un export annuel/semestriel de la base de données est réalisé.

Sur le site <https://data.phytobs.fr/files>, les fichiers téléchargeables sont organisés en dossiers correspondant aux sites d'échantillonnage. Ces dossiers sont nommés d'après leur MRGID. Chaque dossier contient 6 fichiers : 3 jeux de données (Analyst, Phytobs, Combine) chacun sous 2 formats (CSV, ODV).

Les fichiers sont nommés suivant la nomenclature normalisée : nom_du_site-type_de_données.extension

La partie "type_de_données" correspond à :

- "Analyst" : pour les comptages bruts ;
- "Phytobs" : pour les comptages des taxons labellisés ;
- "Combined" : pour l'agrégation des deux précédents.

La partie "extension" correspond à :

- ".csv" : pour des tableaux sans métadonnées ;
- "_odv.txt" : pour les fichiers conformes au standard (bio)ODV).

Supports : Stockage primaire de la base de données dans le centre de données et d'archivage IS-SBR, géré par la plateforme ABiMS, avec des sauvegardes sur disques durs et sur bande. Présence d'un archivage distant sur des serveurs opérés par le CeDoNA.

Localisation des supports :

- Quadriga (REPHY) : localisé à Ifremer Brest ;
- PELAGOS et PHYTOBS-P : localisés à la Station Biologique de Roscoff.

Moyens mis en œuvre pour faciliter la récupération des données archivées en cas d'accident :

Récupération possible des données via les sauvegardes de la base réalisée en continue sur deux sites distincts.

Volume estimé des données	55
Unité	Mo
Dispositions finales	Durée de stockage des données : permanente