

Mémoire de fin d'études d'ingénieur agronome
spécialisation Technologies de l'Information et de la Communication

Apports des SIG dans l'analyse des influences environnementales sur la répartition spatiale des peuplements de poissons

Application aux poissons démerso-benthiques
du lagon sud ouest de Nouvelle Calédonie

Laurent BLOCH

Maître de stage : Jocelyne FERRARIS

Avril – septembre 2001

Remerciements

Je tiens à remercier particulièrement Jocelyne FERRARIS pour m'avoir accueilli dans le laboratoire d'écologie marine de l'IRD de Nouméa et pour les conseils qu'elle m'a apportés tout au long de mon travail.

Mes remerciements s'adressent également à Soizic Pautret et à l'équipe de développeurs du SIG "Savane" au centre IRD de Bondy (Marc Souris, Bernard Lortic et Elisabeth Habert) pour les conseils et le soutien technique qu'ils m'ont apportés dans l'utilisation du logiciel "Savane".

Merci aux équipages du Darmad et de l'Alcyon pour m'avoir permis de participer aux campagnes d'échantillonnage à la palangre et de comptage visuel en plongée, de bien bons moments en mer !

Merci également à Sabrina Virly pour le temps qu'elle a bien voulu consacrer à la relecture et à la correction de ce mémoire.

Merci enfin à l'ensemble des stagiaires, thésards et Volontaires à l'Aide Technique de l'IRD qui ont rendu mon séjour bien agréable et qui m'ont permis de montrer l'étendue de mon talent au volley et au football.

| | |
|--|----|
| CHAPITRE 1 : CONTEXTES GEOGRAPHIQUE ET SCIENTIFIQUE..... | 1 |
| 1. la Nouvelle Calédonie et le cadre géographique de travail..... | 1 |
| 2. l'IRD et le centre de Nouméa | 1 |
| 3. Contexte scientifique | 2 |
| 4. Objectifs du stage | 5 |
| CHAPITRE 2 : MATERIEL ET METHODE..... | 6 |
| 1. Moyens disponibles | 6 |
| 1.1. le SIG "Savane" | 6 |
| 1.1.1. Présentation du logiciel | 6 |
| 1.1.2. Intégration d'objets géographiques dans une base Savane | 6 |
| 1.2. Données disponibles..... | 7 |
| 1.2.1. Données de la base "LagonSO" | 7 |
| 1.2.2. Données disponibles non intégrées..... | 8 |
| 1.3. l'équipe SIG | 8 |
| 2. Traitements, méthodes d'étude | 9 |
| 2.1. Démarche générale | 9 |
| 2.2. Intégration de données..... | 10 |
| 2.2.1. Création de nouvelles relations..... | 10 |
| 2.2.2. Nomenclature des relations | 10 |
| 2.3 Données exploitées..... | 12 |
| 2.3.1. Données ichtyologiques..... | 12 |
| 2.3.2. Autres données | 12 |
| Points de sondages bathymétriques effectués par le SHOM..... | 12 |
| Points d'échantillonnages sédimentologiques..... | 13 |
| 2.4. Pré-traitement des données..... | 13 |
| 2.4.1. Propositions de découpages en saisons | 14 |
| 2.4.2. Recherche du découpage présentant le plus d'influence sur les données ichtyologiques..... | 14 |
| 2.5. Méthode d'interpolation utilisée | 15 |
| 2.6. Identification et description des habitats | 16 |
| 2.6.1. Caractériser et proposer une délimitation de la surface couverte par une espèce..... | 16 |
| 2.6.2. Dégager au sein de son habitat les préférences d'une espèce : | 17 |
| 2.7. Identification d'indicateurs biologiques..... | 19 |
| 2.7.1. Analyses hors SIG | 20 |
| 2.7.2. Utilisation de l'outil SIG : | 20 |
| Méthode de l'indice de distance : | 20 |
| Méthode de recouvrement de zones : | 21 |
| CHAPITRE 3 : RESULTATS | 23 |
| 1. La base de données ichtyologiques..... | 23 |
| 1.1. les relations "HalP8487Lagonso" et "HalP0001Lagonso" | 23 |
| 1.2. les relations "EcoP93Aborevisu" et "EcoP95Aborevisu" | 23 |
| 2. données pré-traitées | 24 |
| 3. description de zones propices..... | 25 |
| 3.1. Caractériser et proposer une délimitation de la zone couverte par une espèce | 26 |

| | |
|--|----|
| 3.1.1. Sélection des palangres où le bec de cane a été repéré..... | 26 |
| 3.1.2. Recherche des limites de tolérance..... | 26 |
| 3.1.4. Intégration de toutes les informations : proposition d'une "zone favorable" | 27 |
| 3.2. Identifier les préférences d'une espèce | 28 |
| 3.2.1. Zonation de l'information ichthyologique : | 28 |
| 3.2.2. Description des préférences :..... | 29 |
| 4. Identification d'indicateurs biologiques de l'effet terrigène | 31 |
| 4.1.1. Analyses hors SIG : | 31 |
| 4.1.2. Utilisation de l'outil SIG : | 32 |
| CHAPITRE 4. DISCUSSION : | 34 |
| 1. Intérêts de l'approche SIG dans ce type de recherche | 34 |
| 2. Validité de la démarche | 34 |
| 2.1. Pré-traitement des données..... | 34 |
| 2.2. Utilisation des résultats d'interpolations..... | 35 |
| 2.3. Délimitation et description d'habitat..... | 36 |
| 2.4. Recherche d'indicateurs biologiques | 36 |
| 3. Validité des résultats | 37 |
| 3.1. Délimitation et description de l'habitat du bec de cane | 37 |
| 3.2. Identification d'indicateur biologiques | 38 |
| Conclusion | 39 |
| Bibliographie | 40 |
| Annexe 1 : Indices granulométriques | 2 |
| Annexe 2 : méthode de l'analyse de variance | 3 |
| Annexe 3 : test des méthodes de comparaison de cartes :..... | 4 |
| Annexe 4 : La "gratte" ou Ciguatera..... | 5 |
| Annexe 5 : Répartition spatiale du <i>Lethrinus nebulosus</i> | 6 |
| Annexe 6 : Détails des résultats de la recherche des limites de tolérance du <i>Lethrinus nebulosus</i> . | 7 |
| Annexe 7 : Détails des résultats de la recherche des préférences environnementales du <i>Lethrinus nebulosus</i> | 9 |
| Annexe 8 : Répartitions spatiales de "PUE classe trophique 1" et "distance à la côte" | 13 |

Introduction

Située dans le Pacifique sud ouest, à proximité du tropique du Capricorne, la Nouvelle Calédonie possède une immense barrière de corail. Cette barrière de corail de 1600km de long délimite de vastes lagons qui représentent une surface de près de 25 000 km². Les écosystèmes récifolagunaires qu'ils abritent sont réputés pour l'extraordinaire diversité des espèces rencontrées, mais présentent également un fonctionnement complexe, soumis aux influences de nombreuses composantes extérieures. Depuis plus de 15 ans, l'IRD s'intéresse aux structures et au fonctionnement du lagon sud ouest de Nouvelle Calédonie.

Les informations scientifiques concernant le lagon sud ouest sont aujourd'hui nombreuses. Elles regroupent des données sur les communautés de poissons échantillonnées (données ichtyologiques), mais aussi des données relatives à d'autres thématiques (sédimentologie, microbiologie, océanographie physique ...). Il devient nécessaire d'ordonner et de synthétiser l'ensemble de ces informations, c'est pourquoi un Système d'Information Géographique (SIG) est en cours de réalisation. L'utilisation d'un SIG présente en effet différents atouts : d'une part la visualisation et la prise en compte de la dimension spatiale dans l'exploitation des données, et d'autre part l'intégration et le croisement de données de sources diverses, et leur restitution sous formes de cartes facilement interprétables.

L'objectif de ce travail est d'exploiter les potentialités offertes par l'outil SIG dans le traitement de données scientifiques. Le travail effectué concerne l'étude des peuplements de poissons récifolagunaires du lagon sud ouest. Il comprend dans un premier temps la constitution de la base de données géographiques rassemblant les informations ichtyologiques, puis l'exploitation de cette base pour la délimitation et la description d'habitats et la recherche d'indicateurs biologiques.

CHAPITRE 1 : CONTEXTES GEOGRAPHIQUE ET SCIENTIFIQUE

1. la Nouvelle Calédonie et le cadre géographique de travail

Située à l'extrémité sud de l'arc mélanésien, à proximité du Tropique du Capricorne, la Nouvelle-Calédonie se trouve au cœur du Pacifique Sud, à 20 000 km de la France et à 1 500 km de l'est de la côte australienne.

Ce territoire d'Outre Mer est constitué de plusieurs îles :

- l'île principale, encore appelée "Grande Terre", s'étire selon un axe sud-est nord-ouest et mesure 400 km de long et 50 km de large. Elle est formée d'une chaîne de montagnes centrale, culminant à 1620 m, qui sépare une côte ouest caractérisée par de grandes plaines à savane de faible altitude et une côte est étroite, encaissée, à la végétation tropicale luxuriante.
- les archipels voisins avec les îles Loyauté, l'île des Pins, les îles Belep et quelques îlots épars.

Au total, la Nouvelle-Calédonie représente une superficie émergée d'environ 19 100 km².

Une barrière de corail de 1 600 km entoure la Grande Terre, l'île des Pins et les îles Belep. Elle délimite un vaste lagon de 24 000 km² qui abrite une très grande diversité d'espèces végétales et animales.

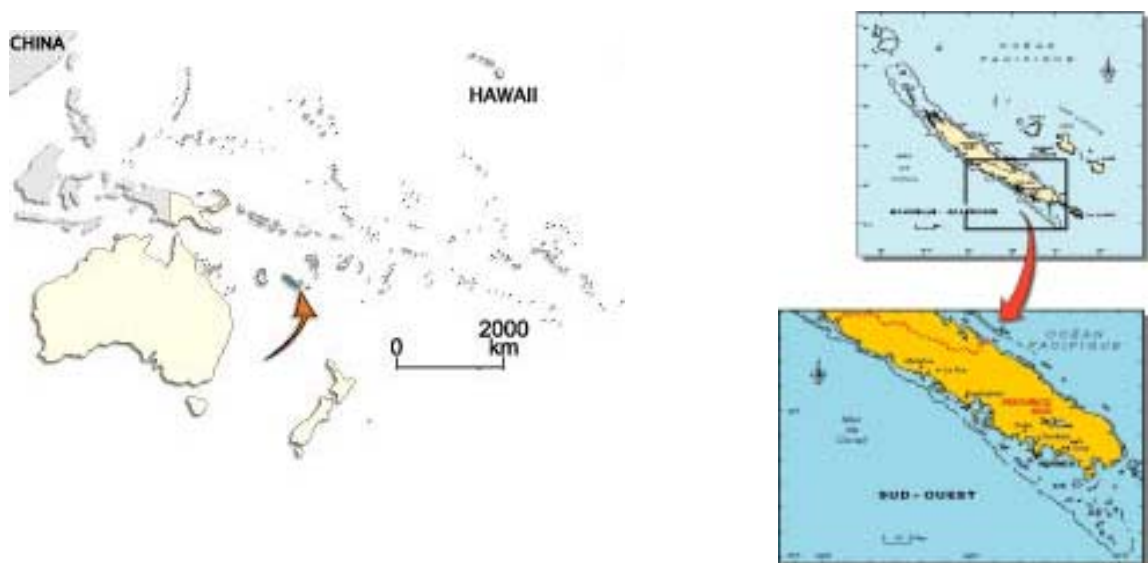


Figure 1 : localisation de la Nouvelle-Calédonie et du cadre de l'étude

La zone étudiée dans ce travail est la partie sud-ouest du lagon de Nouvelle-Calédonie, délimitée par la baie de Saint Vincent au nord et par l'extrémité du récif barrière au sud. Cette zone couvre une surface de 330 km² et abrite des biotopes très différents : mangroves, herbiers de phanérogames marines, champs d'algues, plaines de sable, dalles nues, récifs coralliens et îlots (LABOUTE *et al.*, 1991).

2. l'IRD et le centre de Nouméa

L'IRD (Institut de Recherche pour le Développement) est un établissement public à caractère scientifique et technologique (E.P.S.T.), sous la tutelle conjointe du Ministère de l'Enseignement Supérieur et de la Recherche et du Ministère de la Coopération.

Depuis plus de cinquante ans, il conduit des programmes de recherche sur les milieux intertropicaux : certains pays d'Afrique, d'Amérique Latine et d'Asie, ainsi que dans les Départements et Territoires d'Outre-Mer.

Organisme de recherche en coopération, l'IRD a pour tâche de travailler en partenariat avec les institutions des pays dans lesquels il intervient. Son action vise essentiellement la connaissance et la mise en valeur des milieux physiques, biologiques et humains caractéristiques de ces régions.

Le centre de Nouméa a été créé le 2 août 1946 sous l'appellation d'Institut Français d'Océanie (IFO) pour devenir centre ORSTOM en 1964 et centre IRD fin 1998.

S'il est pratiquement impossible de résumer plus de cinquante années de travaux scientifiques pluridisciplinaires, 3 grands domaines ont fait l'objet de recherches : le milieu marin, le milieu terrestre et les sociétés. Beaucoup de ces travaux ont été réalisés dans le cadre de grands programmes internationaux. Quelles soient fondamentales ou finalisées, les études effectuées ont conduit à des inventaires, collections, cartes, ouvrages, banques de données, brevets, contribuant à une meilleure connaissance et au développement de la Nouvelle-Calédonie et de la région Pacifique sud-ouest.

3. Contexte scientifique

Le lagon sud ouest de Nouvelle Calédonie fait l'objet de nombreux programmes de recherche. Les plus récents s'orientent vers la compréhension du fonctionnement de l'écosystème lagunaire et des effets sur celui-ci des influences terrigène et anthropiques.

- influence terrigène : apport par le ruissellement de particules et de solutés d'origine terrigène dans le milieu; cette influence est présentée par de nombreux auteurs comme une composante déterminante de l'organisation et la productivité des communautés de poissons
- influence anthropique : impact sur le milieu marin du développement des activités humaines :
 - l'urbanisation et ses rejets dans le milieu : rejet d'eaux usées riches en matière organique, rejets industriels et agricoles contenant des contaminants chimiques potentiellement toxiques;
 - l'urbanisation et le développement des activités de plaisance : fréquentation croissante des îlots et récifs du lagon sud ouest, essor de la pêche récréative;
 - le développement de la pêche professionnelle : la croissance démographique entraîne une augmentation de la demande en produits de la mer;
 - réglementations et mesures de protection du lagon : depuis 20 ans, pour des raisons touristique et de gestion de la ressource, de nombreuses mesures de protection ont été mises en place dans le lagon. De nombreux sites ont été soumis à des réglementations et certains ont vu leur niveau de protection évoluer dans le temps.

Les programmes de recherche mis en œuvre associent plusieurs domaines scientifiques. Les plus récents sont :

- **le programme ECOTROPE de l'IRD** (Ecosystèmes COtiers TROPicaux du Pacifique sous influences terrigènes et anthropiques) s'est déroulé entre 1996 et 2000. Son principal objectif était de déterminer comment l'homme influence les écosystèmes littoraux du Pacifique :
Ce programme a rassemblé des actions de recherche sur la compréhension du fonctionnement global de l'écosystème lagunaire (hydrodynamisme, sédimentologie, faune et flore vivant sur le fond (benthos), faune et flore vivant dans l'eau (pélagos).
- **le programme de l'Unité de Recherche CoRéUs de l'IRD** (Approche Ecosystémique des COmmunautés REcifales et de leurs USages dans le Pacifique Sud) a débuté en 2001 et s'achèvera en 2005. Son objectif général est "de proposer des méthodes et des outils d'acquisition et de représentation des connaissances pour améliorer la compréhension des

processus écologiques et des interactions majeures qui gouvernent les ressources marines insulaires". Le but est de généraliser les résultats à l'échelle du Pacifique sud insulaire et de proposer une méthodologie basée sur une approche écosystémique, transférable vers d'autres sites. Il s'agit plus particulièrement de :

- déterminer l'influence des facteurs environnementaux prédominants aux échelles locale et globale sur les communautés de poissons et d'invertébrés, ainsi que sur les caractéristiques biologiques des espèces et sur la diversité biologique, écologique et génétique des communautés récifo-lagonaires;
 - réaliser une première approche globale des interactions entre les ressources côtières insulaires, leurs milieux et leurs usages, plus spécifiquement la pêche;
 - proposer des méthodes standardisées d'observation et d'analyse des écosystèmes récifo-lagonaires et des outils d'aide à la gestion de leurs ressources (indicateurs).
- **le chantier Nouvelle Calédonie du "Programme National Environnement Côtier" (PNEC).** Le PNEC est un programme de recherche national orienté vers l'étude de la zone côtière. Il rassemble différents partenaires : IRD, IFREMER, CNRS, Institut National des Sciences de l'Univers (INSU), Bureau de Recherches Géologiques et Minières (BRGM), Elf, Centre National d'Etudes Spatiales (CNES). Les travaux sont conduits sous la forme de six chantiers correspondant à des sites géographiques:
- Baie de Seine
 - Golfe de Gascogne
 - Golfe du Lion
 - Lagunes méditerranéennes
 - Guyane
 - Nouvelle Calédonie

Le chantier Nouvelle Calédonie du PNEC a pour objectif de déterminer quels sont les principaux effets des activités humaines sur :

- le transport et la transformation des apports terrigènes et anthropiques
- les équilibres structurels et fonctionnels des organismes et des peuplements.

Il se structure en 5 thèmes :

- thème 1 : "circulation et transport des apports terrigènes et anthropiques"
- thème 2 : "fonctionnement biogéochimique du système lagonaire"
- thème 3 : "transfert des métaux dans la chaîne trophique"
- thème 4 : "variations de croissance des organismes lagonaires"
- thème 5 : "formalisation d'indicateurs des milieux et des ressources récifo-lagonaires dans une perspective écosystémique".

Etat des connaissances sur l'écologie des poissons :

De nombreux auteurs considèrent que la présence/absence, la diversité et l'abondance des espèces marines est la conséquence de l'état du milieu dans lequel elles se trouvent : une espèce est abondante dans un lieu si les caractéristiques de ce lieu lui sont favorables.

Les caractéristiques d'un milieu sont les caractéristiques physico-chimiques (température, profondeur, salinité ...). Celles ci sont sous l'influence de deux composantes extérieures :

- l'influence terrigène
- l'influence anthropique

Chacune de ces composantes modifie le milieu et rend un même endroit plus ou moins favorable aux différentes espèces marines :

- les caractéristiques d'une station d'échantillonnage et la communauté de poissons associée sont sous l'influence du substrat, des conditions physico-chimiques ...

- la station d'échantillonnage appartient à un biotope dont les caractéristiques sont sous l'influence de la position par rapport aux vents dominants, de l'influence terrigène, des usages (influence anthropique)...
- le biotope appartient à un écosystème dont les caractéristiques sont sous l'influence des importances relatives et des interactions des biotopes, de l'hydrodynamisme, de la topographie, des actions de pêche ...

L'observation en un point d'une communauté de poissons dépend de la combinaison de toutes ces influences.

Etat des connaissances sur les écosystèmes lagunaires :

Les écosystèmes lagunaires figurent parmi les systèmes écologiques les plus complexes et les plus fragiles de la planète. Ils servent d'abris à des milliers d'espèces végétales et animales.

Bien que des espèces nouvelles soient régulièrement découvertes et décrites, on peut estimer qu'il existe actuellement environ 3000 espèces de poissons récifaux dans le Pacifique tropical. La Nouvelle Calédonie se trouve à proximité de l'archipel indo-malais considéré par les biogéographes comme le centre de dissémination des espèces de la province indo-pacifique. De ce fait, la richesse spécifique y est plus élevée que dans les zones plus orientales du Pacifique. Avec environ 1200 espèces de poissons récifaux, dont 220 d'intérêt commercial, la Nouvelle Calédonie figure parmi les régions les plus riches du Pacifique.

Les écosystèmes lagunaires comportent trois biotopes principaux : les zones côtières, les récifs et la plaine lagunaire. Ces biotopes influencent la diversité et la productivité des peuplements qu'ils abritent. Ainsi, les récifs côtiers présentent en général des richesses spécifiques (nombre d'espèces), des densités (nombre de poissons par unité de surface) et des biomasses moins importantes que les récifs océaniques ou de milieu de lagon (Williams et Hatcher, 1982 ; Letourneur *et al.*, 1997, 1999).

Ces écosystèmes font l'objet de nombreux usages, notamment la pêche côtière particulièrement importante en milieu insulaire. La plupart des espèces ciblées par les pêcheries commerciales, de loisir ou de subsistance ne sont pas nécessairement les plus abondantes. Le nombre de familles présentant un intérêt commercial ou alimentaire est limité, mais le nombre d'espèces par famille est souvent important (Dalzell, 1996).

On peut opposer les espèces pêchées, mobiles, changeant de milieu en fonction de leur besoins et les autres espèces, sédentaires, donc inféodées à un milieu restreint.

L'étude de l'ensemble des espèces d'une communauté corallienne nécessite donc un effort d'échantillonnage considérable. Pourtant, la connaissance de l'ensemble de la communauté est indispensable si l'on désire comprendre les effets qu'ont une influence naturelle ou anthropique :

Ainsi, deux types d'échantillonnages ichtyologiques ont été réalisés dans le lagon sud ouest :

- des comptages visuels en plongée qui permettent le recensement et l'acquisition des caractéristiques biologiques (estimations de taille et de poids) des espèces rencontrées.
- des échantillonnages avec utilisation d'engin de pêche qui permettent d'estimer de stocks d'espèces d'intérêt commercial. La technique de pêche utilisée est la technique de la palangre :

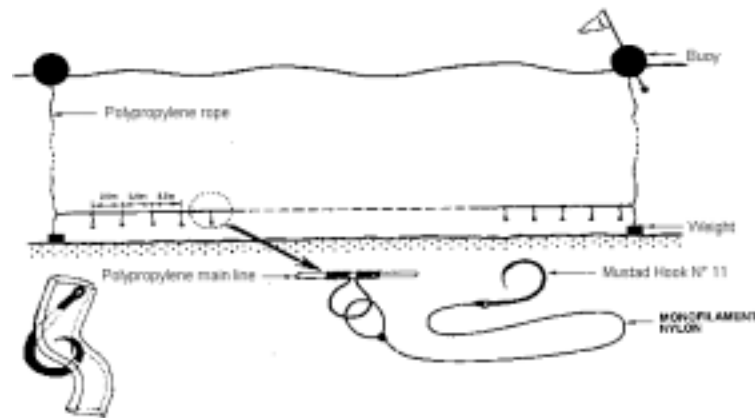


Figure n 2 : Schéma d'une palangre de fond "posée". source : VIRLY, 1997.

4. Objectifs du stage

Dans le contexte de recherche décrit précédemment, le stage s'inscrit principalement dans l'Unité de Recherche CoRéUs et dans le thème 5 du chantier Nouvelle Calédonie du PNEC. Ce thème a pour objectif de développer des démarches de construction d'indicateurs de l'état d'un écosystème côtier soumis aux influences terrigènes et anthropiques. L'étude est abordée sous l'angle de différents compartiments écologiques : faune et flore benthiques, ichtyofaune. Les grands points qui doivent être éclaircis sont :

- quels sont les espèces et les paramètres écologiques les plus sensibles aux gradients d'influences terrigènes et anthropiques, notamment à la pêche et aux mesures de protection?
- quels sont les habitats de ces espèces (caractéristiques physiques et biologiques) ?
- quels indicateurs peuvent être définis pour suivre l'état de ces habitats et des populations associées compte tenu des influences terrigènes et anthropiques ?

L'intérêt de ce projet est de proposer une approche synthétique de l'écosystème en intégrant dans un SIG des informations multi-sources. Cette approche permet d'étudier les éventuelles relations pouvant s'établir entre les différentes grandeurs mesurées en tenant compte de l'aspect spatial des données (structures et répartitions spatiales).

Au sein de ce projet, le stage correspond à l'utilisation de l'outil SIG pour l'étude du seul compartiment ichtyofaune. Les objectifs sont les suivants:

- 1- création de la base de données géographiques relative aux données d'ichtyofaune;
- 2- étude de la répartition et de la structuration spatiale des peuplements ichtyologiques du lagon sud ouest de Calédonie, mais aussi de leur évolution dans le temps par SIG. Applications sur certaines espèces.
- 3- identification, par SIG, des caractéristiques environnementales qui pourraient expliquer les répartitions spatiales des espèces considérées.
- 4- participer à la définition d'indicateurs écologiques : apports de l'outil SIG dans la recherche d'indicateurs écologiques des effets anthropiques et terrigènes. On entend par indicateur écologique une variable (espèce, famille, groupe trophique ...) ou une combinaison de variables suffisamment sensible à une influence pour pouvoir être utilisée pour évaluer l'importance de celle-ci.

CHAPITRE 2 : MATERIEL ET METHODE

1. Moyens disponibles

1.1. le SIG "Savane"

1.1.1. Présentation du logiciel

Savane est un SIG conçu et mis au point par une équipe de l'IRD, développé par Marc Souris. Il a été conçu pour fonctionner sous Unix et est en cours d'adaptation pour une utilisation sur PC. C'est pourquoi certaines fonctions ne sont pas encore disponibles ou incomplètes.

Dans "Savane", une couche géographique est appelée "relation", et les variables associées aux objets géographiques "attributs".

La particularité de ce logiciel est d'être conçu pour fonctionner sur une base de données partagée. Son architecture repose sur 4 modules :

- **Savateca** : module d'administration de la base. Il assure la gestion de l'organisation de la base de données. On peut ainsi déclarer, intégrer, supprimer et modifier les relations et les attributs qu'elles contiennent. Ce module permet également de gérer les accès des utilisateurs aux données. Ce module gère **la description** de la base, et non les données elles mêmes;
- **Savedit** : module de digitalisation. Il permet de convertir l'information géographique des données au format d'information géographique Savane (fichiers "mygale");
- **Savamer** : module de recalage et de géoréférencement d'image;
- **Savane** : module d'exploitation de la base, interface destinée aux utilisateurs. Il permet les opérations de sélections, croisements, mises en relation, regroupements, masques, jointures et agrégation de données géographiques, requêtes et calculs sur les attributs, calculs (statistiques, métriques, morphologiques), classifications, interpolations graphiques et modèles numériques. Un utilisateur ne disposant pas des autres modules ne peut effectuer que des modifications temporaires de la base de données.

1.1.2. Intégration d'objets géographiques dans une base Savane

On considère un fichier de données du type :

| Id | Longitude | Latitude | % sable | % vase | % graviers |
|----|-----------|----------|---------|--------|------------|
| 1 | 166.2324 | 22.5689 | 20 | 35 | 25 |
| 2 | 166.2357 | 22.5691 | 10 | 10 | 15 |
| 3 | 166.4251 | 22.5754 | 25 | 10 | 45 |
| 4 | 166.5421 | 22.5632 | 20 | 20 | 30 |

L'intégration de ces informations dans Savane s'effectue de la manière suivante :

- intégration d'un fichier de localisation comportant les informations : clé, longitude en degrés–minutes–secondes, latitude en degrés–minutes–secondes. Cette information est convertie au format d'information géographique de Savane (fichiers "Mygale").

Il faut définir, lors de l'intégration, la projection et l'ellipsoïde dans lesquelles ont été saisies les positions géographiques. La projection employée est toujours la même, il s'agit de la zone 58 de la projection Universal Transverse Mercator (UTM). En revanche, les ellipsoïdes sont variables : pour les positions acquises sur des cartes marines, il s'agit de positions rapportées au réseau

géodésique IGN 1972, avec pour ellipsoïde de référence l'International 1924; pour les positions acquises par GPS, il s'agit de l'ellipsoïde WGS 84.

Cette intégration s'effectue avec le module "Savedit".

- association du fichier de localisation "Mygale" et des variables descriptives des objets géographiques pour former une nouvelle relation. Cette étape s'effectue avec le module "Savateca".

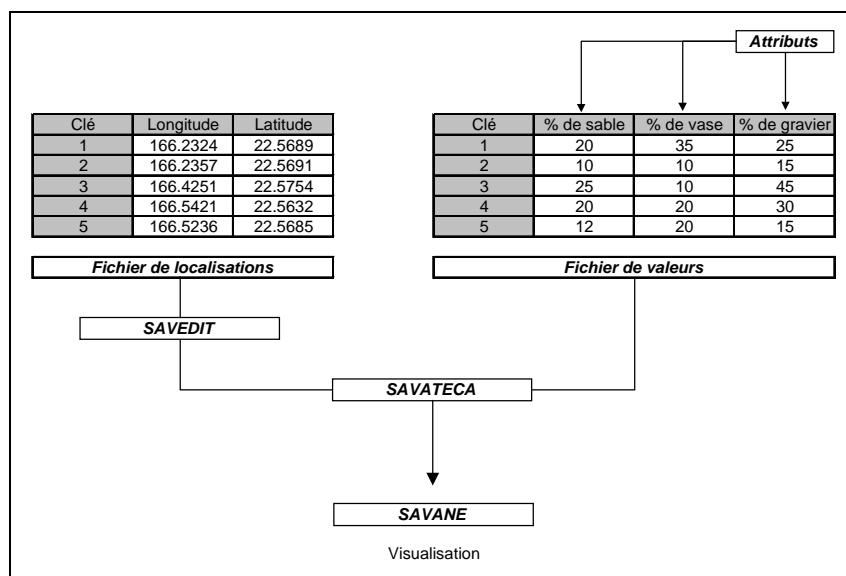


Figure n°3 : Principe de l'intégration de données dans SAVANE

L'intégration d'image se fait avec le module SAVAMER, qui permet également d'effectuer les opérations de géoréférencement nécessaires.

1.2. Données disponibles

1.2.1. Données de la base 'LagonSO'

Une base de données géographique "LagonSO", concernant le lagon sud-ouest de Nouméa, existe depuis Septembre 2000. Les informations qui s'y trouvent sont issues du programme ECOTROPE et des programmes précédents.

Cette base de données géographique contient les informations suivantes :

- informations topographiques (trait de côte du sud-ouest calédonien; contour de la barrière de corail, contour des îlots et des platiers du lagon sud ouest)
- relevés bathymétriques effectués par le Service Hydrographique et Océanique de la Marine (SHOM) dans et à l'extérieur du lagon sud ouest;
- données de la campagne d'échantillonnage à la palangre menée entre 1984 et 1987 (poissons);
- données de campagnes de comptages visuels en plongée réalisées sur le récif Aboré (substrat, poissons et organismes vivants);
- données de campagnes d'échantillonnages sédimentologiques réalisées sur tout le lagon sud-ouest.

Les objets géographiques présents dans cette base sont de plusieurs types :

- point : ce sont les stations d'échantillonnage
- ligne : par exemple le trait de côte extrait des cartes SHOM digitalisées;
- polygone : c'est le cas du récif barrière.
- mosaïque : il s'agit des données pixel (images géoréférencées ou résultats d'interpolation).

1.2.2. Données disponibles non intégrées

D'autres jeux de données sont disponibles et n'ont pas été intégrés dans la base "LagonSO". Il s'agit de :

1. données ichtyologiques IRD qui proviennent :

- de campagnes menées pour l'estimation des stocks d'espèces commerciales des fonds meubles : pêches expérimentales à la palangre réalisées sur le lagon sud-ouest en 2000-01
- de campagnes menées pour la connaissance des peuplements ichtyologiques récifaux : comptage visuels en plongée réalisés sur les récifs du lagon sud-ouest effectués entre 1985 et 1999.

2. données ichtyologiques de partenaires extérieurs :

- Thollot et Wantiez Consultants : suivi par comptages visuels en plongée de :
 - quelques espèces de poissons commerciaux et d'invertébrés sur 7 sites de la Province Sud;
 - peuplements rencontrés sur les îlots du lagon sud-ouest soumis à des mesures de protection (échantillonnages en 1990; 1994; 1998).
- Université de Nouvelle Calédonie :
 - comptages visuels en plongée réalisés dans le cadre de la thèse de S. Sarraména : "Contributions à l'étude des réserves marines du lagon sud-ouest de Nouvelle Calédonie".

3. autres données IRD :

Il s'agit de toutes les données géoréférencées provenant des actions de recherche menées dans le lagon sud-ouest par les différents laboratoires de l'IRD (océanographie physique, biogéochimie, ...).

4. autres données de partenaires extérieurs :

- Direction des Ressources Naturelles de la Province Sud : contours des zones du lagon sud ouest soumises à des mesures de protection.

1.3. l'équipe SIG

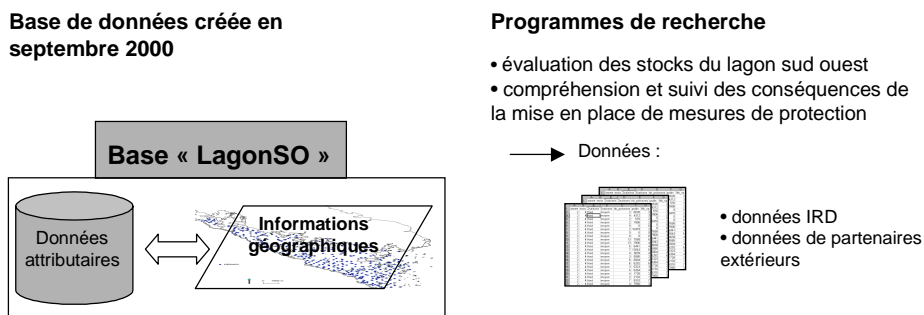
Plusieurs personnes collaborent au projet SIG :

- Jocelyne Ferraris, biostatisticienne à l'IRD et responsable du projet SIG;
- L'équipe de concepteurs de Savane au centre IRD de Bondy qui assurent une aide technique par messages électroniques;
- Les différents chercheurs thématiques de l'IRD, propriétaires des données qui sont et seront intégrées dans la base;
- Soizic Pautret, étudiante du Mastère SILAT, chargée de la mise en place et de l'administration de la base de données globale, mais aussi de la réalisation d'une utilisation du SIG sur les habitats benthiques;
- Alexis Raillard, Volontaire à l'Aide Technique informaticien. Travaille à temps partiel sur la future mise en réseau de la base de données partagée;
- Laurent Bloc'h, chargé de la constitution et de l'exploitation des données géographiques ichtyologiques.

2. Traitements, méthodes d'étude

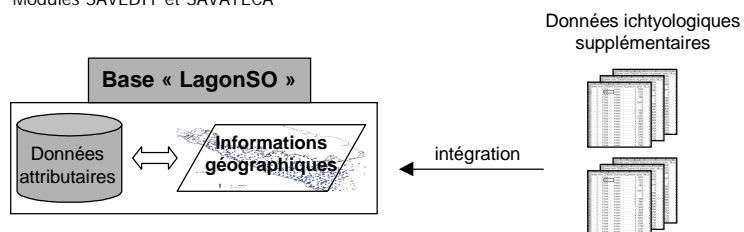
2.1. Démarche générale

La démarche générale de ce travail est représentée sur le schéma suivant :



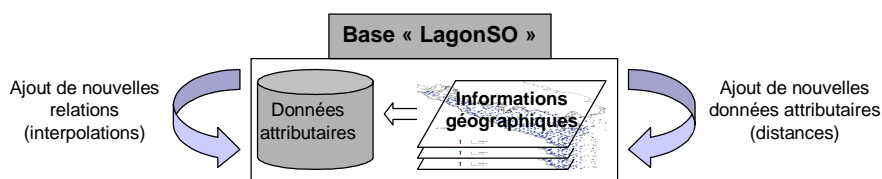
1. COMPLÉTER LA BASE DE DONNÉES ICHTYOLOGIQUES

→ Modules SAVEDIT et SAVATECA



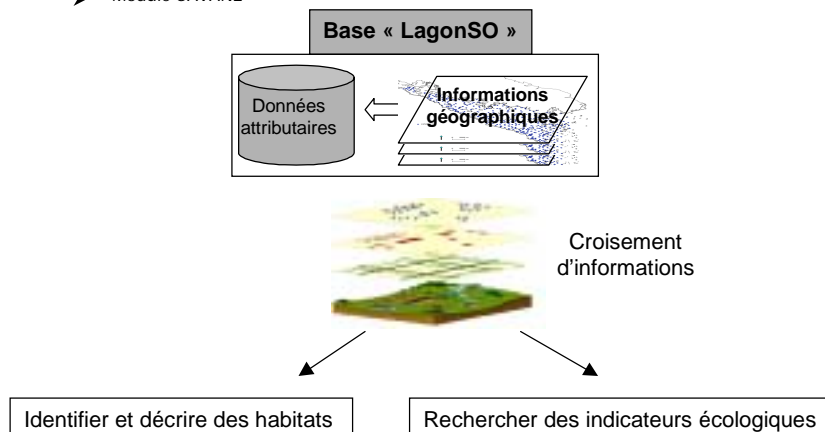
2. OPTIMISER LA BASE DE DONNÉES ICHTYOLOGIQUES

→ Modules SAVANE et SAVATECA



3. EXPLOITER LA BASE DE DONNÉES

→ Module SAVANE



2.2. Intégration de données

La constitution de la base de données ichthyologique a nécessité l'ajout de nouvelles données dans la base "LagonSO", ainsi que la mise en conformité de ces informations avec l'organisation globale de cette base de données. Cette mise en conformité a été conduite en collaboration avec Soizic Pautret, administrateur de la base "LagonSO".

2.2.1. Création de nouvelles relations

La base de données "LagonSO" ne contenait pas toutes les données ichthyologiques disponibles. Afin d'intégrer ces informations et de ne pas endommager la base existante; la base de données "Laurent" a été créée. Cette base contient :

- des relations copiées à partir de "LagonSO" :
 - trait de côte;
 - contours du récif barrière, des îlots, des platiers du lagon sud ouest;
 - points de relevés bathymétriques;
 - données des deux campagnes d'échantillonnages sédimentologiques menées sur le lagon sud ouest;
- des relations obtenues à partir des relations contenues dans "LagonSO" : l'information géographique (fichiers "Mygale") est conservée, mais de nouvelles données attributaires y sont associées :
 - données issues des deux campagnes de comptages visuels en plongée réalisés sur le récif Aboré;
 - données issues de la campagne d'échantillonnage à la palangre réalisées entre 1984 et 1987;
- des relations nouvelles :
 - données issues de la campagne d'échantillonnage à la palangre réalisée en 2000-01;
 - contour des zones du lagon sud ouest soumises à réglementations;
 - scène SPOT géoréférencée couvrant la zone récif Aboré-Nouméa;

Certaines informations présentées précédemment n'ont pas été intégrées dans cette base, pour différentes raisons (fichiers en cours d'harmonisation, perte de données ...).

La base "Laurent" a servi de base de travail à partir de laquelle les relations créées et validées ont été copiées vers la base "LagonSO".

2.2.2. Nomenclature des relations

La base de données géographiques en construction rassemble des données relatives à des domaines et des thématiques différents. Ayant pour but d'être partagée entre tous les chercheurs du centre IRD de Nouméa, la base nécessite la création d'une nomenclature homogène des relations qu'elle contient. "Savane" limite la taille des noms de relation à 16 caractères ; c'est pourquoi l'utilisation de codes est nécessaire.

Une première nomenclature a été proposée en septembre 2000 par G. Bouvet (Bouvet, 2000). Elle se basait sur la classification des fichiers de données en dossiers et sous-dossiers ayant chacun un code composé de trois lettres. Ces codes de classification étaient repris dans la nomenclature des relations obtenues à partir de ces fichiers.

Cette organisation a été reprise et améliorée :

Classification des fichiers de données :

Les dysfonctionnements observés au cours de l'année d'utilisation ont permis de redéfinir cette classification afin de la rendre plus fonctionnelle. De nombreux programmes de recherches associent plusieurs thématiques différentes. La classification de G. Bouvet conduisait à fractionner

par thématique les fichiers de données qui en étaient issus. Or ce fractionnement des fichiers de données est inutile, c'est pourquoi nous proposons une classification des données par programme de recherche. Ceci permet de conserver les jeux de données dans leur intégralité et de conserver la cohérence des programmes scientifiques dont ils sont issus.

L'arborescence des dossiers dans les répertoires de la base de données LagonSO s'insère donc désormais dans une logique de programmes scientifiques. La nomenclature des fichiers est la suivante : *"nom du programme_date_chiffre"*

date : deux derniers chiffres de l'année ou des années de début et de fin du programme

chiffre : numérotation utilisée s'il existe, pour un même programme, plusieurs jeux de données dans la même année.

Nomenclature des relations :

La nomenclature proposée reprend largement celle de G. Bouvet. Le nom d'une relation intégrée dans la base "LagonSO" comprend :

- trois caractères indiquant le thème de la relation
- un ou deux caractères indiquant le sous-thème de la relation
- deux ou quatre caractères indiquant la date des observations (deux derniers chiffres de l'année ou des années de début et de fin des observations)
- sept à dix caractères libres, destinés à l'ajout d'un autre thème dans le cas d'une relation pluri-thématique, l'ajout du nom du programme, des commentaires personnels, etc.

| code | thème | code | sous-thème |
|------------|-----------------|-----------|-----------------------------------|
| Fon | Fonds de cartes | V | Dessins vectorisés |
| | | T | Cartes topographiques |
| | | P | Photographies aériennes |
| | | I | Image satellite |
| | | C | Image satellite classifiée |
| Eau | Eau | H | Hydrodynamisme |
| | | R | hydrologie des rivières |
| | | P | physico-chimie de l'eau |
| Sed | Sédiments | B | biogéochimie et biosédimentologie |
| | | G | géochimie |
| | | S | sédimentologie |
| Eco | Ecologie | P | poissons |
| | | PB | phytobenthos |
| | | ZB | zoobenthos |
| | | PP | phytoplancton |
| | | ZP | zooplancton |
| | | B | bactéries |
| Bio | Biologie | | |
| Hal | Halieutique | P | palangre |

Figure n°5 : nomenclature des relations de S. Pautret et L. Bloc'h

La nomenclature prévoit des thèmes et sous-thème ne contenant pour l'instant aucune information (thème "biologie", sous thème "hydrologie des rivières"). Ils figurent pourtant dans la nomenclature de la base car celle-ci évoluera vers le recueil de toutes les informations scientifiques disponibles sur le lagon sud-ouest.

Cette nomenclature ne suffit pas pour renseigner totalement l'utilisateur sur le contenu d'une relation. Afin de compléter la description des relations et d'archiver les métadonnées (conditions de l'échantillonnage, technique de localisation, précision de la position, géodésie, ellipsoïde, projection, ...) un dictionnaire des données est élaboré par S. Pautret (Pautret, 2001).

2.3 Données exploitées

2.3.1. Données ichtyologiques

Il s'agit de deux relations de type "point" rassemblant les informations relatives aux deux campagnes d'échantillonnage à la palangre de 1984-87 et de 2000-01 dans le lagon sud ouest (363 couples de palangres en 1984-87, 350 en 2000-01).

Ces relations contiennent les informations :

- date
- profondeur
- nombre d'hameçons fixés sur la palangre
- poids total des poissons capturés par palangre
- nombre de poissons capturés
- nombre d'espèces capturées par palangre
- Prise par Unité d'Effort (PUE en g/hameçon) de chacune des 112 espèces pêchées.
exemple : pour une palangre de 100 hameçons ayant pêché 10 kg de *L.nebulosus*, la PUE correspondante est de $10\ 000 / 100 = 100$ g/hameçon
- distances à la côte, au récif barrière, aux platiers calculées à partir des positions géographiques des points.

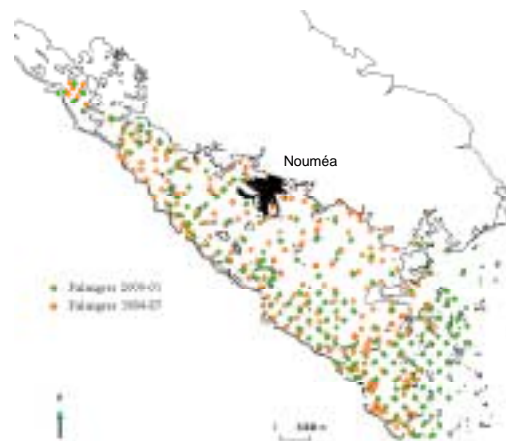


Figure n°5 : campagnes d'échantillonnages à la palangre réalisées dans le lagon sud ouest

2.3.2. Autres données

Les autres informations exploitées sont considérées comme des descripteurs de l'environnement. On distingue :

Points de sondages bathymétriques effectués par le SHOM

Il s'agit d'une relation de type "point" contenant les points de sondages bathymétriques effectués par le SHOM :

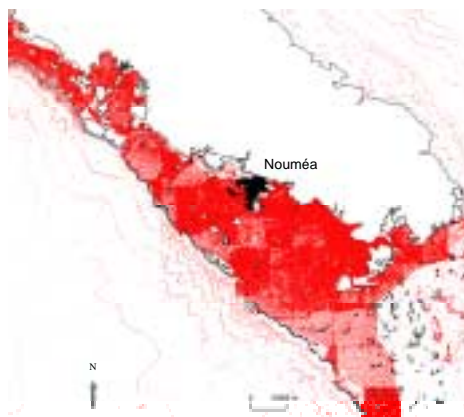


Figure n° 6 : points de sondages bathymétriques effectués par le SHOM

Points d'échantillonnages sédimentologiques

Il s'agit de deux relations de type "point" contenant les informations relatives aux points d'échantillonnages sédimentologiques relevés lors des campagnes de J.P. Debenay (1975) et de C. Chevillon (1986).

Ces relations contiennent les informations :

- profondeur
- fractions du sédiment retenues par les tamis : 20mm, 0.5mm, 0.25mm et 0.063mm (vases)
- indices granulométriques construits à partir de ces fractions (annexe 1)
- teneur en éléments carbonatés des vases (seulement mesurée par C. Chevillon).

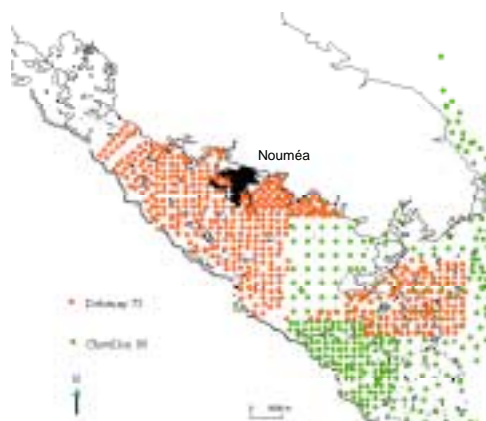


Figure n° 7 : points d'échantillonnages sédimentologiques sur le lagon sud ouest

2.4. Pré-traitement des données

L'observation en un point d'une communauté de poissons dépend de la combinaison des influences de nombreux paramètres. Or ces paramètres connaissent des variations spatio-temporelles dont il faut tenir compte pour l'exploitation des données :

comment associer et exploiter des données qui n'ont pas été recueillies dans les mêmes conditions (lieu et date) ?

Le rôle du SIG est de prendre en compte, dans l'exploitation des données, les variabilités spatiales. En revanche, "Savane" ne permet pas de prendre en compte les variabilités temporelles des différents paramètres.

L'objectif du pré-traitement des données est d'estimer l'importance des répercussions, observées sur les données ichthyologiques, des variabilités temporelles afin d'en tenir compte dans l'exploitation des données.

Par rapport aux objectifs du stage et aux données exploitées, seules les répercussions de la variabilité saisonnière des données ichthyologiques ont pu être estimées.

La méthode de pré-traitement proposée est basée sur un découpage de l'année en saisons obtenu à partir des variations de la température de l'eau.

2.4.1. Propositions de découpages en saisons

La température de l'eau connaît tout au long de l'année des variations continues qui lui font atteindre en moyenne 26 °C en été et 20 °C en hiver. Les poissons sont des animaux à température variable (hétérothermes), c'est à dire que leur température interne varie avec celle de l'eau qui les entoure. Celle ci influence grandement la biologie de ces espèces et détermine la vitesse de croissance, les phases de reproduction, etc.

Afin de mesurer simplement la variabilité induite par les variations de température de l'eau sur les données ichthyologiques, on propose de découper l'année en saisons de température:

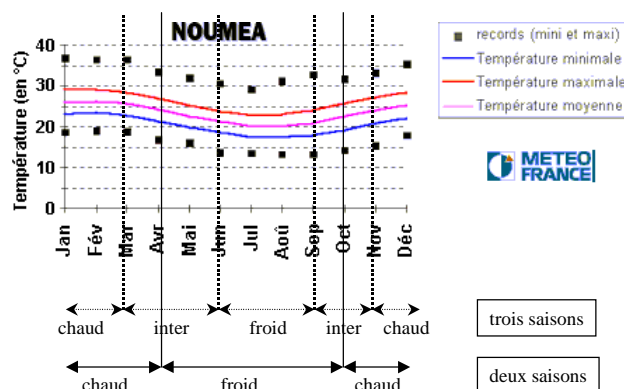


Figure n°8 : propositions de découpage en saisons de température de l'eau

Deux types de découpage sont possible :

- un découpage séparant une saison fraîche et une saison chaude, la température moyenne annuelle (23°C) étant utilisée comme seuil;
- un découpage définissant une saison chaude (températures comprises entre 24 et 26 °C), une saison fraîche (températures comprises entre 20 et 22 °C) et une saison intermédiaire (température comprises entre 22 et 24 °C).

2.4.2. Recherche du découpage présentant le plus d'influence sur les données ichthyologiques

Cette recherche s'appuie sur la méthode de **l'analyse de variance** (annexe 2):

On pose comme hypothèse que la variable ichthyologique y est fonction de la variable explicative V ="saison de température de l'eau". Nous proposons deux formes à cette variable:

- V_1 : deux saisons de température de l'eau;
- V_2 : trois saisons de température de l'eau.

On utilise donc pour chaque variable ichthyologique les modèles :

$$\text{pour tout } i \in [1; N], \quad y_i = \alpha_0 + \alpha_1(V_1)_i + \varepsilon_i$$

$$\text{pour tout } i \in [1; N], \quad y_i = \alpha_0 + \alpha_1(V_2)_i + \varepsilon_i$$

avec y_i = variable ichthyologique

V_1 : deux saisons de température de l'eau
 V_2 : trois saisons de température de l'eau

Si la probabilité p_1 issue de chacun des modèles (test de l'hypothèse $H_1 : \alpha_1 = 0$) est significative, elle est utilisée comme outil de comparaison pour choisir, entre les variables V_1 et V_2 , celle qui présente le plus d'influence sur la donnée ichtyologique (probabilité p_1 la plus faible).

Le modèle donne également les coefficients attribués à chacune des modalités de la variable explicative. Ces coefficients traduisent, pour chaque modalité, les écarts mesurés entre la moyenne observée pour la modalité et la moyenne globale.

Les coefficients correspondant aux modalités de la variable V retenue sont utilisés pour "corriger" les valeurs de la variable ichtyologique et tenter de réduire la variabilité due à V .

2.5. Méthode d'interpolation utilisée

La démarche de ce travail repose sur l'exploitation de l'aspect spatial des données. A partir de données ponctuelles, l'exploitation de l'aspect spatial n'est possible qu'après interpolation.

La méthode d'interpolation utilisée est celle proposée par "Savane". Il s'agit d'une interpolation barycentrique sur les voisins, à 2 passages :

- **barycentrique sur les voisins** : le plan sur lequel se situe le point à interpoler est découpé en huit octants (figure n° 9). La méthode de calcul consiste en une interpolation sur les huit points de référence les plus proches du point à interpoler, répartis dans les huit octants ; l'importance d'un point de référence étant d'autant plus grande dans le résultat du calcul que sa distance au point à interpoler est faible.
- **2 passages** : l'algorithme utilisé est très rapide si l'on dispose de beaucoup de points bien répartis dans la fenêtre de calcul. Dans le cas inverse, il est très lent. C'est pourquoi l'option "deux passages" est utilisée. Le premier passage permet la création de points valués régulièrement espacés qui serviront de points de référence lors du deuxième passage.

Cet algorithme est intéressant pour sa rapidité. En revanche, les valeurs obtenues sont le résultats de deux interpolations successives. Cela a pour conséquences d'amplifier les artefacts créés et d'obtenir un interpolateur non exact : il ne restitue pas exactement les valeurs mesurées sur les points d'échantillonnage.

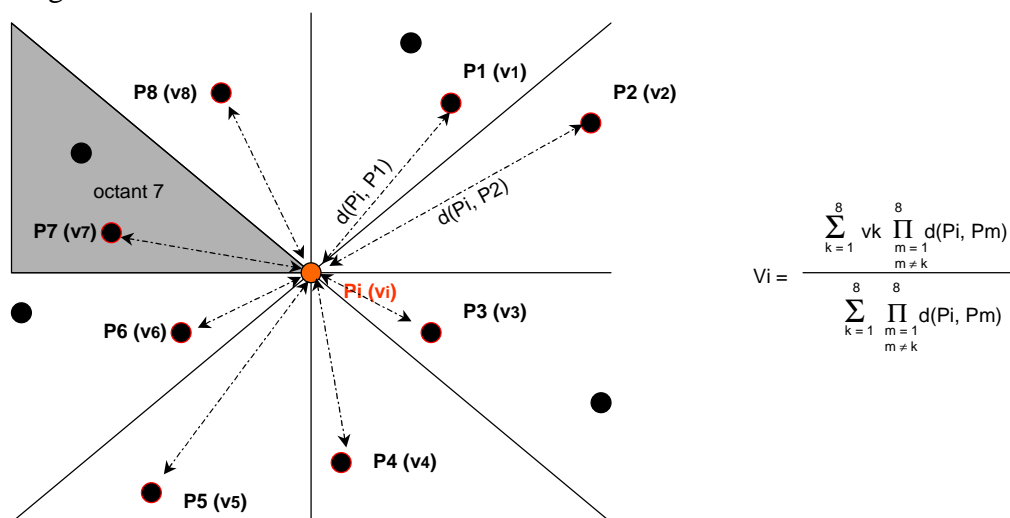


Figure n°9 : méthode d'interpolation barycentrique utilisée

La précision de l'interpolation (taille du pixel valué) d'une variable i dépend de l'abondance et de la répartition de ses points d'échantillonnage. Pour des variables présentant des variations continues et similaires sur l'ensemble de la zone d'étude, cette taille du pixel peut être définie à l'aide d'un variogramme. Dans notre cas, la réalisation de variogrammes n'est pas satisfaisante :

- certaines variables ne répondent pas aux critères d'utilisation du variogramme. Ainsi, la bathymétrie et toutes les variables biologiques présentent une hypervariabilité spatiale qui rendraient inexacts des résultats issus d'un variogramme couvrant toute la zone d'étude.
- l'effort d'échantillonnage pratiqué pour les autres variables n'est pas suffisant : les portées indiquées par les variogrammes (distance maximale séparant deux points pour lesquels les valeurs de la variable i sont liées) sont inférieures aux distances séparant les points d'échantillonnage.

Dans notre démarche, nous avons donc choisi de fixer la taille du pixel d'interpolation d'une variable en fonction de son effort d'échantillonnage : si la distance minimale séparant deux points d'échantillonnage d'une variable est de 200m, l'interpolation de cette variable s'effectue pour un pixel de 200m de côté.

2.6. Identification et description des habitats

Il s'agit d'une **première étape** vers l'identification et la description des habitats de différentes espèces de poissons lagunaires. C'est pourquoi le terme "zone favorable" sera utilisé par la suite.

L'habitat d'une espèce est une notion difficile à définir; d'après le Petit Robert, "un habitat est un milieu géographique propre à la vie d'une espèce animale ou végétale".

Dans ce travail, nous considérons l'habitat d'une espèce comme l'ensemble des milieux dans lesquels **cette espèce a été repérée** (par la pêche ou par comptage visuel). Les caractéristiques de ces milieux peuvent varier, c'est pourquoi l'habitat d'une espèce est une zone géographique d'abord décrite par la présence potentielle de cette espèce puis par les intervalles de valeurs des paramètres environnementaux associés.

La recherche et la description de l'habitat d'une espèce passe donc par :

- l'identification des limites de tolérance (min-max) de l'espèce pour les différents paramètres du milieu disponibles
- la délimitation de l'aire géographique correspondante, potentiellement couverte par l'espèce.

La méthode proposée pour la recherche de l'habitat d'une espèce par SIG comprend deux étapes :

2.6.1. Caractériser et proposer une délimitation de la surface couverte par une espèce

Cette première étape correspond à une recherche de seuils. Elle repose sur la **confrontation de données ponctuelles** : à chaque point d'échantillonnage ichtyologique est associée la valeur correspondante des variables environnementales disponibles.

Pour chacune de ces variables, les limites de tolérance de l'espèce sont identifiées. A partir de ces limites, on définit une "zone favorable" à cette espèce qui représente une première tentative de délimitation de la surface géographique couverte par cette espèce. Cette méthode comporte :

- une interpolation des données environnementales;
- la restriction des points d'échantillonnage ichtyologique aux seuls points pour lesquels l'espèce a été repérée (observée ou capturée);
- l'association, pour chaque point d'échantillonnage ichtyologique, des valeurs des variables environnementales correspondantes (issues des interpolations faites précédemment);

- une identification, pour chacune des variables environnementales disponibles, des limites de tolérance de l'espèce. On considère que ces limites correspondent, pour chaque variable environnementale, aux valeurs minimale et maximale relevées sur les points d'échantillonnage conservés;
- le seuillage des répartitions spatiales de chaque variable environnementale selon les limites de tolérance définies. On obtient, pour chaque variable environnementale, la zone délimitée par ces seuils;
- la délimitation d'une "zone favorable" à l'espèce par intersection des zones obtenues pour chaque variable environnementale.

Schéma :

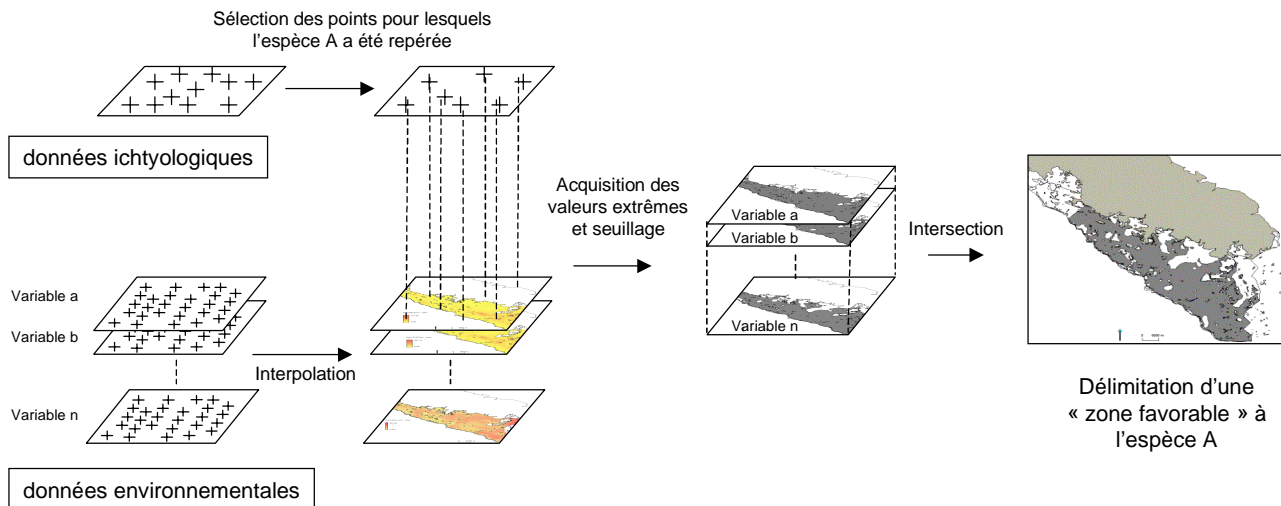


Figure n°10 : méthode de délimitation d'une zone favorable

Il est ensuite intéressant de préciser ce premier résultat, c'est à dire d'identifier, au sein de sa zone favorable, les préférences d'une espèce. Pour cela, on affine l'étude en exploitant les données d'abondance de l'espèce :

2.6.2. Dégager au sein de son habitat les préférences d'une espèce :

La méthode proposée repose sur une **zonation de l'information biologique** et sur la comparaison des caractéristiques environnementales des zones ainsi délimitées. Ces comparaisons s'effectuent sur les distributions des différentes variables environnementales disponibles.

Elle comprend :

1. une interpolation des données ichtyologiques et environnementales; une zonation de l'information ichtyologique (abondance). Cette zonation repose sur une classification des valeurs interpolées en 5 classes d'abondance. Afin de respecter la distribution de la variable, la classification s'effectue par écart à la moyenne, en fixant la taille des intervalles à un écart type. On aboutit à une partition en 5 classes : très forte, forte, moyenne, faible et très faible abondance. Représentées spatialement, ces classes conduisent à une zonation de l'information ichtyologique.

Remarque : afin de prendre en compte toutes les valeurs prises par la variable, la taille des intervalles des classes extrêmes n'est pas fixe.

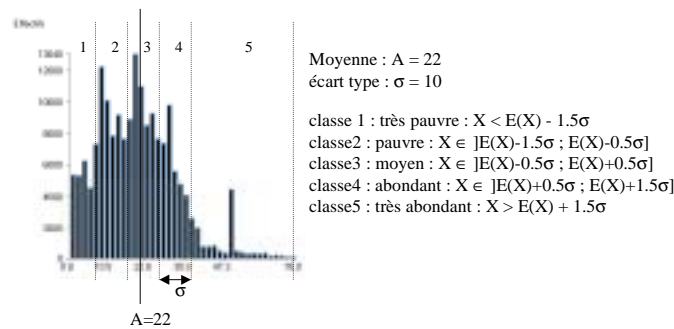


Figure n°11 : Exemple de classification

2. la recherche des préférences de l'espèce . Cette recherche s'appuie sur la comparaison, pour chaque variable environnementale, des distributions observées dans les différentes zones d'abondance.

Contraintes : les différentes zones d'abondance n'ont pas la même surface, donc les effectifs des classes d'abondance sont différents. La comparaison des distributions nécessite donc :

- la construction des distributions **des fréquences** des valeurs observées dans chaque zone d'abondance.
- La construction d'histogrammes suivant la règle de *Sturge* :
nombre de classes de fréquences = $1+(3.3 \log n)$ avec n l'effectif de la classe d'abondance

Ces comparaisons permettent, pour chaque variable, d'observer si une tendance se dégage :

- glissement progressif du mode;
- resserrement autour du mode des distributions des fréquences. Ce resserrement se mesure grâce à un indice d'aplatissement α_4 . Cet indice est égal au quotient du moment du quatrième ordre par la quatrième puissance de l'écart type :

$$\alpha_4 = \frac{m_4}{s^4} \quad \text{avec} \quad m_4 = \frac{\sum (x - E(x))^4}{n-1} \quad s^4 = \sigma^4$$

Cet indice est égal à 3 pour une distribution normale. Il est inférieur à trois si l'aplatissement de la distribution est supérieur à celui de la distribution normale et inversement.

Schéma :

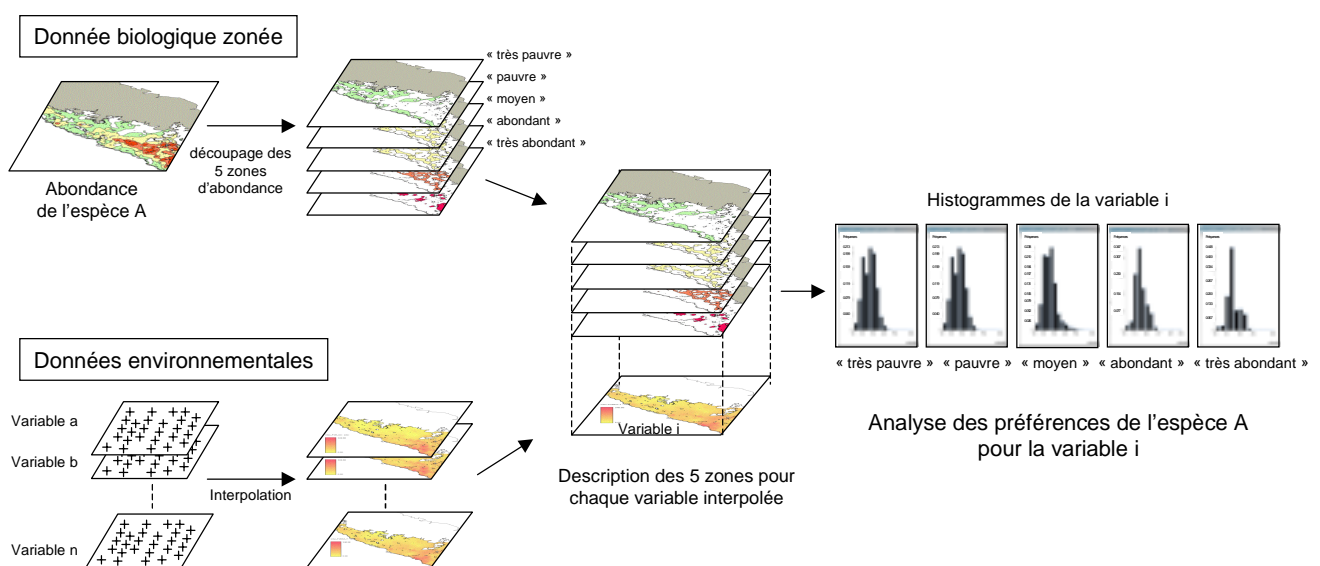


Figure n°12 : méthode d'identification des préférences d'une espèce

2.7. Identification d'indicateurs biologiques

Nous avons vu que la présence et l'abondance d'une espèce est la conséquence de l'état du milieu. Pour chaque paramètre environnemental, l'espèce présente des limites de tolérance : en dehors de ces limites, l'espèce n'est pas présente, et à l'intérieur de celles-ci, elle présente des préférences qui ont des conséquences sur sa biologie (abondance, taille, ...).

Dés lors, pourquoi ne pas utiliser les espèces animales comme des descripteurs de l'état du milieu ? Perez *et al* (2000) définissent un indicateur biologique comme : "Un taxon ou un ensemble de plusieurs taxons (peuplement, communauté) qui, par ses caractéristiques qualitatives et/ou quantitatives, témoigne de l'état d'un système écologique et qui, par des variations de ses caractéristiques, permet de détecter d'éventuelles modifications".

Notre démarche a pour objectif d'identifier les indicateurs biologiques des différentes influences qui s'exercent sur le lagon sud ouest. La connaissance de ces indicateurs pourra permettre dans un deuxième temps de définir l'indicateur biologique décrit dans la définition.

L'identification de l'indicateur biologique d'une influence suppose la recherche de la variable biologique **la plus sensible, donc la plus corrélée à cette influence.**

L'apport du SIG dans cette recherche est de permettre la prise en compte des corrélations spatiales.

On considérera par la suite l'utilisation des termes suivants :

- une corrélacion spatiale positive signifie que les répartitions spatiales de deux variables présentent des structures semblables : si la variable A augmente dans une zone, la variable B fait de même.
- une corrélacion spatiale négative signifie que les répartitions spatiales de deux variables présentent des structures opposées : si la variable A augmente dans une zone, la variable B diminue.

La méthode proposée ici est celle appliquée à la recherche d'un indicateur biologique de l'effet terrigène. Cet effet correspond aux conséquences de l'influence terrigène sur les communautés de poissons lagunaires. L'influence terrigène en un point est fonction de la quantité des apports terrigènes, solides et dissous, reçus en ce point. Dès lors, l'estimation de cette influence nécessite l'estimation de la quantité des apports. Deux approches sont proposées dans la littérature :

- l'utilisation de mesures sédimentologiques : les sédimentologues considèrent que la fraction vaseuse du sédiment (particules dont le diamètre est inférieur à 0.063 mm) ne présente que deux types de constituants :
 - les particules carbonatées résultant de l'activité biologique sur place (particules endogènes);
 - les autres particules, non carbonatées donc considérées comme exogènes et d'origine terrigène.

On dispose donc, par **la mesure de la teneur en carbonates des vases** en un point, d'une estimation des apports terrigènes en ce point (cette méthode ne permet de mesurer que les apports terrigènes solides).

- l'utilisation de **la distance à la côte** : pour de nombreux auteurs, cette distance est un bon estimateur de l'influence terrigène. Cette approche permet de considérer l'ensemble des apports (solides, dissous) d'origine terrigène.

La méthode proposée comprend deux étapes :

- identifier, entre la teneur en carbonates des vases et la distance à la côte, la variable qui quantifie le mieux une influence terrigène sur les données ichtyologiques utilisées
- d'identifier la variable biologique la plus liée à cette variable.

2.7.1. Analyses hors SIG

Ces analyses ont deux objectifs :

- identifier le meilleur estimateur de l'influence terrigène. Il s'agit d'identifier laquelle des variables "teneur en carbonates des vases" et "distance à la côte" présente le plus d'influence sur les espèces échantillonnées;
- effectuer une première sélection des variables biologiques les plus corrélées à l'estimateur retenu.

Ces analyses reposent sur une Analyse en Composante Principale normée (ACP). Les variables biologiques sont projetées sur un plan principal défini par les variables descriptives de l'environnement (distances à la côte, au récif, aux platiers, taille moyenne du sédiment, teneur en vases du sédiment, teneur en carbonates des vases).

2.7.2. Utilisation de l'outil SIG :

Le SIG permet l'exploitation de la répartition spatiale des variables. Cette répartition spatiale est utilisée comme un critère de sélection des variables. On calcule, pour les variables biologiques retenues à l'issue de l'ACP, le degré de proximité existant entre leur répartition spatiale et celle de l'estimateur de l'influence terrigène. Ces mesures de proximité des répartitions spatiales complètent les résultats de l'ACP pour permettre d'identifier la variable biologique la plus corrélée à l'influence terrigène.

Deux méthode de calcul des degrés de proximité sont proposées :

Méthode de l'indice de distance :

Cette méthode est basée sur le calcul d'un indice de distance existant entre deux répartitions spatiales. Cet indice correspond à la moyenne des valeurs absolues des écarts, mesurés pour chaque point d'une grille arbitraire, entre les valeurs centrées réduites des deux cartes.

Cette méthode comporte :

- l'acquisition, selon une grille de 150 000 points arbitraires, de valeurs interpolées de la carte A;
 - l'acquisition, selon la même grille de points arbitraires, de valeurs interpolées de la carte B;
 - la normalisation des données acquises sur la carte A. Cette normalisation ramène les écarts mesurés entre les valeurs à des distances à la moyenne exprimées en écarts type;
 - la normalisation des données acquises sur la carte B;
- Remarque** : La normalisation des données est nécessaire pour exploiter des cartes représentant des variables exprimées dans des unités différentes.
- le calcul, pour chaque point de la grille, de la valeur absolue de l'écart entre les valeurs normalisées des deux cartes;
 - le calcul de la moyenne des valeurs absolues des écarts.

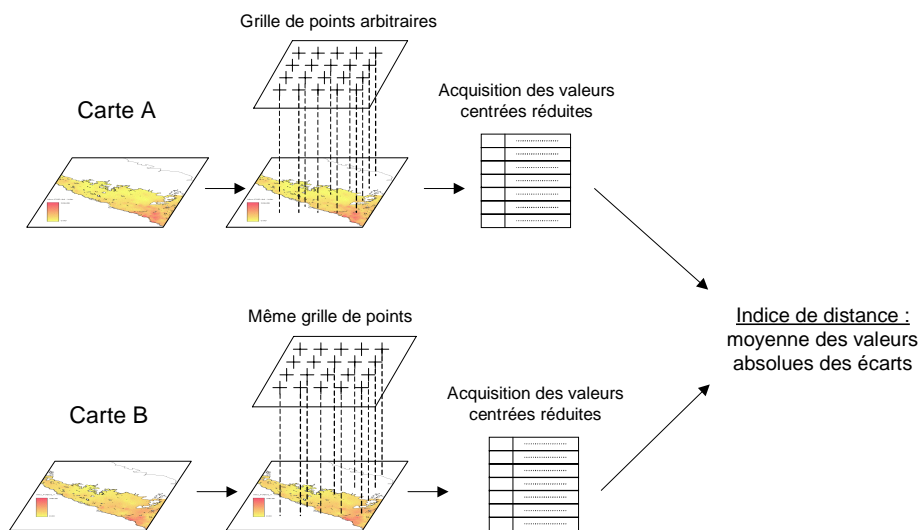
Schéma :

Figure n°13 : méthode de l'indice de distance

On obtient, pour chaque couple de cartes, un indice de distance positif représentant le niveau moyen des différences mesurées entre les deux cartes.

Cet indice est facilement interprétable : il est égal à 0 si les deux cartes sont identiques, et augmente à mesure que celles-ci diffèrent. (annexe 3)

Afin de compléter les informations apportées par l'indice de distance, une autre méthode, basée sur la mesure du taux de recouvrement de zones est proposée.

Méthode de recouvrement de zones :

Cette méthode repose sur l'exploitation de répartitions spatiales zonées. Chacune des répartitions spatiales à comparer subit une zonation obtenue par la méthode de classification détaillée précédemment (classification des valeurs par écart à la moyenne, avec fixation des intervalles à un écart type et obtention de cinq classes).

Cette méthode comporte :

- une classification des valeurs de la variable A. On obtient une typologie des valeurs en 5 classes (très pauvre, pauvre, moyen, abondant, très abondant) respectant la distribution intrinsèque de A. La représentation de ces classes donne la carte zonée A;
- une classification des valeurs de la variable B qui permet la réalisation de la carte zonée B;
- une comparaison des cartes zonées et la construction d'un tableau des surfaces des intersections des zones observées;
- un calcul de la valeur du χ^2 obtenue sur ce tableau observé à partir d'une hypothèse d'indépendance des deux variables.

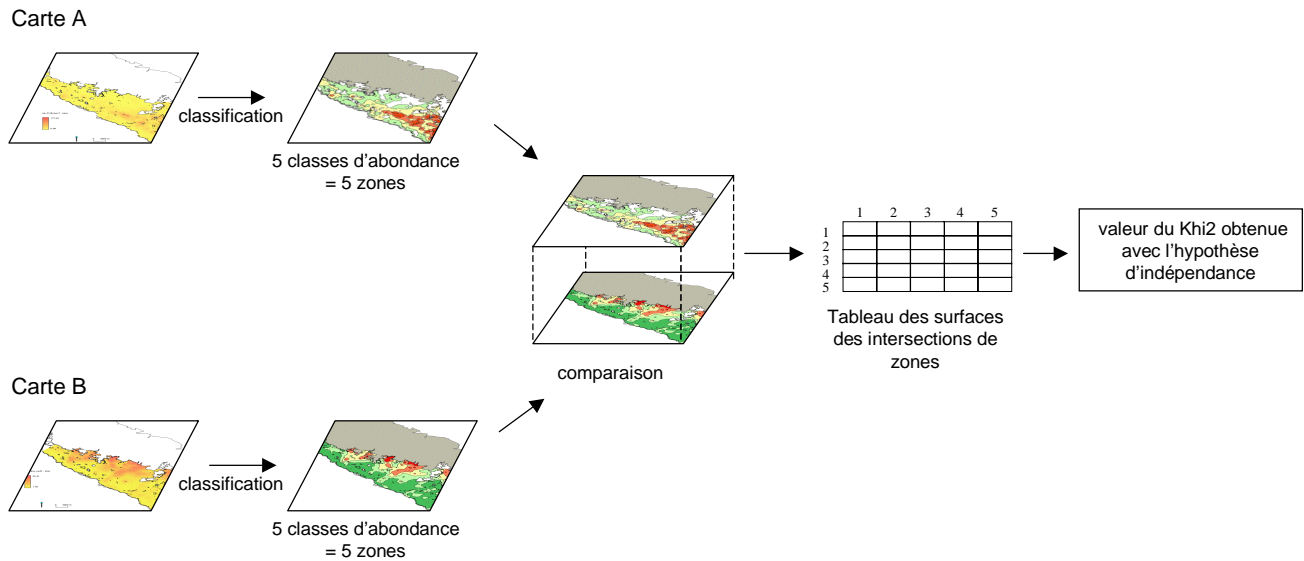
Schéma :

Figure n°14 : méthode de recouvrement de zones

Pour chaque couple de variables, on obtient une valeur du khi2 calculée à partir des surfaces des intersections des zones d'abondance : plus les répartitions spatiales des deux variables sont indépendantes, plus les surfaces des intersections observées sont proches des surfaces des intersections théoriques et plus la valeur du khi2 obtenue pour ce couple est faible (annexe 3).

La partition des valeurs de toutes les variables en 5 classes autorise la comparaison des valeurs du Khi2 de chaque couple de cartes.

👉 **On retiendra que** le SIG permet le traitement et l'exploitation de données spatialisées. Ses apports dans un contexte de recherche sont :

- utilisation de l'outil cartographique : représentation **visuelle** des phénomènes;
- croisement d'informations d'origines diverses (informations biologiques, sédimentologiques, bathymétriques);
- exploitation de la répartition spatiale des variables :
 - interpolation qui permet, à partir de données ponctuelles, de disposer d'une information sur l'ensemble de la zone d'étude;
 - délimitation de zones géographiques à partir de critères divers;
 - considération de la répartition spatiale comme une caractéristique des variables représentées, au même titre que la moyenne ou l'écart type.

CHAPITRE 3 : RESULTATS

Les résultats des méthodes mises en œuvre ne sont présentés que pour le Bec de cane (*Lethrinus nebulosus*).



Figure n°15 : photo de bec de cane (*Lethrinus nebulosus*).
Source : <http://www.wa.gov.au/westfish/rec/broc/ning/ningdet.html>

Le bec de cane est un poisson très abondant en Nouvelle Calédonie. La qualité de sa chair associée à un niveau de toxicité à la ciguatera (annexe 4) pratiquement nul font de cette espèce une des plus appréciées par les calédoniens.

Il mesure communément jusqu'à 60 cm, mais peut atteindre 86 cm et plus de 7 kg.

Les connaissances actuelles sur son écologie se résument par : le bec de cane fréquente les grandes étendues sédimentaires du lagon avec une préférence pour les fonds blancs de l'arrière du récif barrière où il trouve en abondance mollusques, crustacés et oursins dont il broie coquilles et carapaces ; il ne dédaigne pas les zones sableuses de la pente externe. Les jeunes se rencontrent fréquemment autour des récifs coralliens côtiers. Le *Lethrinus nebulosus* se situe dans une tranche de profondeur allant de 10 à 75 m. (LABOUTE, 2000)

1. La base de données ichtyologiques

La base "LagonSO" contient désormais quatre relations rassemblant des informations ichtyologiques :

1.1. les relations "HalP8487Lagonso" et "HalP0001Lagonso"

Ces relations contiennent les données des campagnes d'échantillonnages à la palangre de 1984-87 et 2000-01. On dispose, pour chacune d'elles et pour chaque palangre, des informations :

- date (mois, année)
- profondeur
- distances à la côte, au récif barrière et aux platiers
- nombre d'hameçons, pourcentage d'hameçons intacts, pourcentage d'hameçons perdus
- nombre de poissons pêchés, nombre d'espèces pêchées, poids total pêché
- PUE totale, PUE des 112 espèces pêchées à la palangre, PUE de ces espèces classées par familles et par groupes trophiques.

1.2. les relations "EcoP93Aborevisu" et "EcoP95Aborevisu"

Ces relations contiennent les données des campagnes de comptages visuels en plongée réalisées en 1993 et 1995 sur le récif Aboré. On dispose, pour chacune d'elles et pour chaque comptage, des informations :

- date (mois, année)
- profondeur (profondeurs minimale, maximale et moyenne)
- biotope (platier, tombant récifal, lagon)

- zone réserve (une moitié du récif Aboré était interdite à la pêche en 1993, l'ensemble du récif en 1995)
- densités (g/m^2) des espèces observées regroupées par familles, catégories trophiques et classes de mobilité (territoriaux, sédentaires, mobiles, très mobiles)

2. données pré-traitées

On rappelle que ce pré-traitement n'a été appliqué que pour mesurer et tenter de minimiser au sein des données ichtyologiques la variabilité due aux variations saisonnières de la température de l'eau.

Ce pré-traitement s'est révélé nécessaire après la visualisation de ce type de représentation :

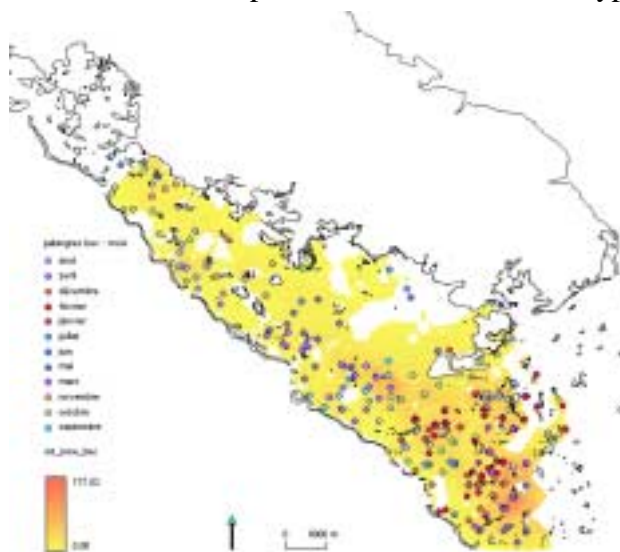


Figure n°16 : représentation de l'interpolation des PUE de bec de cane et des mois au cours desquels ont été posées les palangres correspondantes

L'effort d'échantillonnage des deux campagnes de pêche à la palangre sur le lagon sud ouest s'est réparti sur les douze mois des années 1984-87 et 2000-01. Cet étalement des échantillonnages induit une variabilité temporelle dans les données observées : on constate par exemple que les zones de plus grande abondance de bec de cane correspondent à la partie sud du lagon, principalement échantillonnée durant les mois de décembre, janvier et février (été austral).

Cette constatation pose les questions suivantes:

- comment exploiter des données qui n'ont pas été acquises aux mêmes dates ?
- quelle est l'influence de la saison sur les données échantillonnées et comment la quantifier ?

Les analyses de variance réalisées sur l'ensemble des données de PUE de bec de cane acquises lors des campagnes de palangres 1984-87 et 2000-01 ont donné les résultats suivants :

modèle 1 : pour tout $i \in [1; N]$, $\text{PUE}_{\text{bec de cane}} = \alpha_0 + \alpha_1(V_{\text{deux saisons}})_i + \epsilon_i$

| | Df | Sum of Sq | Mean Sq | F Value | Pr(F) |
|-------------|-----|-----------|----------|----------|--------------------|
| deuxsaisons | 1 | 9806.9 | 9806.932 | 7.663951 | 0.005784892 |
| Residuals | 689 | 881657.0 | 1279.618 | | |

modèle 2 : pour tout $i \in [1; N]$, $\text{PUE}_{\text{bec de cane}} = \alpha_0 + \alpha_1(V_{\text{trois saisons}})_i + \epsilon_i$

| | Df | Sum of Sq | Mean Sq | F Value | Pr(F) |
|--------------|-----|-----------|----------|----------|---------------------|
| troissaisons | 2 | 21134.0 | 10567.02 | 8.353282 | 0.0002603525 |
| Residuals | 688 | 870329.8 | 1265.01 | | |

Les modèles montrent que les influences des variables V_1 et V_2 sont significatives ($P < 0.05$). Puisque $P(V_2) < P(V_1)$, la variable V_2 présente plus d'influence sur la PUE du bec de cane que la variable V_1 . C'est donc la variable V_2 (découpage de l'année en trois saisons) qui doit être retenue pour la suite de la démarche.

Les coefficients issus du modèle 2 sont :

Stroissaisons:

| froid | moyen | chaud |
|-----------|-----------|----------|
| -8.546421 | 0.5219305 | 8.024491 |

Ces coefficients permettent d'estimer les effets de la variabilité saisonnière de la température de l'eau sur la PUE de bec de cane : ils signifient que la moyenne des PUE de bec de cane en "saison chaude" est supérieure de 8.02 g/hameçon à la moyenne globale des PUE de bec de cane.

La connaissance de ces coefficients permet de réduire grossièrement la variabilité due aux variations saisonnières de température de l'eau sur les données de PUE de bec de cane : il suffit d'appliquer les "corrections" suivantes :

- ajouter 8.55 g/hameçon aux PUE non nulles acquises en saison froide;
- retrancher 0.52 g/hameçon aux PUE acquises à l'intersaison;
- retrancher 8.02 g/hameçon aux PUE acquises à la saison chaude.

A l'issue de ces corrections, les PUE négatives sont ramenées à la valeur 0.

On dispose alors de données de PUE pour lesquelles la variabilité saisonnière de température de l'eau a été réduite, leur représentation spatiale donne :

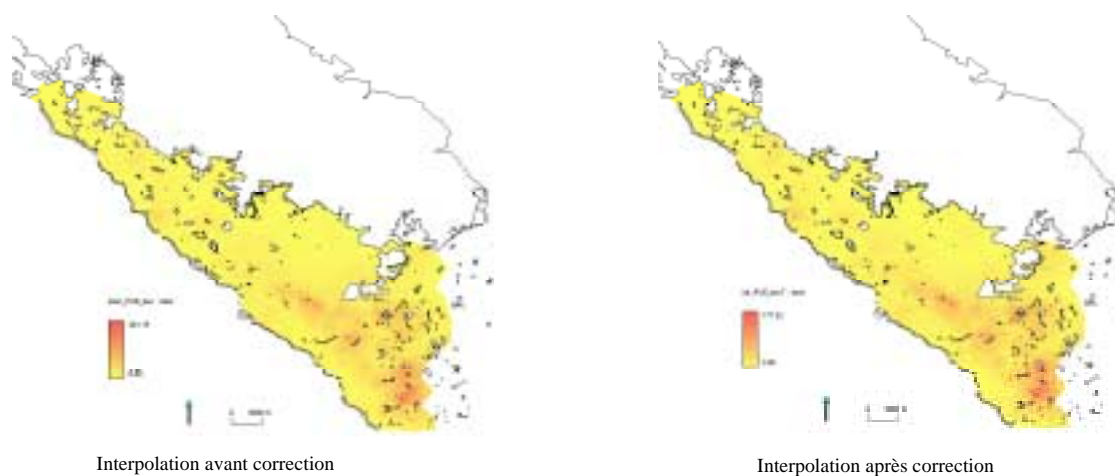


Figure n°17 : représentations spatiales des PUE de bec de cane avant et après "correction" de la variabilité saisonnière de température de l'eau.

Il apparaît à l'issue de cette méthode que l'influence des variations saisonnières de température de l'eau ne modifie pas la répartition spatiale de la variable "PUE du bec de cane", ce qui constitue déjà un résultat. La mise en œuvre de cette méthode nous permet surtout de considérer que les données de PUE de bec de cane ont été acquises en une même saison, et donc de rendre possible leur comparaison.

3. description de zones propices

On connaît le type de milieux favorables au bec de cane, mais la description plus précise de ses exigences et la délimitation de son aire de couverture ne sont pas connues. L'utilisation des SIG permet d'apporter les éléments de réponse suivants :

3.1. Caractériser et proposer une délimitation de la zone couverte par une espèce

Cette méthode repose sur l'acquisition, au niveau des points d'échantillonnage ichthyologique, des valeurs des descripteurs environnementaux correspondantes.

3.1.1. Sélection des palangres où le bec de cane a été repéré

Afin de connaître les valeurs minimale et maximale associées à la présence de bec de cane pour chaque variable environnementale, il est nécessaire de restreindre les points d'échantillonnage aux seuls points où le bec de cane a été repéré :

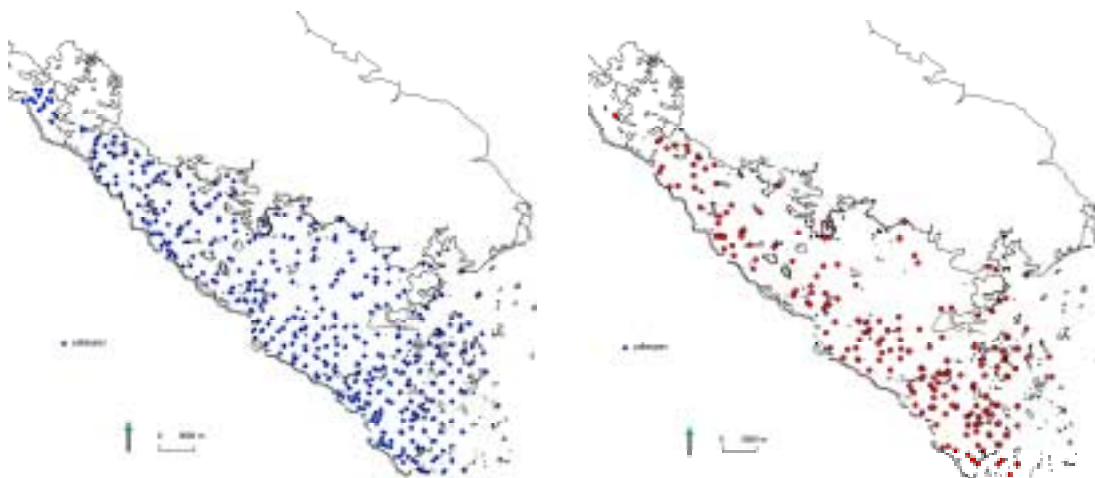


Figure n°18 : restriction aux seuls échantillonnages pour lesquels le bec de cane a été repéré

3.1.2. Recherche des limites de tolérance

A l'issue de l'acquisition des valeurs des variables environnementales, on dispose, pour tous les points d'échantillonnage à la palangre sélectionnés, des valeurs des variables "profondeur", "taille moyenne", "% vases", "% carbonates". On peut alors obtenir, pour chaque variable environnementale, les gammes de valeurs associées à la présence du bec de cane.

Ainsi, pour le pourcentage de vase, on obtient l'histogramme suivant:

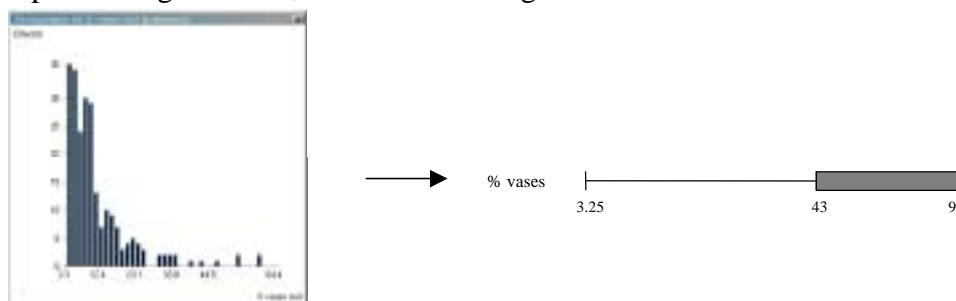


Figure n°19 : histogramme des teneurs en vases du sédiment pour les palangres ayant pêché du bec de cane et seuil associé

Cet histogramme, éclairé par les valeurs de PUE correspondantes, permet de définir un seuil au delà duquel le milieu semble défavorable au bec de cane.

On connaît la répartition spatiale de la teneur en vases du sédiment. A partir du seuil défini précédemment, on peut déterminer les zones du lagon sud ouest pour lesquelles la teneur en vase du sédiment rend le milieu favorable au bec de cane :



figure n°20 : carte des teneurs en vase du sédiment favorables au bec de cane

3.1.4. Intégration de toutes les informations : proposition d'une "zone favorable"

En rassemblant tous les seuils obtenus à l'issue de l'étape précédente (annexe 5), on obtient les limites suivantes :

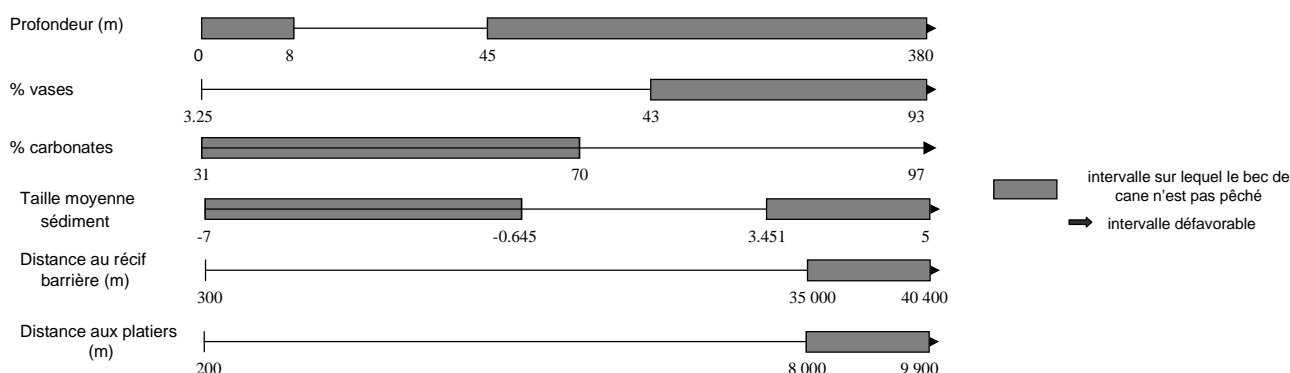


Figure n°21 : limites de tolérance du bec de cane

Chacune des variables présentée ci-dessus présente sa propre répartition spatiale. La connaissance des limites de tolérance du bec de cane pour la variable 1 permet de seuiller la répartition spatiale de cette variable et donc de représenter pour cette variable la zone des valeurs favorables au bec de cane.

En seuillant les répartitions spatiales de toutes les variables environnementales disponibles et en les recoupant, on peut définir une zone qui satisfait toutes les conditions définies précédemment. On obtient une carte de la "zone favorable" au bec de cane.

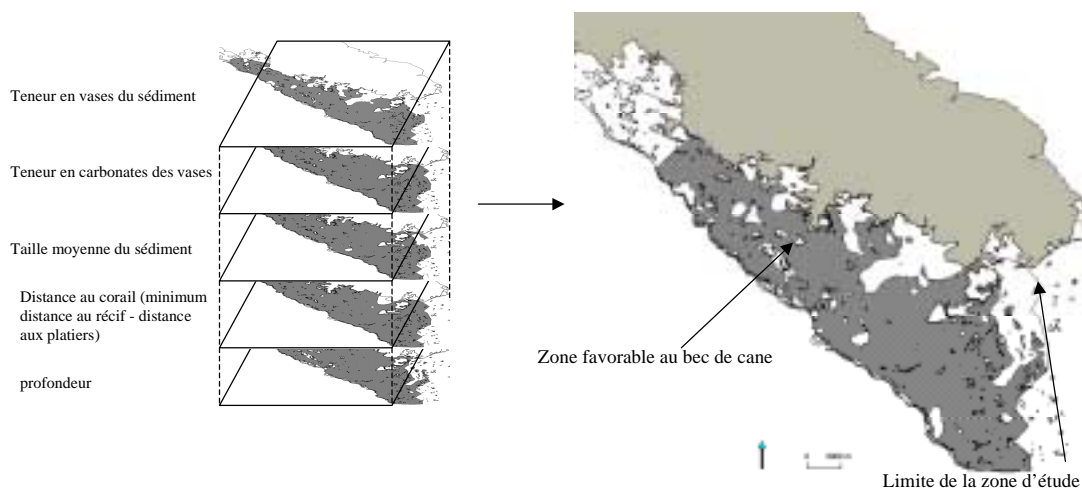


Figure n°22 : construction de la zone favorable au bec de cane

En considérant l'hypothèse : ["pêche de bec de cane" signifie "présence de l'espèce" donc "milieu favorable"], la zone ainsi définie correspond à une **estimation de la délimitation de la zone potentiellement couverte par le bec de cane**. Le niveau de précision de cette estimation dépend du nombre de variables environnementales mises en œuvre. On peut considérer cette estimation comme une première approche de délimitation de l'habitat du bec de cane.

On observe que la zone mise en évidence couvre pratiquement toute la zone d'étude. Le bec de cane est en effet connu pour être une espèce ubiquiste.

Néanmoins cette espèce ubiquiste présente nécessairement des préférences, des conditions du milieu optimales qui conditionnent la répartition de son abondance au sein de cette zone favorable.

3.2. Identifier les préférences d'une espèce

Cette méthode a pour objectif d'identifier les préférences du bec de cane au sein de sa zone favorable, c'est pourquoi elle ne concerne que les variables environnementales interpolées sur la zone délimitée auparavant. Ces informations se trouvent donc sous la forme de pixels valués de 134 m de coté.

Les préférences sont identifiées grâce aux informations données par les PUE observées dans le lagon sud ouest.

3.2.1. Zonation de l'information ichthyologique :

Les PUE de bec de cane obtenues par interpolation sur la zone favorable subissent une partition par écart à la moyenne en 5 classes.



Figure n°23 : histogramme et valeurs limites des classes de PUE de bec de cane

Cette classification permet de proposer une typologie de 4 classes d'abondance de bec de cane :

- "pauvre" : PUE comprise entre 0 et 11.51 g/hameçon;
- "moyen" : PUE comprise entre 11.51 et 31.33 g/hameçon;
- "abondant" : PUE comprise entre 31.33 et 51.14 g/hameçon;
- "très abondant" : PUE supérieure à 51.14 g/hameçon.

Cette typologie est utilisée pour zoner l'information ichthyologique, on obtient la représentation suivante :

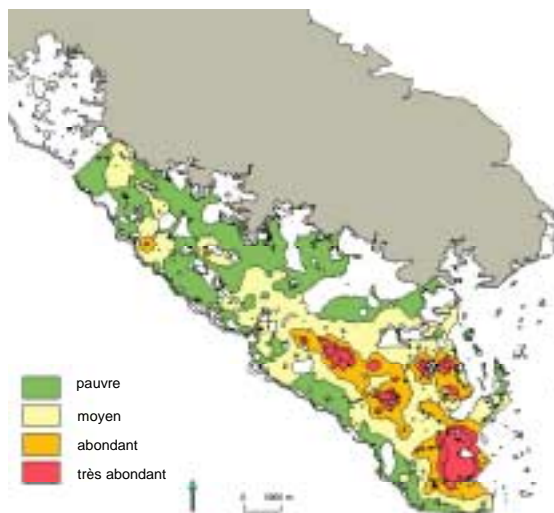


Figure n°24 : zonation des PUE de bec de cane selon la typologie proposée

3.2.2. Description des préférences :

La description des préférences du bec de cane repose sur la comparaison, pour chaque variable environnementale, des distributions des valeurs des pixels présents dans chacune des zones d'abondance.

Ces comparaisons permettent, pour chaque variable, :

- d'observer si une tendance se dégage : glissement progressif du mode ou resserrement autour du mode des distributions des fréquences ;
- d'identifier éventuellement l'intervalle des valeurs optimales pour le bec de cane.

Ainsi, pour la distance au récif, on exploite les histogrammes suivants :

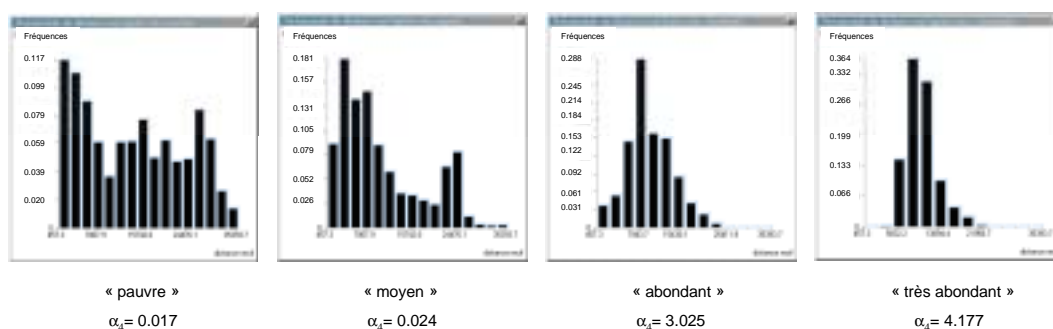


Figure n°25 : histogrammes des fréquences des distances au récif mesurées dans chaque zone d'abondance

La comparaison des distributions des distances au récif mesurées dans les différentes zones d'abondance vérifie ce que l'on observe sur la représentation visuelle des zones d'abondance du

bec de cane (Figure n°23) : dans le sud du lagon, les abondances augmentent à mesure que l'on s'éloigne du récif barrière. En effet, on observe sur les distributions :

- un glissement des modes qui traduit un gradient croissant des abondances sur l'axe récif-côte : situé à 600m du récif pour la zone "pauvre", il se situe à 11000m pour la zone "très abondant";
- un resserrement progressif des distributions autour de ces modes qui indique l'apparition d'un intervalle de distances au récif optimales:

$$\alpha_4 \text{ pauvre} < \alpha_4 \text{ moyen} < \alpha_4 \text{ abondant} < \alpha_4 \text{ très abondant}$$

Il apparaît donc que le bec de cane, dans sa zone favorable correspondant à des distances au récif comprises entre 300 et 35000m , soit plus abondant dans les zones situées entre 5000 et 25000m du récif barrière, avec un optimum observé à 11000m.

On constate qu'au delà de 25000m du récif, l'abondance du bec de cane diminue : on observe des modes secondaires à 30000m pour la zone "pauvre" et à 26000m pour la zone "moyen".

En appliquant cette analyse à toutes les variables environnementales disponibles, on identifie les préférences du bec du cane au sein de la zone favorable que nous avons délimitée précédemment:

- profondeur : aucune préférence n'est apparue
- teneur en vases du fond : entre 3.5 et 35 %, avec un optimum observé à 7 %
- teneur en carbonates des vases : autour de 88 %
- taille moyenne du sédiment : entre -0.5 et 2.2
- distance au récif : entre 5000 et 25000m
- distance à la côte : entre 10 000 et 48 000m, avec un optimum observé à 19 000m.

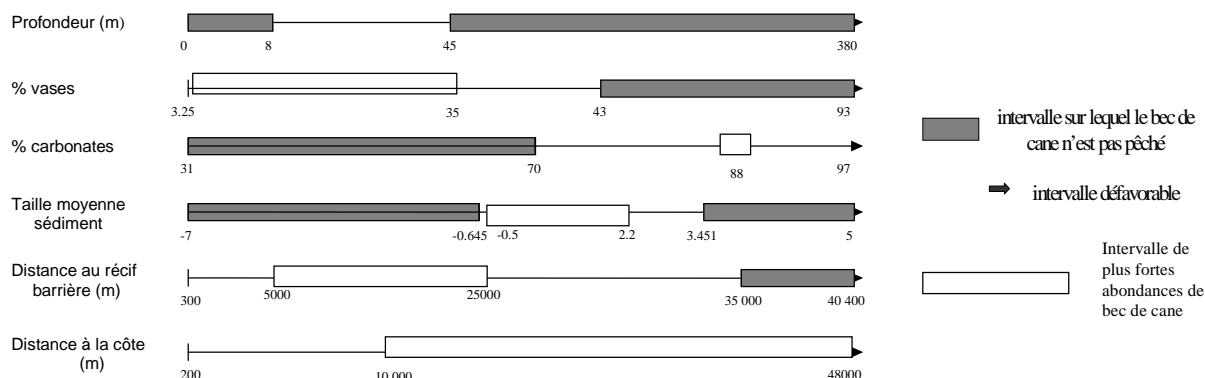


Figure n°26 : bilan de la recherche des préférences du bec de cane

👉 On retiendra que :

Nous avons délimité une zone favorable qui représente une estimation de l'habitat potentiel du bec de cane. Cette zone favorable correspond à un ensemble de zones dont des caractéristiques environnementales se situent dans les limites de tolérance de cette espèce.

En dehors de ces limites l'espèce n'est pas présente, et à l'intérieur de celles-ci, l'espèce présente des préférences qui ont des conséquences sur sa biologie (abondance, taille, ...).

C'est sur ce postulat que repose la notion d'indicateur biologique : peut on déceler des changements de l'état du milieu en étudiant les variations de certaines caractéristiques biologiques des espèces ou familles rencontrées habituellement dans ce milieu ?

les espèces représentées ; cette variable est finalement retenue comme estimateur de l'influence terrigène.

Sélection des variables biologiques les plus corrélées à l'estimateur de l'influence terrigène :

La plupart des variables biologiques sont corrélées positivement à la distance à la côte. Parmi elles sont sélectionnées les variables "PUE totale", "PUE labridés", "PUE classe trophique 1 (macrocarnivores-piscivores)" et "PUE B.perditio".

A l'inverse, quelques variables sont corrélées négativement à la distance à la côte. Ces variables peuvent également être utilisées dans la recherche d'indicateur biologiques, et les variables "PUE N.metopias" et "PUE S.undosquamis" sont sélectionnées.

On dispose donc à l'issue de cette première analyse de 6 variables biologiques potentiellement utilisables comme indicateur biologique de l'effet terrigène. Or la sélection de ces variables est basée sur l'analyse de leur corrélation avec la distance à la côte sans qu'il y ait prise en compte de l'aspect spatial des données. Un SIG permet cette prise en compte, et peut ainsi apporter des éléments déterminants pour la sélection des variables.

4.1.2. Utilisation de l'outil SIG :

Un SIG peut se révéler déterminant dans la recherche d'indicateur biologique puisqu'il permet l'étude des répartition spatiales des variables disponibles. Ainsi, il est possible d'utiliser la répartition spatiale comme un critère de sélection des variables biologiques.

Pour cela, on compare les niveaux de proximité existant entre les répartitions spatiales des couples "distance à la côte" – "variable biologique".

Résultats obtenus avec la méthode de l'indice de distance :

| Couple de variables | indice de distance |
|--|--------------------|
| distance à la côte - PUE classe trophique 1 | 0.661 |
| distance à la côte – PUE totale | 0.735 |
| distance à la côte – PUE B. perditio | 0.759 |
| distance à la côte – PUE labridés | 0.847 |
| distance à la côte – PUE S. undosquamis | 1.294 |
| distance à la côte – PUE N. metopias | 1.412 |

L'indice de distance traduit l'écart moyen mesuré entre les pixels des répartitions spatiales de deux variables normalisées : plus l'indice est élevé, plus l'écart moyen entre les répartitions spatiales est important.

La comparaison des indices obtenus indique que, parmi les 6 variables biologiques sélectionnées, la variable présentant la répartition spatiale la plus proche de celle de la distance à la côte est la variable "PUE classe trophique 1".

D'après cette méthode, parmi toutes les variables biologiques disponibles à l'issue des échantillonnages à la palangre, c'est cette variable qui présente la plus forte corrélation avec l'estimateur de l'influence terrigène : c'est elle qui doit être retenue pour servir d'indicateur biologique de l'effet terrigène.

Néanmoins, on observe que les deux variables qui montraient après l'ACP des corrélations négatives avec la distance à la côte présentent ici les indices de distance les plus élevés. En effet, la méthode de calcul de l'indice de distance conduit à écarter les variables qui présentent des corrélations négatives avec la variable estimateur. Or un indicateur biologique d'une influence n'est pas nécessairement une variable corrélée positivement à cette influence. Les résultats

obtenus à l'issue de cette méthode doivent donc être complétés par ceux de la méthode des recouvrements de zones.

Résultats obtenus avec la méthode des recouvrements de zones :

| Couple de variables | valeur du Khi2 |
|---|--------------------|
| distance à la côte - PUE classe trophique1 | 271 638.989 |
| distance à la côte – PUE totale | 247 758.799 |
| distance à la côte – PUE B. perditio | 175 202.922 |
| distance à la côte – PUE N. metopias | <u>171 671.124</u> |
| distance à la côte – PUE labridés | 143 192.413 |
| distance à la côte – PUE S. undosquamis | 51 957.828 |

La valeur du khi2 obtenue pour chaque couple de variables constitue une mesure du degré d'indépendance des répartitions spatiales : plus deux variables ont des répartitions spatiales indépendantes, plus la valeur de khi2 obtenue est faible.

La comparaison de ces valeurs indique à nouveau que la variable "PUE classe trophique1" présente la répartition spatiale la plus liée à celle de la distance à la côte. Parmi les variables biologiques disponibles à l'issue des échantillonnages à la palangre, c'est donc cette variable qui pourra être retenue pour constituer un indicateur biologique de l'effet terrigène.

On constate également que, par rapport aux résultats obtenus par la méthode de l'indice de distance, le classement des variables biologiques selon leur corrélation spatiale avec la distance à la côte n'est pas le même : la variable "PUE N. metopias", qui présentait l'indice de distance le plus élevé apparaît ici en quatrième position.

Finalement, les deux méthodes proposées convergent : parmi les variables biologiques disponibles à l'issue des échantillonnages à la palangre de 1984-87 et de 2000-01, c'est la variable "PUE classe trophique 1" qui est la plus corrélée à la distance à la côte. C'est elle qui pourra être retenue pour constituer un indicateur biologique de l'effet terrigène.

Etant corrélée positivement à la distance à la côte, on pourra l'utiliser selon : "en dehors de toute autre influence, plus les PUE de la classe trophique 1 en un point sont élevées, plus l'influence terrigène en ce point est faible".

CHAPITRE 4. DISCUSSION :

1. Intérêts de l'approche SIG dans ce type de recherche

L'utilisation d'un SIG permet de disposer de **données spatialisées**. Cette prise en compte de l'aspect spatial permet :

- une représentation visuelle des données : la carte constitue un outil d'analyse puissant (représentation des caractéristiques des données, mise en évidence de tendances, de possibles relations avec d'autres paramètres ...)
- une interpolation des données : on obtient, à partir de points de référence, une information estimée couvrant toute la zone étudiée
- le croisement d'informations de thématiques différentes : cela permet l'étude de leurs relations, le résultat de leur combinaison
- une zonation de l'espace selon différents critères
- l'utilisation de la répartition spatiale comme caractéristique d'une variable, au même titre que la moyenne ou l'écart type.

2. Validité de la démarche

2.1. Pré-traitement des données

Les communautés de poissons sont sous l'influence de très nombreuses composantes qui modifient le milieu et le rendent plus ou moins favorable aux différentes espèces. Ces composantes présentent des variations spatio-temporelles qui compliquent l'association et l'exploitation de données acquises à des dates et dans des lieux différents. Néanmoins, nous avons choisi dans notre démarche d'exploitation d'associer l'ensemble des données issues des campagnes d'échantillonnage à la palangre de 1984-87 et 2000-01. Une telle association conduit à introduire dans les données plusieurs sources de variabilité :

- variabilité spatiale
- variabilités temporelles (saisonnière et interannuelle).

La méthode de pré-traitement des données proposée représente une tentative de réduction, sur les données ichtyologiques échantillonnées, des conséquences de la variabilité saisonnière. Les saisons sont identifiées à l'aide de la température de l'eau. Cette méthode repose sur :

- **l'observation, sur les données échantillonnées, des conséquences de la variabilité saisonnière.** Les connaissances sur la biologie des poissons ne sont pas suffisantes pour permettre de connaître le comportement des espèces en fonction des saisons. C'est pourquoi nous utilisons l'ensemble des données échantillonnées à la palangre pour considérer un peuplement de poissons type sur lequel nous mesurons les conséquences de l'effet des saisons.

Or les données disponibles ne concernent pas le suivi temporel d'un même site, mais des échantillonnages effectués dans des lieux différents. Les saisons ne sont donc pas la seule source de variabilité des données disponibles.

- **l'identification des saisons.** Nous avons choisi de découper l'années en saisons à l'aide de la courbe de variation des températures moyennes observées sur 30 ans à Nouméa (Figure n°8). Cette courbe moyenne nous a conduit à appliquer le même traitement aux données de 1984-87 et à celles de 2000-01.

Il aurait été préférable, pour prendre en compte de possibles variations interannuelles (notamment phénomènes El Niño, La Niña), d'utiliser les courbes de variations de température annuelles et d'appliquer un pré-traitement correspondant aux années d'échantillonnage de chaque campagne.

- **la correction des données ichtyologiques.** Nous avons appliqué à chaque variable ichtyologique un coefficient de correction. Ce coefficient correspond à l'écart observée entre les moyennes de cette variable pour les différentes saisons et la moyenne globale. Les données sont ainsi corrigées de façon grossière.

Dans notre démarche, par rapport aux objectifs et aux données disponibles, seule la variabilité intra annuelle a été étudiée, en prenant pour référence la température de l'eau. Cette méthode pourrait s'appliquer de la même manière pour l'étude de la variabilité d'autres paramètres. Elle constitue une tentative de réduction des conséquences d'une source de variabilité extérieure sur les données. Elle ne peut et ne doit être appliquée que dans des cas où son manque de rigueur n'est pas préjudiciable pour les résultats attendus.

2.2. Utilisation des résultats d'interpolations

L'utilisation d'interpolations permet de disposer, à partir de données ponctuelles, d'une information couvrant toute la zone d'étude. Ceci permet d'une part le croisement de données n'ayant pas été échantillonnées aux mêmes endroits, d'autre part un découpage du lagon sud ouest en différentes zones.

Néanmoins, cela conduit notre méthode à :

- **utiliser les résultats de la seule méthode d'interpolation de "Savane".** Cette méthode barycentrique permet de prendre en compte, pour estimer la valeur en un point, les valeurs des huit points les plus proches, répartis dans huit octants. Elle affecte aux valeurs des points les plus proches un poids plus important qu'à celles des points éloignés. Son principal défaut est de nécessiter l'utilisation de l'option "deux passages". Cette option, prévue pour réduire les temps de calcul, permet également de lisser les résultats et de faire disparaître les structures en "étoiles" observées lors d'interpolations à un seul passage. Ce lissage est à double tranchant : il diminue les artefacts obtenus à l'issue du premier passage mais l'interpolateur final n'est plus exact.
- **être soumis aux effet de bord.** Ces effets induisent un biais dans les données estimées dans les zones situées en limite de la zone interpolée (côte et récif barrière).
- **exploiter des interpolations de "qualités" différentes.** La qualité d'une interpolation correspond à la précision avec laquelle la zone interpolée est décrite. Cette précision dépend directement de la taille du pixel d'interpolation, et nous avons fixé celle ci à la distance minimale séparant deux points d'échantillonnage. Notre choix permet d'adapter la précision de l'interpolation à la quantité d'information échantillonnée, mais nous amène également à comparer des données n'ayant pas subi le même effort d'échantillonnage et les mêmes traitements.
- **ignorer l'hypervariabilité spatiale des données.** Le nombre insuffisant de points d'échantillonnage et le lissage du à la méthode d'interpolation ne reflète pas l'hypervariabilité rencontrée dans ce type de milieux et ne permet pas l'étude des microhabitats.

2.3. Délimitation et description d'habitat

La méthode proposée repose sur les possibilités de croisement d'informations qu'offre un SIG pour permettre :

- **une délimitation de la zone potentiellement couverte par une espèce**

La méthode s'apparente à une recherche des seuils de présence-absence de l'espèce pour tous les paramètres environnementaux. La délimitation est basée sur le postulat suivant : "la pêche de l'espèce A en un point signifie la présence de cette espèce en ce point et inversement". Nous avons ainsi considéré les pêches à la palangre comme des échantillonnages représentatifs de la présence ou de l'abondance de la faune en un point.

Or les plans d'échantillonnages à la palangres ont été conduits dans un objectif **halieutique** d'estimation des stocks des fonds meubles du lagon, et l'utilisation d'un engin de pêche ne garantit pas un échantillonnage exhaustif des espèces présentes. La pêche à la palangre cible ainsi les espèces carnivores.

Néanmoins, cette méthode permet de délimiter, **à partir d'observations ichtyologiques**, une zone potentiellement couverte par une espèce. Cette zone satisfait toutes les conditions environnementales favorables à cette espèce apparues lors du croisement des données ichtyologiques et des paramètres environnementaux interpolés.

- **une description des préférences de l'espèce pour les différents paramètres environnementaux décrivant cette zone**

La méthode repose sur l'exploitation des données d'abondance de l'espèce : si une espèce est abondante dans une zone, cela signifie que les caractéristiques de cette zone lui conviennent. Cette abondance peut se trouver largement surestimée si un banc est rencontré. De nombreuses espèces capturées à la palangre présentent un comportement grégaire.

Encore une fois c'est l'observation de la grandeur biologique qui permet d'identifier les caractéristiques des zones correspondantes.

2.4. Recherche d'indicateurs biologiques

La méthode proposée repose sur les possibilités offertes par un SIG dans l'exploitation de la répartition spatiale d'une variable.

Méthode de l'indice de distance :

Cette méthode repose sur le calcul d'un écart moyen observé entre les valeurs relevées sur deux cartes au moyen d'une grille de points arbitraires. Les points de la grille doivent être en nombre suffisant et répartis de manière homogène pour permettre une comparaison complète des deux cartes. La grille utilisée ici est constituée des 150 000 points de relevés bathymétriques effectués par le SHOM dans la zone d'étude. Ce nombre de points est excessif. Cela a pour conséquence d'utiliser une grille de points de comparaison dont le pas est plus faible que la résolution des cartes comparées (150-200m). Ceci n'a pas de conséquence sur les résultats obtenus, mais il serait préférable d'utiliser une grille de points ordonnés et répartis de manière homogène, et d'adapter le pas de la grille utilisée à la résolution des cartes comparées.

Cette méthode permet de mesurer la proximité existant entre deux cartes présentant des répartitions spatiales **corrélées positivement** (on rappelle que cela signifie des répartitions spatiales présentant des structures similaires : là où la variable A augmente, la variable B fait de même et inversement). Dans le cas contraire, l'indice de distance ne peut être utilisé.

Une méthode similaire a été développée par Herzfeld et Merriam (1988). La comparaison des cartes s'effectuait pour chaque pixel et conduisait à la création d'une carte des écarts obtenus après normalisation.

Méthode de recouvrement de zones :

Cette méthode repose sur l'exploitation d'une information zonale et mesure le degré d'indépendance des répartitions spatiales des variables.

Elle repose sur une classification en 5 modalités (très pauvre, pauvre, moyen, abondant, très abondant) appliquée aux distributions des variables étudiées. Les classes obtenues définissent des zones sur chacune des répartitions spatiales. Les zones de différentes répartitions spatiales peuvent être comparées puisqu'elles les mêmes modalités, appliquées aux différentes variables.

La comparaison des surfaces des intersections observées avec les surfaces des intersections théoriques aboutit à une valeur du χ^2 : plus cette valeur est faible, plus les répartitions spatiales étudiées sont indépendantes. Dans cette méthode, la sensibilité d'une variable à une influence se mesure par le degré d'indépendance existant entre leur répartitions spatiales.

La partition des valeurs de toutes les variables en 5 classes autorise la comparaison des valeurs du χ^2 obtenues pour chaque couple de cartes.

Par rapport à la précédente, cette méthode mesure le degré d'indépendance de deux répartitions spatiales. Elle peut donc être utilisée pour mesurer la proximité existant entre deux cartes présentant des répartitions spatiales corrélées positivement **et** négativement.

Cette méthode est ici mise en œuvre pour tenter d'identifier un indicateur biologique de l'influence terrigène, mais elle est transposable à toute mesure de la corrélation spatiale existant entre deux variables.

3. Validité des résultats

3.1. Délimitation et description de l'habitat du bec de cane

Notre recherche de l'habitat du bec de cane s'est basé sur la recherche des caractéristiques physiques des milieux dans lesquels le bec de cane est rencontré. Cette définition de l'habitat est discutée.

Ainsi, pour C. Lévêque (1995), un habitat ne serait pas seulement un ensemble de caractéristiques physiques du milieu mais la conséquence d'un héritage phylogénique et de stratégies adaptatives. L'héritage phylogénique correspond à l'ensemble des traits vitaux de l'espèce, sélectionnés par l'évolution (comportements reproducteurs, habitudes et spécialisations alimentaires ...). Ces traits vitaux font peser sur l'espèce un certain nombre de contraintes abiotiques, biologiques et comportementales qui vont déterminer ses besoins en terme d'habitat.

Les stratégies adaptatives correspondent à la possibilité offerte aux espèces de développer des tactiques, des réponses adaptatives aux modifications du milieu dans lequel elles vivent. Ces stratégies (qui peuvent s'avérer nécessaires à la survie de l'espèce) permettent à l'espèce de développer des comportements alternatifs, selon les conditions ambiantes qui lui sont offertes et auxquelles elle ne peut échapper.

Il faudrait donc estimer la qualité de la zone délimitée pour une espèce en fonction de ses capacités d'adaptation.

On observe que la zone proposée pour le bec de cane couvre pratiquement toute la zone d'étude. Ce résultat vérifie le caractère ubiquiste déjà connu de cette espèce. Néanmoins, la précision de la

zone proposée dépend de la quantité d'informations disponibles. Pour proposer une délimitation précise, il faudrait disposer de :

- davantage de paramètres environnementaux (couverture algale, faune benthique (oursins, holothuries ...))
- davantage de points d'échantillonnage ichtyologique, réalisés à cette fin, afin de mieux évaluer la présence ou l'absence de l'espèce
- davantage de mesures des paramètres environnementaux afin de percevoir les différents micro-habitats présents dans le lagon.

La recherche des préférences du bec de cane pour les différents paramètres disponibles permet de connaître l'environnement qui lui semble optimal. Cette description pourrait être approfondie afin de tenter de dégager la variable la plus restrictive pour cette espèce. Cette variable pourrait devenir un indicateur de la présence du bec de cane.

3.2. Identification d'indicateur biologiques

Les méthodes mises en œuvre ont permis d'identifier, parmi toutes les variables biologiques acquises lors des campagnes d'échantillonnage à la palangre et intégrées dans la base "LagonSO", la variable la plus corrélée à la distance à la côte. Il s'agit de la variable "PUE de la classe trophique 1". La classe trophique "macrocarnivores-piscivores est peut être le groupe d'espèces le plus sensible à l'influence terrigène, mais la PUE n'est peut être pas le paramètre qui exprime le mieux cette sensibilité. Il faudrait donc poursuivre cette analyse avec les données de taille, de poids, ou d'indices construits à partir de ces paramètres pour identifier, parmi les espèces échantillonnées à la palangre, l'espèce et le paramètre décrivant cette espèce qui pourront être retenus pour constituer un indicateur biologique de l'influence terrigène.

Une démarche identique pourra être conduite pour tenter d'identifier des indicateurs biologiques d'autres influences (pêche, fréquentation touristique, eutrophisation, ...).

A terme, de tels indicateurs pourront être utilisés pour estimer la qualité du milieu rencontré en un point. Ils pourraient également être utilisés, dans un objectif de protection et de gestion du milieu, pour constituer des outils d'aide à la décision.

Conclusion

Le principal intérêt dans l'utilisation d'un SIG à des fins de recherche est la prise en compte et l'exploitation de l'aspect spatial des données. La démarche présentée dans ce travail a permis de montrer que les fonctionnalités offertes par un SIG s'avèrent pertinentes pour atteindre les objectifs de recherche fixés :

Etudier de la répartition et de la structuration spatiale des peuplements ichtyologiques du lagon sud ouest de Nouvelle Calédonie :

- interpolation des données ponctuelles et disposer d'une information couvrant toute la zone d'étude;
- visualisation des répartitions spatiales et observation de la nécessité de réduire les conséquences des saisons sur les données ichtyologiques;
- restitution sous forme cartographique des résultats du pré-traitement des données;
- zonation de l'information ichtyologique faisant apparaître les structures spatiales.

Délimiter et décrire des habitats :

- croisement des données ichtyologiques avec des données sédimentologiques et bathymétriques (description du milieu au niveau des points d'échantillonnage mais aussi des zones d'abondance);
- proposition d'une délimitation de la zone potentiellement couverte par une espèce.

Participer à la définition d'indicateurs biologiques de l'effet terrigène :

- exploitation de la répartition spatiale comme critère de sélection des variables biologiques les plus sensibles à l'influence terrigène.

Les différentes méthodes proposées, améliorées et adaptées en fonction des objectifs à atteindre, sont utilisables pour d'autres problématiques. Elles constituent un exemple des potentialités offertes par un SIG dans l'exploitation de données scientifiques.

Bibliographie

ARNAUD M., EMERY X., Estimation et interpolation spatiale, Hermes, 2000.

BOURDERIE J., Implémentation de la base de données géographiques du programme ECOTREOPE à l'aide du SIG SAVANE et traitement des données bathymétriques et sédimentologiques du lagon sud-ouest de Nouvelle-Calédonie, Rapport de stage pour l'obtention du diplôme de Technicien Supérieur de la Mer (Intchmer), 2000.

BOUVET G., Identification de biotopes lagunaires par imagerie Landsat 7 et couplage SIG, Mémoire de fin d'études, 2000.

CHARDY P., CHEVILLON C., CLAVIER J., Major benthic communities of the south west lagoon of New Caledonia, Coral Reefs n°7, pp 69-75, 1988.

COREUS, Approche écosystémique des Communautés Récifales et de leurs USages dans le Pacifique sud, présentation détaillée du projet scientifique, document interne IRD, février 2001.

GRIMAUD J., KULBICKI M., Influence de la distance à l'océan sur les peuplements ichthyologiques des récifs frangeants de Nouvelle Calédonie, Sciences de la vie n°321, pp923-931, 1998.

HABERT E., Savane version 7.05, IRD de Bondy, version provisoire de janvier 2000

HABERT E., Savateca version 7.05, IRD de Bondy, version provisoire de janvier 2000

HABERT E., Savedit version 7.05, IRD de Bondy, version provisoire de janvier 2000

HABERT E., Savamer version 7.05, IRD de Bondy, version provisoire de janvier 2000

HERZFELD U.C., SONDERGARD M.A., MAPCOMP a fortran program for weighted thematic map comparison, Computers & Geosciences Vol 14 n°5, pp 699-713, 1988.

KULBICKI M., GALZIN R., LETOURNEUR Y., MOU-THAM G., SARRAMEGNA S., THOLLOT P., SCHERRER B., Biostatistique, Gaëtan Morin Editeur, pp 104, 165-168, 1984.

LEBART L., MORINEAU A., PIRON M., Statistique exploratoire multidimensionnelle, Dunod (3^{ème} édition), pp 33-67, 223-250, 2000.

LETOURNEUR Y., KULBICKI M., LABROSSE P., Spatial structure of commercial reef fish communities along a terrestrial runoff gradient in the northern lagoon of New Caledonia, Environmental Biology of Fishes n°51, pp 141-159, 1998.

LEVEQUE C., L'habitat : être au bon endroit au bon moment ?, Bulletin français de la pêche et de la pisciculture n°337/338/339, pp 9-20, 1995.

PNEC (Programme National Environnement Côtier), Formalisation d'indicateurs des écosystèmes récifo-lagonaires dans une perspective écologique, chantier Nouvelle Calédonie, fichier informatique, 2000.

THOLLOT P., Caractérisation des communautés marines et suivi des peuplements ichthyologiques des îlots du parc du lagon sud placés en réserve, T&W Consultants, avril 1999.

THOLLOT P., Observatoire des récifs coralliens Actions menées en 1998, T&W Consultants, juin 1999.

WANTIEZ L., THOLLOT P., KULBICKI M., Effects of marine reserves on coral reef fish communities from five islands in New Caledonia, Coral Reefs n°16, pp 215-224, 1997.

WANTIEZ L., CHAUVET C., Les peuplements de poissons de la réserve marine du récif Aboré (Nouvelle Calédonie) : composition spécifique, structures trophique et démographique avant l'ouverture de la pêche, Documents scientifiques et techniques ORSTOM, janvier 1996.

Documentation Internet :

<http://www.obs-vlfr.fr/~enseigne/anado/series/chap3.htm>

Documentation sur le variogramme.

<http://www-fourier.ujf-grenoble.fr/~parisse/geostat/geostat.html>

documentation sur les méthodes d'interpolation et les géostatistiques.

<http://web2.uqat.quebec.ca/certinfo/cours/certinfo/gen3307/archives/ProjetA98/baccalau.htm>

documentations en pdf sur le krigage et liens vers d'autres sites de documentation sur le krigage..

<http://www.gypsymoth.ento.vt.edu/~sharov/PopEcol/lec3/geostat.html>

tutoriel de géostatistique.

ANNEXES

Annexe 1 : Indices granulométriques

Classification appliquée à la Taille moyenne ou « MEAN SIZE »

Limites selon Wentworth (1922) :

| | | | | | | | | | | | | | | | | |
|------------|--------------|-----------------------|--------------|------------------|-------------|--------------------|------------|--------------------|----------|-------------------------|-----------|----------------|-----------|---------------|-----------|-------------------------|
| Vase | 3.98 | Sable très fin | 3 | Sable fin | 2 | Sable moyen | 1 | Sable grossier | 0 | Sable très grossier | -1 | Granule | -2 | Gravier | -6 | Galet (φ) |
| VA | | STF | | SF | | SM | | SG | | STG | | GN | | GV | | GT |
| <i>Mud</i> | 0.063 | <i>Very fine sand</i> | 0.125 | <i>Fine sand</i> | 0.25 | <i>Medium sand</i> | 0.5 | <i>Coarse sand</i> | 1 | <i>Very coarse sand</i> | 2 | <i>Granule</i> | 4 | <i>Pebble</i> | 64 | <i>Cobble</i> mm |

Classification appliquée au Triage ou « STANDARD DEVIATION »

Limites adaptées à l'étude des milieux récifaux par Weydert (1971) :

| | | | | | | | | |
|------------------------------|--|-------------------------|--|--------------------|--|----------------------|--|---------------------------|
| Extrêmement bien trié | | Très bien trié | | Bien trié | | Mal trié | | Très mal trié |
| EBT | | TBT | | BT | | MT | | TMT |
| 0.5 | | 1.1 | | 1.5 | | 2 | | 1 |
| <i>Extremely well sorted</i> | | <i>Very well sorted</i> | | <i>Well sorted</i> | | <i>Poorly sorted</i> | | <i>Very poorly sorted</i> |

Classification appliquée à l'Asymétrie ou « SKEWNESS »

Limites selon Falk & Ward (1957)

| | | | | | | | | | | |
|-----------|-----------------------------|--|--------------------|--|---------------------------|--|----------------------|--|-------------------------------|--|
| | Très négative | | Négative | | Symétrie | | Positive | | Très positive | |
| | AT- | | A- | | S | | A+ | | AT+ | |
| -1 | 0.3 | | -0.1 | | 0.1 | | 0.3 | | 1 | |
| | <i>Strongly fine skewed</i> | | <i>Fine skewed</i> | | <i>Nearly symmetrical</i> | | <i>Coarse skewed</i> | | <i>Strongly coarse skewed</i> | |

Classification appliquée à la Normalité ou « KURTOSIS »

Limites selon Falk & Ward (1957)

| | | | | | | | | | | |
|-------------------------|--|--------------------|--|---------------------|--|--------------------|--|-------------------------|--|------------------------------|
| Très peu accusée | | Peu accusée | | Moyennement accusée | | Accusée | | Très accusée | | Extrêmement accusée |
| TPA | | PA | | MA | | A | | TA | | EA |
| 0.67 | | 0.9 | | 1.11 | | 1.5 | | 3 | | 1 |
| <i>Very platykurtic</i> | | <i>Platykurtic</i> | | <i>Mesokurtic</i> | | <i>Leptokurtic</i> | | <i>Very leptokurtic</i> | | <i>Extremely leptokurtic</i> |

Annexe 2 : méthode de l'analyse de variance

On pose comme hypothèse que la variable y est une combinaison linéaire de variables explicatives supposées connues sans erreur.

Pour cela, on pose le modèle :

$$\text{pour tout } i \in [1; N], y_i = \alpha_0 + \alpha_1(V1)_i + \alpha_2(V2)_i + \dots + \alpha_n(Vn)_i + \varepsilon_i$$

- avec
- $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n$ les coefficients qui doivent être définis par le modèle;
 α_0 peut être considéré comme coefficient d'une variable explicative particulière V_0 dont les valeurs seraient toujours égales à 1;
 - V_1, V_2, \dots, V_n les variables indépendantes explicatives du modèle;
 - ε_i est le résidu représentant l'écart entre la valeur observée y_i et la partie expliquée de l'observation ($\alpha_0 + \alpha_1(V1)_i + \alpha_2(V2)_i + \dots + \alpha_n(Vn)_i$).

On dispose, pour évaluer les coefficients inconnus du modèle, d'un système de N équations linéaires ayant $N + n + 1$ inconnues. Le système admet donc une infinité de solutions.

Soient a_0, a_1, \dots, a_n les coefficients correspondants à une des solutions possibles. On cherchera la solution qui minimise, globalement, suivant un critère à définir, l'ensemble des écarts à la linéarité. C'est à dire :

choisir (a_0, a_1, \dots, a_n) qui minimise l'ensemble des e_i

$$\text{avec } e_i = y_i - \tilde{y} = y_i - (a_0 + a_1(V1)_i + a_2(V2)_i + \dots + a_n(Vn)_i)$$

Parmi tous les critères possibles de minimisation, citons les méthodes des moindres carrés ($\min(\sum e_i^2)$), des moindres valeurs absolues ($\min(\sum |e_i|)$) et celle du "minimax" ($\min(\max e_i)$). La méthode employée ici est celle des moindres carrés.

Pour savoir si une variable explicative V_k a une influence réelle sur la variable à expliquer y , on procède à un test d'hypothèse sur le coefficient de régression α_k .

L'hypothèse nulle H_0 est l'éventuelle non influence qui se traduit par :

$$H_0 \quad \alpha_k = 0 \text{ (autres coefficients quelconques)}$$

On écrit alors la statistique : $t = \frac{a_k}{s_k}$

où s_k est l'estimation de l'écart type du $k^{\text{ième}}$ coefficient de régression a_k

Si (H_0) est vraie, la statistique suit une loi de Student à $(N-n)$ degrés de liberté. A partir de cette distribution, on obtient la probabilité p_k correspondant à la valeur t_k prise par t :

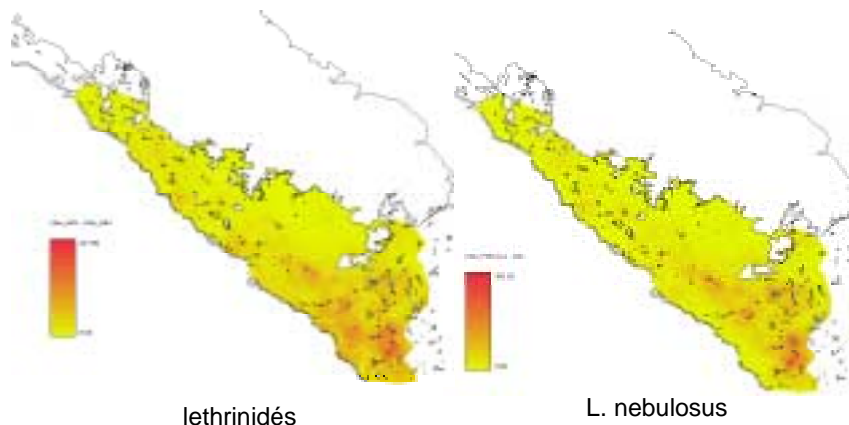
$$p_k = P(|t| \geq t_k)$$

Cette probabilité permet, selon le seuil de confiance retenu, d'accepter ou de rejeter H_0 .

Le modèle donne également les coefficients attribués à chacune des modalités des variables explicatives. Ces coefficients traduisent pour chaque variable les écarts mesurés entre la moyenne observée pour la modalité et la moyenne globale.

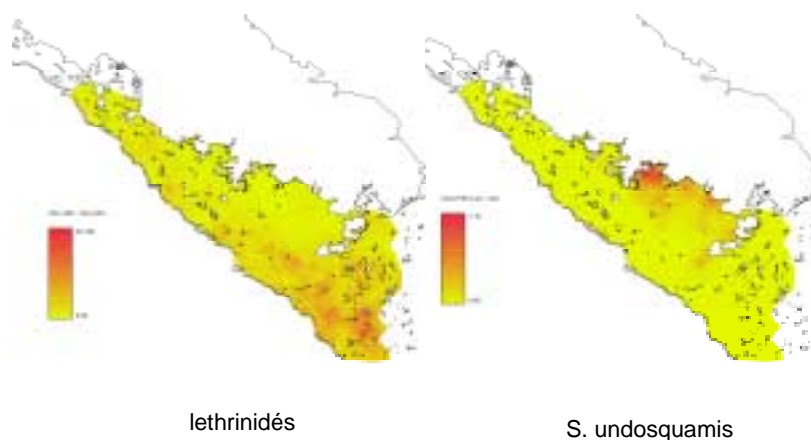
Annexe 3 : test des méthodes de comparaison de cartes :

Les tests ont été réalisés sur deux couples de cartes, représentant des situations fréquemment rencontrées. Le premier couple associe des cartes représentant des variables aux répartitions géographiques proches :



Ces cartes représentent la répartition spatiale après interpolation de variables mesurées lors d'échantillonnages à la palangre. Ce sont des prises par unité d'effort (PUE, mesurées ici en gramme par hameçon) d'une espèce (*Lethrinus nebulosus*) et d'une famille (lethrinidés).

Le deuxième couple associe des cartes représentant des variables aux répartitions spatiales différentes, voire opposées. Cela permet d'observer le comportement de nos deux mesures de ressemblance pour ce type de comparaison.



Résultats obtenus pour l'indice de ressemblance :

- couple 1 : 0.353
- couple 2 : 1.159

Cet indice mesure la différence moyenne qui existe entre les valeurs normalisées des pixels des deux cartes. Il mesure donc le niveau de différence de deux répartitions spatiales, pas le niveau de proximité de celles-ci.

Résultats obtenus par la méthode de recouvrement des classes :

- couple 1 : $\text{Khi}2 = 408\,738.728$
- couple 2 : $\text{khi}2 = 50\,481.318$

On rappelle que plus la valeur du $\text{Khi}2$ obtenue est grande, moins les répartitions spatiales des deux variables sont indépendantes.

Cette mesure permet donc de dégager des variables ayant des corrélations spatiales positives **et** négatives. Cet aspect est utile puisque la recherche de variables représentatives ne se limite pas aux variables réagissant positivement à une perturbation.

Annexe 4 : La "gratte" ou Ciguatera

Le nom de ciguatera provient de Cuba où un petit mollusque gastéropode nommé "cigua" est responsable de nombreuses intoxications alimentaires. La ciguatera concerne l'ensemble des régions coralliennes des trois océans.

Le terme "gratte", propre à la Nouvelle Calédonie, traduit le fait qu'elle est susceptible d'engendrer des troubles d'ordre neurologiques, notamment au contact du froid, tels que démangeaisons et picotements des extrémités. L'intensité des symptômes varie selon la quantité de toxine absorbée, ou la quantité de toxine accumulée. De fortes douleurs articulaires et musculaires, des frissons accompagnés d'une faiblesse générale sont fréquents. Les cas les plus graves, rares fort heureusement, entraînent la paralysie, le coma, voire la mort.

D'importants travaux de recherche ont été et sont consacrés à l'étude de la ciguatera. Ils montrent que ce phénomène est d'une extrême complexité. Ils ont abouti à la mise en évidence de plusieurs toxines accumulées par les poissons. Leur origine se situe dans des algues benthiques unicellulaires broutées par les espèces herbivores et qui s'accumulent ensuite le long de la chaîne alimentaire. Si ces algues sont présentes sur des substrats divers, leur développement semble favorisé par la mise en place de "surfaces neuves" dans le milieu marin. C'est ainsi que des flambées cyguatoxiques ont été fréquemment observées dans des zones concernées par des implantations de structures portuaires, des échouages de bateaux, et parfois des cyclones dévastateurs entraînant de profonds bouleversements des biotopes.

Si elle affecte principalement les espèces carnivores, toutes ne sont pas toxiques.

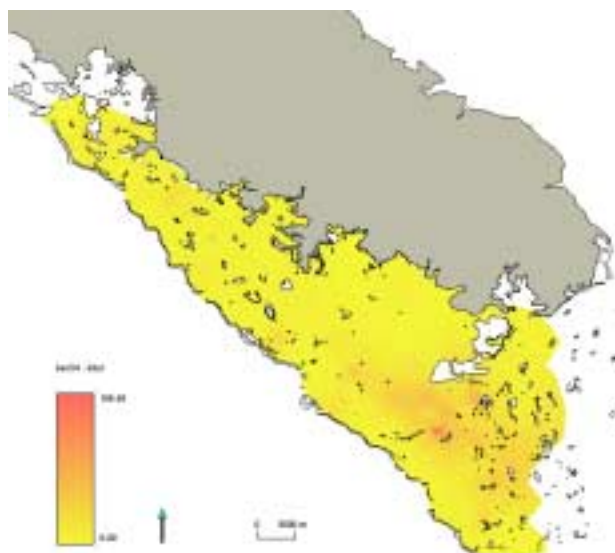
De très nombreux traitements inspirés des pharmacopées traditionnelles sont utilisées pour lutter contre le "gratte". Ils agissent avec des succès variés, largement fonction de la sensibilité individuelle des personnes atteintes.

Source : Poissons de Nouvelle Calédonie, LABOUTE L., GRANDPERRIN R., IRD Editions Catherine Ledru, 2000.

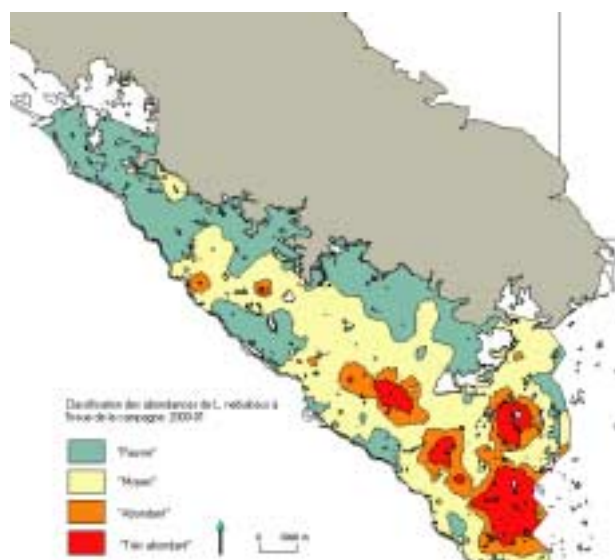
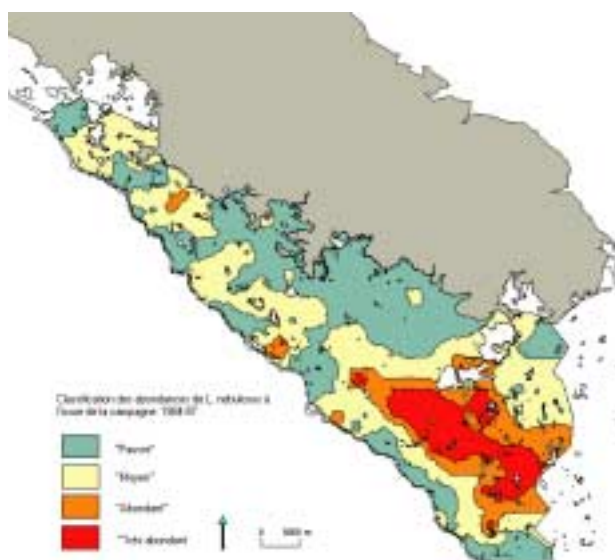
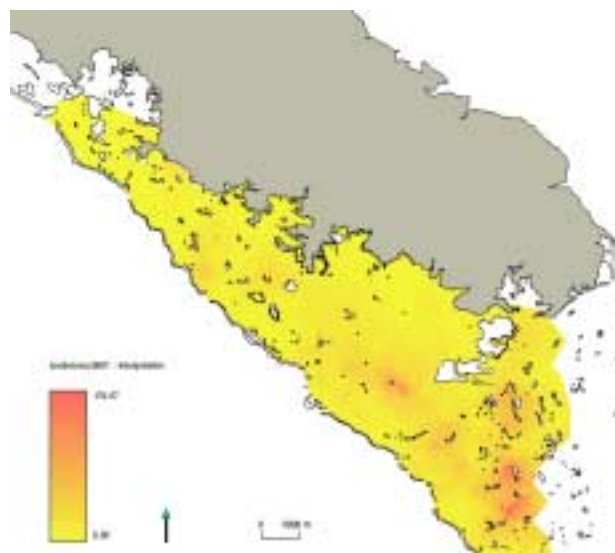
Annexe 5 : Répartition spatiale du *Lethrinus nebulosus*

La représentation des interpolations et des classifications effectuées sur les PUE du *Lethrinus nebulosus* mesurées lors des campagnes d'échantillonnage à la palangre est la suivante :

Campagne 1984-87



Campagne 2000-01

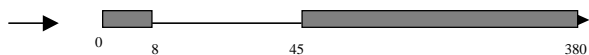
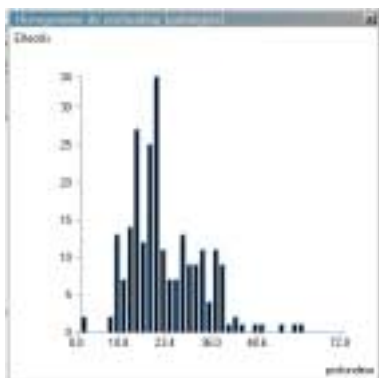


On observe que le sud du lagon correspond à la zone où le bec de cane se rencontre en abondance. On remarque toutefois que la localisation des maxima obtenus diffère d'une campagne sur l'autre. Ces différences sont peut être dues à une modification de l'influence anthropique sur cette espèce.

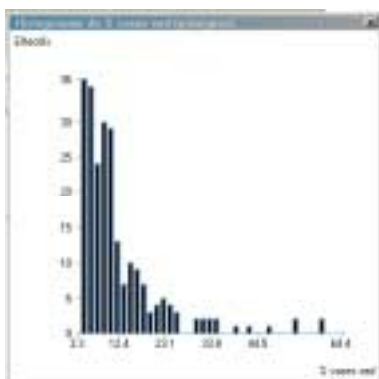
Annexe 6 : Détails des résultats de la recherche des limites de tolérance du *Lethrinus nebulosus*

La recherche des limites de tolérance s'est effectuée à partir des histogrammes des valeurs des différents paramètres environnementaux disponibles relevées là où le *Lethrinus nebulosus* a été pêché.

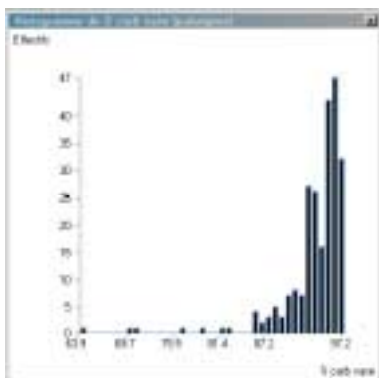
Profondeur



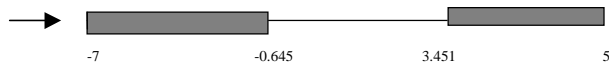
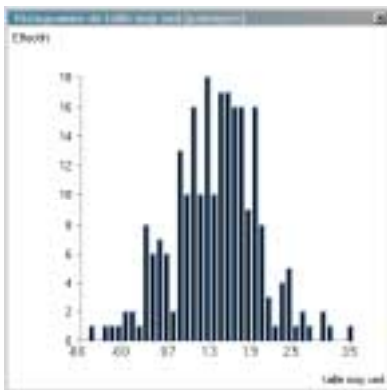
% vases du sédiment



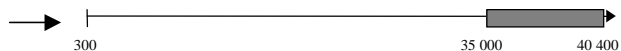
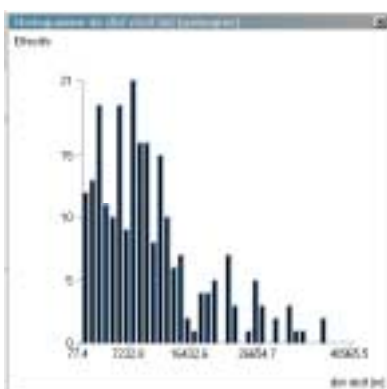
% carbonates des vases



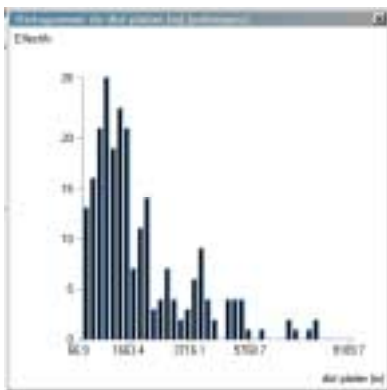
Taille moyenne du sédiment



Distance au récif



Distance aux platiers



L'ensemble de ces intervalles est utilisé pour proposer une délimitation de la zone potentiellement couverte par *Lethrinus nebulosus*, satisfaisant toutes les conditions apparues.

Annexe 7 : Détails des résultats de la recherche des préférences environnementales du *Lethrinus nebulosus*

La description des préférences environnementales du *Lethrinus nebulosus* repose sur la comparaison, pour chaque variable environnementale, des distributions des valeurs des pixels présents dans chacune des zones d'abondance.

Ces comparaisons permettent, pour chaque variable, :

- d'observer si une tendance se dégage : glissement progressif du mode ou resserrement autour du mode des distributions des fréquences ;
- d'identifier éventuellement l'intervalle des valeurs optimales pour le bec de cane.

Profondeur :

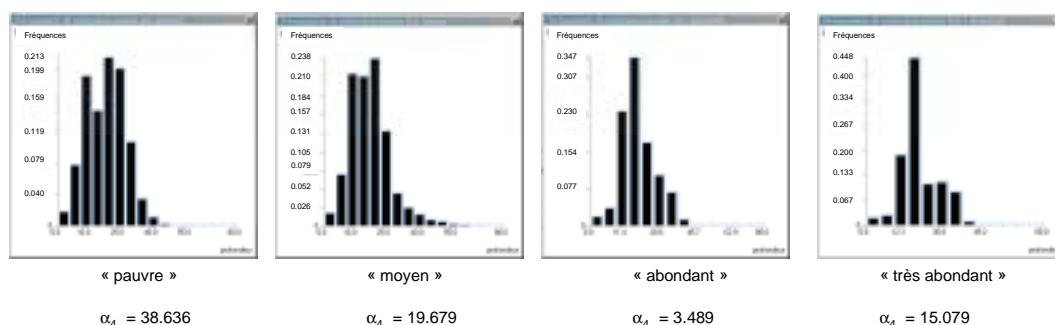


figure n° : histogrammes des fréquences des profondeurs rencontrées dans chaque zone d'abondance

L'observation de ces distributions ne révèle pas de tendance générale :

- les modes sont les mêmes, situés entre 18 et 20 m;
- on observe une tendance à l'aplatissement des distributions : α_4 pauvre > α_4 moyen > α_4 abondant;

Il semble donc que le bec de cane ne présente pas de préférences de profondeur : il est présent, quelle que soit son abondance, de 0 à 40m, avec un pic à 18-20 m qui correspond à la profondeur moyenne de la plaine lagonaire.

Teneur en vases du fond :

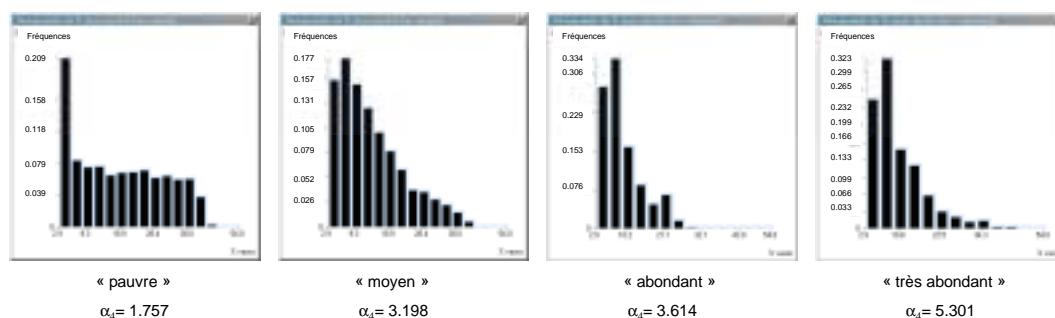


figure n° : histogrammes des fréquences des teneurs en vases rencontrées dans chaque zone d'abondance

La comparaison des distributions des teneurs en vases du fond dans les différentes zones d'abondance montre :

- un glissement progressif des modes : il se situe à 5% de vases dans la zone "pauvre" et à 7 % dans la zone "très abondant";
- un resserrement progressif des distributions autour des modes :

$$\alpha_4 \text{ pauvre} < \alpha_4 \text{ moyen} < \alpha_4 \text{ abondant} < \alpha_4 \text{ très abondant}$$

Il apparaît donc que le bec de cane, dans sa zone favorable correspondant à des teneurs en vases du fond comprises entre 3.25 et 43 %, soit plus abondant pour des teneurs en vases comprises entre 3.25 et 35 %, avec un optimum observé à 7 %.

Teneur en carbonates des vases :

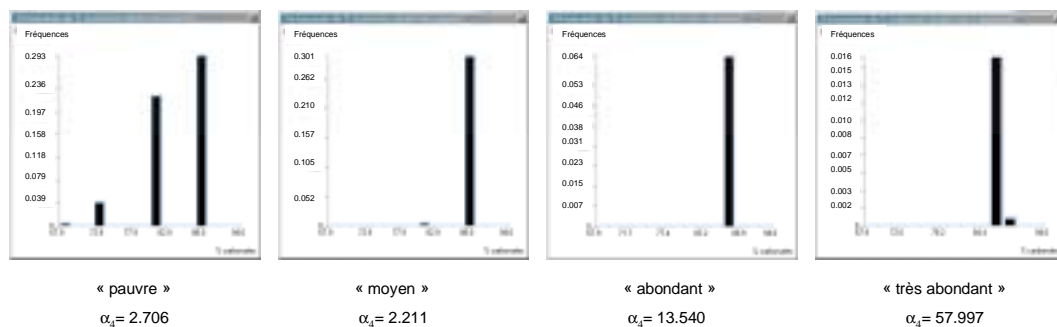


figure n° : histogrammes des fréquences des teneurs en carbonates des vases rencontrées dans chaque zone d'abondance

La comparaison des distributions des teneurs en vases du fond dans les différentes zones d'abondance montre :

- les modes sont les mêmes, situés à 88 % de carbonates;
- un resserrement des distributions autour du mode : $\alpha_4 \text{ moyen} < \alpha_4 \text{ abondant} < \alpha_4 \text{ très abondant}$;

Il semble donc que le bec de cane, dans sa zone favorable correspondant à des teneurs en carbonates des vases comprises entre 70 et 97 %, soit plus abondant pour des teneurs en carbonates égales à 88 %.

Taille moyenne du sédiment :

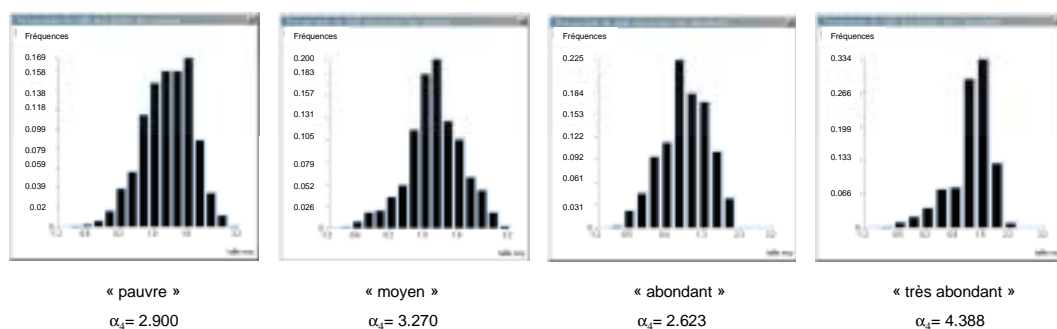


figure n° : histogrammes des fréquences des tailles moyennes du sédiment rencontrées dans chaque zone d'abondance

La comparaison des distributions des tailles moyennes du sédiment dans les différentes zones d'abondance montre que la distribution de la zone "abondant" se démarque :

- son mode se situe à 0.7 alors que les modes des autres distributions se situent autour de 1.6;
- son coefficient d'aplatissement est le seul à être inférieur à celui de la zone "pauvre".

La distribution de la zone "abondant" ne permet pas de dégager une tendance générale. Néanmoins, on observe que les étendues des distributions diminuent et que, pour les zones "abondant" et "très abondant", les distributions sont comprises entre les valeurs -0.5 et 2.2.

S'il n'est pas possible à partir de ces histogramme de dégager l'optimum de taille moyenne du bec de cane, on peut considérer que, dans sa zone favorable correspondant à des tailles moyennes du

sédiment comprises entre -0.64 et 3.45 , le bec de cane est plus abondant pour des tailles moyennes comprises entre -0.5 et 2.2 .

Distance au récif :

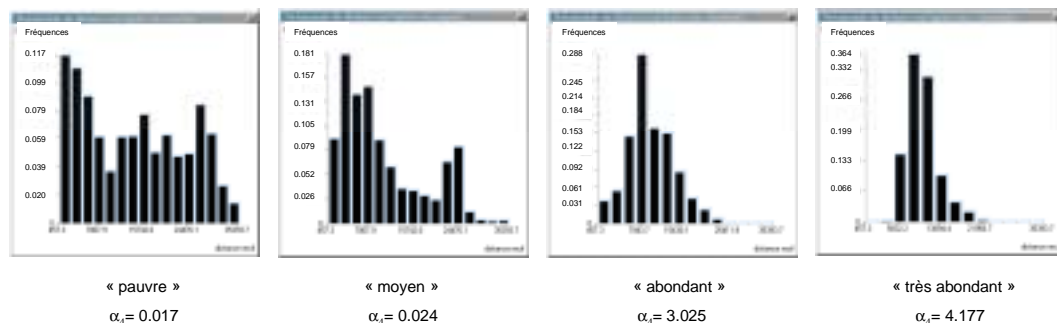


figure n° : histogrammes des fréquences des distances au récif mesurées dans chaque zone d'abondance

La comparaison des distributions des distances au récif mesurées dans les différentes zones d'abondance vérifie ce que l'on observe sur la représentation visuelle des zones d'abondance du bec de cane (figure n° :) : dans le sud du lagon, les abondances augmentent à mesure que l'on s'éloigne du récif barrière. En effet, on observe sur les distributions :

- un glissement des modes qui traduit un gradient croissant des abondances sur l'axe récif-côte : situé à 600m du récif pour la zone "pauvre", il se situe à 11000m pour la zone "très abondant";
- un resserrement progressif des distributions autour de ces modes qui indique l'apparition d'un intervalle de distances au récif optimales:

$$\alpha_4 \text{ pauvre} < \alpha_4 \text{ moyen} < \alpha_4 \text{ abondant} < \alpha_4 \text{ très abondant}$$

Il apparaît donc que le bec de cane, dans sa zone favorable correspondant à des distances au récif comprises entre 300 et 35000m, soit plus abondant dans les zones situées entre 5000 et 25000m du récif barrière, avec un optimum observé à 11000m.

On constate qu'au delà de 25000m du récif, l'abondance du bec de cane diminue : on observe des modes secondaires à 30000m pour la zone "pauvre" et à 26000m pour la zone "moyen".

Distance à la côte :

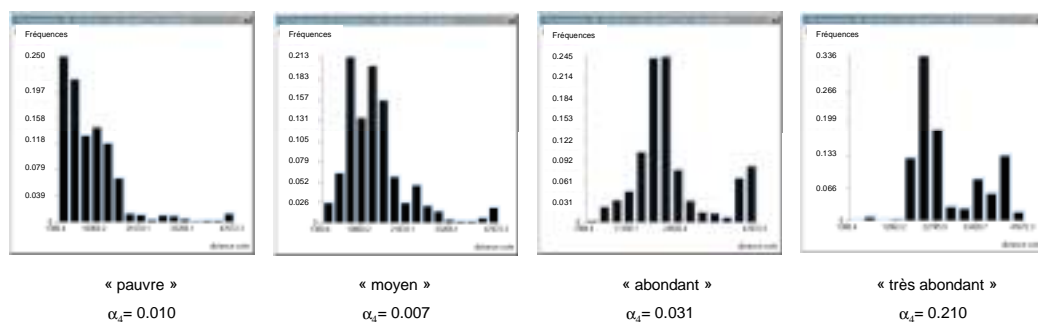


figure n° : histogrammes des fréquences des distances à la côte mesurées dans chaque zone d'abondance

La comparaison des distributions des distances à la côte mesurées dans les différentes zones d'abondance montre :

- un glissement progressif des modes vers des distances à la côte plus élevées (1400m pour la zone "pauvre", 19000m pour les zones "abondant" et "très abondant"). On observe également

l'apparition d'un mode secondaire au delà de 35000m pour les zones "abondant" et "très abondant";

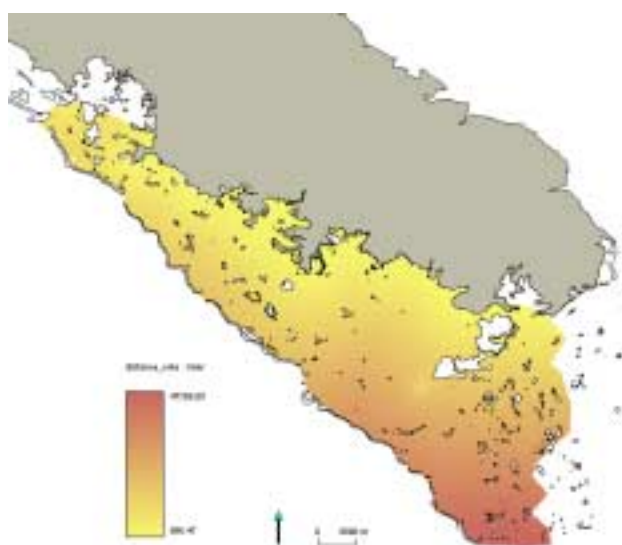
- un resserrement des distributions autour des modes : $\alpha_4 \text{ moyen} < \alpha_4 \text{ abondant} < \alpha_4 \text{ très abondant}$;

Il semble donc que le bec de cane, dans sa zone favorable (qui comprend toutes les distances à la côte possibles), soit plus abondant dans les zones situées entre 10000 et 48000m de la côte, avec un optimum observé à 19000m.

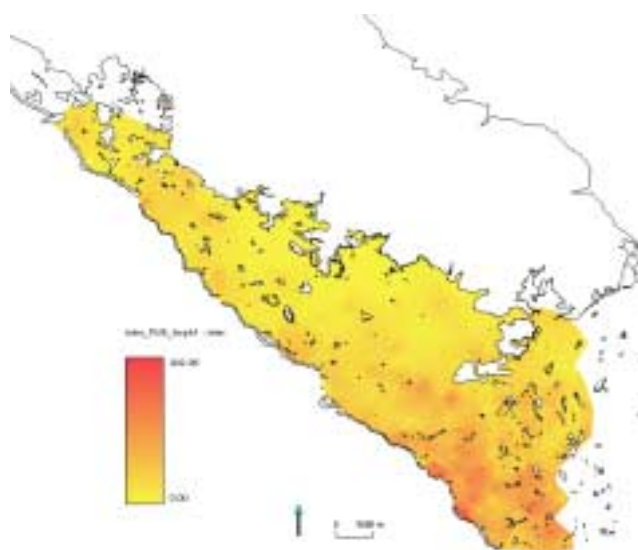
Annexe 8 : Répartitions spatiales de "PUE classe trophique 1" et "distance à la côte"

La variable "distance a la côte" a été retenue comme estimateur de l'importance de l'influence terrigène en un point. Notre démarche de recherche d'indicateurs biologiques de l'effet terrigène s'est donc basée sur la recherche des variables biologiques les plus corrélées à cet estimateur. La variable retenue est la PUE de la classe trophique 1.

Les répartitions spatiales de ces deux variables sont les suivantes :



distance à la côte



PUE classe trophique 1