Méthodologie statistique de sélection d'un indicateur d'usage pertinent et efficace à partir d'enquêtes sur la pêche plaisancière

Comparaison des résultats de la Réserve Naturelle Marine à la Réunion, la Réserve Naturelle Marine de Cerbère-Banyuls et du lagon sud-ouest de la Nouvelle-Calédonie.

GONSON Charles, FERRARIS Jocelyne,
THOMASSIN Aurélie, PAYROT Jérôme et JOLLIT Isabelle

[mai 2011]

Introduction **p.2**

I. MÉTHODOLOGIE **p.5**

A.Sélection des données p.6
B.Sélection des métriques p.7
C.Test de l'indicateur p.11

II. Cas de Cerbère-Banyuls **p.15**

A.Identification des thèmes p.16
B.Thème Pression p.16
C.Thème Pratique p.24
D.Thème Perception p.29
E.Thème Socio-économie p.35

III. Cas de la Nouvelle-Calédonie **p.40**

A.Identification des thèmes p.41
B.Thème Pression p.41
C.Thème Pratique p.48
D.Thème Impact p.54
E.Thème Socio-économie p.62

IV. Cas de la Réunion **p.68**

A.Identification des thèmes p.69
B.Thèmes Pression p.69
C.Thèmes Pratique p.73
D.Thèmes Perception p.78
E.Thèmes Socio-économie p.83

ANNEXES **p.88**

Introduction

Les Aires Marines Protégées (AMP) sont maintenant considérées comme les principaux outils de conservation de la biodiversité et de gestion des pêcheries. À travers le monde, le besoin d'évaluer l'efficacité de ces AMP est en constante augmentation. Cependant les manques de moyens et de personnels auxquels sont confrontés les gestionnaires peuvent être un obstacle à la réalisation de suivis réguliers nécessaires à cette évaluation et donc à l'achèvement des objectifs de gestion (Pomeroy et al. 2005). L'identification d'indicateurs peut permettre d'établir un diagnostic et d'évaluer les risques liés aux changements environnementaux et anthropiques à moindre coût. C'est pourquoi depuis la fin des années 1990, les indicateurs biologiques sont devenus un enjeu important pour les sciences environnementales (Garcia et al. 2000). C'est dans cette optique que les projet "**Indicateurs de la Performance d'Aires Marines Protégées pour la gestion des écosystèmes côtiers, des ressources et de leurs usAges - PAMPA** -" et "**Gouvernance des Aires marInes protégées pour la gestion durable de la biodiversité et des USages côtiers - GAIUS** -" visent proposer des méthodologies d'évaluation et de suivi de l'efficacité des AMP, par l'identification d'indicateurs associés autour de plusieurs cas d'études méditerranéens et d'outre mer.

Deux problèmes majeurs sont à la base des difficultés d'évaluation de l'efficacité des AMP. D'une part, les questions de recherche ont souvent été déconnectées des besoins réels des gestionnaires dans ce domaine. D'autre par, les données disponibles ne permettent pas la comparaison des conditions environnementales et bio-écologiques au sein des AMP avec les usages et la gestion qui y sont pratiqués. Le projet **PAMPA** vise à répondre à ces problèmes en a) instaurant un dialogue entre scientifiques et gestionnaires permettant de fournir un outil d'aide à la gestion applicable, b) en lançant une vaste campagne de collecte et de mise au format de données comprenant à la fois des informations sur les ressources et communautés associées et sur les caractéristiques des usagers et c) proposant des tableaux de bord d'indicateurs. Le projet **GAIUS** vise à repositionner la réflexion sur les indicateurs dans un cadre plus conceptuel et théorique, pluridisciplinaire, impliquant juristes, écologistes, économistes, géographes et modélisateurs.

Un indicateur est une variable qualitative ou quantitative qui peut être obtenue à partir de données de terrain ou de modèles. Il peut être relié directement à des objectifs de gestion

ou à des questions de recherche (Claudet et al. 2006, Pelletier et al. 2005, Amand et al. 2004) avec deux principales caractéristiques requises la **pertinence** et l'**efficacité**. On entend par pertinence le lien entre l'indicateur et le phénomène qu'il est censé représenter tandis que son efficacité fait référence à sa précision, sa sensibilité, sa variabilité et sa puissance statistique. Une grille de lecture associée à un indicateur permet de faciliter son interprétation ainsi que de le relier à des objectifs de gestion par l'identification de valeurs seuils. Cependant cette identification ne peut se faire qu'à partir de valeurs de référence issues de la bibliographie, de la connaissance d'experts ou par l'étude des variations spatiales et temporelles de l'indicateur à partir de données historiques. Cette étude porte sur la sélection et l'analyse d'indicateurs relatifs aux usages dans et autour des AMP.

Les activités touristiques liées au milieu marin (pêche, plongée, activités nautiques) sont devenues des sources de pression non négligeables pour la gestion des zones côtières. En particulier la pêche plaisancière au sein des AMP est souvent plus importante que la pêche professionnelle, que ce soit en nombre de bateaux, en nombre de pêcheurs et parfois même en biomasse prélevée (Lloret 2008). Cependant cette activité est très peu étudiée (Stelzenmuller et al. 2008). La pêche plaisancière rassemble plusieurs types d'activités (pêche à pied, à la ligne, sous-marine et embarquée) et rend sa définition difficilement explicite. La définition de la pêche plaisancière que l'on retiendra dans cette étude est l'activité de pêche dont le produit est destiné à la consommation exclusive du pêcheur et de ses proches sans pouvoir être colporté, exposé ou vendu (art. 1er du Décret 90-618 du 11 juillet 1990). De plus, on distinguera le pêcheur plaisancier du pêcheur vivrier par le fait que le pêcheur plaisancier pratique cette activité pour le plaisir qu'elle lui procure. Bien que les perceptions et les comportements des usagers soient pris en compte dans la plupart des objectifs de gestion des AMP (réduction des conflits, acceptation sociale de l'AMP, …) et que de nombreux auteurs aient présenté l'importance du rôle des usagers dans le succès des AMP (Himes 2007; Russ 1999), il s'agit d'un domaine de recherche encore peu étudié (Suman et al. 1999; Thomassin et al. 2010).

Cette étude a pour but la mise en place d'une méthodologie d'analyse statistique visant à sélectionner et décrire un panel d'indicateurs d'aide à la gestion de la pêche plaisancière. Trois points sont abordés : i) la sélection des données pour la réalisation de typologies des

usagers ; ii) la sélection de métriques[1] ; iii) le test d'indicateurs. La sélection des données présente, à partir des variables disponibles[2], celles qui sont analysées en vue de caractériser la pêche plaisancière. Cinq thèmes liés à cette activité sont explorés a) la "pression" de pêche, b) la "pratique" de l'activité, c) la "perception" de l'AMP par l'usager, d) la caractéristique "socio-économique" de l'usager et e) l' "impact" du pêcheur sur les ressources. À chacun de ces thèmes peuvent correspondre un ou plusieurs objectifs de gestion généraux ou inhérents à l'AMP. La sélection des métriques a pour but d'identifier des indicateurs potentiels à partir de typologies des usagers pour chacun des thèmes développés. La métrique discriminant le mieux ces groupes d'usagers est ainsi sélectionnée comme indicateur potentiel. Le test de l'indicateur consiste à étudier en priorité les variations spatiales et temporelles de l'indicateur à l'aide de modèles statistiques mais aussi à identifier les facteurs influençant ses variations.

La méthode mise en place dans ce travail vise, à être transférable entre sites d'étude. En effet les caractéristiques des AMP étant différentes (taille, date de création, géomorphologie) et les usages étant variables entre ces AMP (usages traditionnels, espèces cibles) l'utilisation d'indicateurs identiques à l'ensemble des AMP pourrait les rendre inadaptés autant du point de vue de leur pertinence que de leur efficacité. La méthodologie statistique proposée a été mise au point sur les données de la Réunion, puis a été optimisée sur les données de Cerbère-Banyuls et de Nouvelle-Calédonie afin de garantir son caractère transférable.

L'hypothèse de base de la démarche est que le groupe d'usagers échantillonnés, représentatifs de la population statistique inventoriées, est hétérogène du point de vue de la pression exercée sur l'AMP, de l'impact sur les ressources, de leurs pratiques, de leurs perceptions et de leurs caractéristiques socio-écononomiques (thèmes étudiés). Trouver un indicateur de l'effet de l'AMP sur les usagers nécessite donc de prendre en compte cette hétérogénéité. La sélection de l'indicateur le plus pertinent est au départ basé sur la sélection de la métrique la plus pertinente : cette métrique étant celle qui reflète le plus la variabilité des usagers. La métrique sera ensuite proposée en tant qu'indicateur efficace si et seulement si, cette métrique montre des différences significatives avec des facteurs d'intérêt dont on cherche à évaluer l'effet (figure n°1).

---

[1] On appelle « **métrique »**, une variable mesurable, avec donc une dimension, une unité de mesure et des bornes minimales et maximales.

[2] Les variables observées font référence aux protocoles standardisés proposés par le projet PAMPA et à la base de données PAMPA.

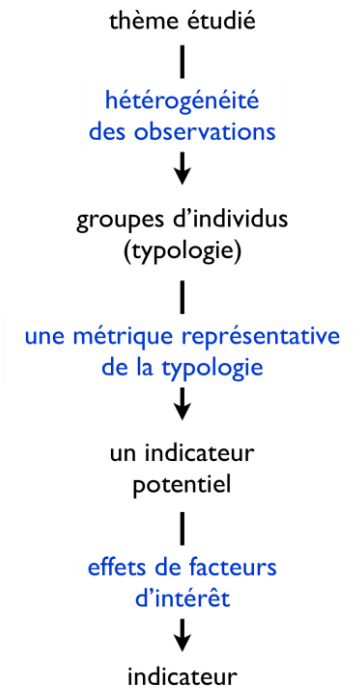


Figure n°1: schéma représentant la démarche (en noir) et les hypothèses sous-jacentes (en bleu) de la méthodologie statistique

I. MÉTHODOLOGIE

La démarche proposée part du principe qu'une structure de groupe peut être identifiée, pour chacun des thèmes présentés en introduction, par une approche typologique constituée de deux étapes. 1) une ordination (ACP ou ACM) permet d'explorer la structure des données (relation entre les variables, structure de groupes) et d'associer chacune des observations à des coordonnées quantitatives sur un plan factoriel. 2) une classification non hiérarchique (K-means) permet de former des groupes à partir des coordonnées des premiers axes factoriels afin de ne pas prendre en compte le « bruit blanc » propre à chaque jeu de données. L'hypothèse sous-jacente, dont dépend la pertinence de cette méthode, est qu'une variabilité suffisante pour l'identification de groupes distincts, existe au sein de l'échantillon pour un nombre suffisant de variables d'un même thème. Les analyses statistiques, dont l'ensemble de la démarche est résumé à l'annexe 1, ont été réalisées sur le logiciel R dans sa version 2.10.1. Sur l'annexe 1 est présenté la démarche pour les cas de la Réunion et de Cerbère-Banyuls. Le cas de la Nouvelle-Calédonie a nécessité une mise au format des données, comme pour le cas de la Réunion, une sélection des variables, une sélection des métrique et le test de l'indicateur, de la même façon que pour le cas de Cerbère Banyuls.

A. Sélection des données

Le choix des variables est crucial dans la formation de groupes puisque cela aura une influence sur toutes les étapes de la méthodologie, sur les résultats et sur les interprétations. Ce choix a un effet direct sur la pertinence de l'indicateur. En effet, la méthodologie permet de sélectionner parmi la liste de variables retenues, celles qui résument le mieux la structure de groupes d'un thème donnée. De ce fait chacune des variables sélectionnées doit être pertinente en temps qu'indicateur de l'AMP.

Pour chacune des variables, la présence de valeurs extrêmes pour les données de nature quantitative ou de modalités très peu représentées pour les données de nature qualitative, peut avoir un impact important dans la construction du plan factoriel et finalement sur la formation des groupes. Ainsi, afin de ne pas influencer l'analyse par la présence des individus « extrêmes », les variables présentant ces valeurs doivent répondre à certains critères. Ces critères sont bien évidemment différents selon que la variable soit qualitative ou quantitative:

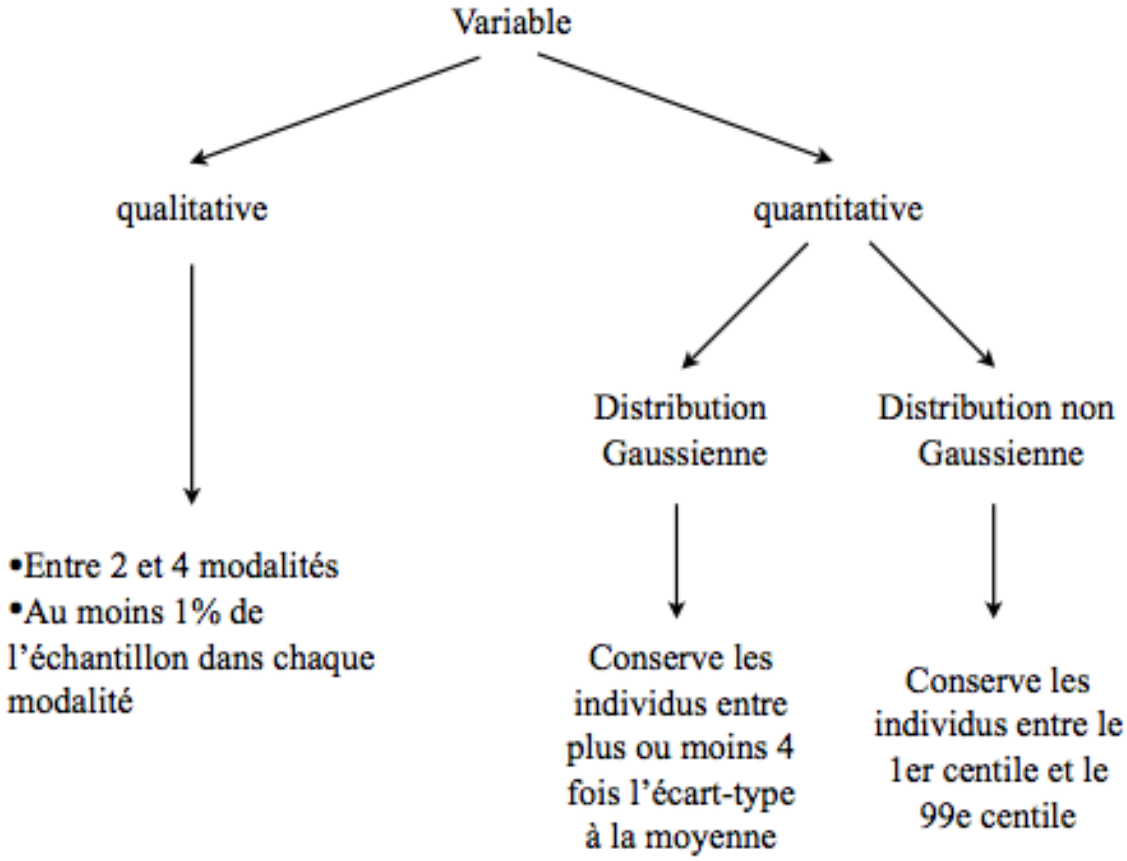


Figure n°I-1: Arbre de décision présentant les choix a effectuer pour la sélection des individus

Le respect ou non des critères aura des conséquences sur la sélection des individus de l'échantillon et/ou le recodage des variables (ex : regroupement des modalités). Les individus « extrêmes » pourront alors faire l'objet d'une analyse séparée en tant qu'éléments atypique de l'échantillon mais peut-être représentatif de comportement à prendre en compte.

Dans le cas où le nombre d'observations est inférieur à 100, on décide de conserver toutes les données, la présence de valeurs extrêmes ou de modalités peu représentées ayant une représentativité relative plus importante que dans des jeux de données de grande dimension.

B. Sélection des métriques[3]

Une typologie est effectuée pour chaque thème qui caractérise les usagers de la pêche plaisancière dans et autour de l'AMP. Une description des étapes et des choix à effectuer est présentée.

Ordination

Les méthodes d'ordination utilisées sont différentes selon que les variables soient quantitatives (ACP) ou qualitatives (ACM). Les deux méthodes ont le même but, il s'agit de réduire l'espace des variables en définissant des axes factoriels, qui concentrent la variabilité maximale et qui permettent de générer des coordonnées pluridimensionnelles pour chaque individu sous forme quantitative. Le choix du nombre d'axes factoriels, et donc le nombre de coordonnées pour chaque individu, revêt une importance capitale pour la suite de l'analyse. En effet le nombre d'axes sélectionnés influera sur le résultat de la typologie puisqu'il déterminera la position relative des individus dans l'espace factoriel. Il s'agit de choisir un nombre d'axes nécessaires et suffisants pour conserver l'information principale et écarter celle superflue (ou bruit). Étant donné le caractère générique de la méthode développée, on propose des règles générales pour le choix du nombre d'axes à conserver.

Le premier critère sur lequel on se base pour le choix du nombre d'axes est la corrélation des axes avec les variables de départ. Le deuxième critère concerne les valeurs propres des axes. La valeur propre d'un axe représente la part de variance des données qui est expliquée par cet axe. Le détail de ces règles est présenté dans le tableau n°I-1.

[3] On appelle « **métrique »**, une variable mesurable, avec donc une dimension, une unité de mesure et des bornes minimales et maximales.

Tableau n°I-1: Tableau présentant les règles sur lesquelles se basent le choix du nombre d'axes selon l'analyse d'ordination utilisée

Méthode d'ordination	Règle selon les corrélation	Règles selon les valeurs propres des axes
ACP	Corrélation entre les axes conservés et au moins une variable supérieure ou égale à 0.5	Nombre d'axes conservés présentant un pourcentage cumulé des valeurs propres supérieur à 66%
ACM		Méthode du coude: on conserve les axes précédents la première décroissance inférieure à 10%. Si le nombre d'observation est inférieur à 100 individus, on conserve les axes précédents la dernière décroissance supérieure à 10%.

Les méthodes de mise en évidence d'une structure de données, telles que les ordinations, posent le problème de la robustesse des résultats vis à vis de l'effort d'échantillonnage. On propose donc d'étudier l'évolution des corrélations entre les variables pour l'ACP et entre les modalités pour l'ACM, en fonction de la taille d'échantillon considérée. Cette étape est effectuée dans le but de détecter une taille d'échantillon minimale qui garantisse la pertinence de la structure mise en évidence par l'analyse factorielle.

Pour ce faire, on commence par créer des sous-échantillons de tailles croissantes issus de tirages aléatoires sans remises de l'échantillon total. Il est important que pour les variables qualitatives chacune des modalités soit représentée par un nombre suffisant d'individus. Par la suite et sur chacun de ces échantillons, la méthode d'ordination correspondante est effectuée. Enfin on calcule l'indice Δi avec la formule:

$$\Delta_i = \frac{1}{(P-1)!}\sum_{j}^{P}\sum_{k}^{P-1}\left|\delta_{jk}\right|$$

avec, P= le nombre de variables ou modalités

∂_{jk}= différence entre coefficient de corrélation de la variable ou modalité j du sous échantillon avec la variable ou modalité k de l'échantillon total

L'interprétation de la robustesse porte principalement sur la stabilité de Δi vis à vis de l'effort d'échantillonnage. Plus Δi sera faible, plus la structure (corrélations des variables et projections des individus) des plans factoriels sera proche. Cette étape permettra de déterminer un effort d'échantillonnage minimal à respecter par le gestionnaire pour l'utilisation de cette méthode.

Le tableau n°1 fournit des règles de décision pour réduire la dimension des variables tandis que Δi permet de déterminer la dimension minimale de l'échantillon (nombre des

individus) pour conserver la structure de groupe présente au sein de l'échantillon. Il s'agit dans les deux cas de réduire la dimension des données tout en conservant l'information pertinente à la formation de la typologie.

Classification - partitionnement

À partir des coordonnées générées par l'ordination, on cherche à former des groupes d'individus homogènes. Pour ce faire on utilise une méthode de partitionnement non hiérarchique des k-means qui vise à classer chacune des *n* observations en *k* groupes, de manière à ce que la variabilité (somme des carrés intragroupes) soit la plus faible possible. Chacune des *n* observations est attribuée au groupe dont le centroïde (moyenne multivariée) est le plus proche.

Afin de garantir le minimum global de la somme des carrés intragroupes, et de ne pas proposer une partition dépendant uniquement des points de départ, on réitère la partition en changeant le point de départ des centroïdes des classes. La méthode des k-means est une méthode objective en ce sens qu'elle minimise la valeur d'un certain critère numérique. Cependant, le choix du nombre de groupes est important, et aura une répercussion sur la suite de l'analyse. Afin de ne pas laisser la subjectivité de l'utilisateur inférer dans la méthode, on propose de se baser sur le critère de Calinski-Harabasz (1974). Ce critère (Ca) s'écrit sous la forme:

$$Ca = \frac{(SSB/(k-1))}{(SSW/(n-k))}$$

avec,
k = nombre de groupes
n = nombre d'individus
SSB = somme des carrés inter-groupe
SSW = somme des carrés intra-groupe

Le critère de Calinski est donc un rapport de la variance intergroupe sur la variance intragroupe. Autrement dit plus ce critère est élevé, plus les individus sont homogènes à l'intérieur de chaque groupe et plus les groupes seront différents. On choisit de se donner la possibilité d'avoir entre 2 et 5 groupes. On limite le nombre de groupes à 5 pour faciliter la lecture des résultats. Le critère de Calinski-Harabasz est calculé pour chaque nombre de groupes. On conserve le nombre de groupes pour lequel ce critère est le plus élevé.

Interprétation de la structure

Une fois les groupes formés il est important de savoir quelles sont les particularités de chacun de ceux-ci et de se rendre compte si les différences observées (entre deux groupes ou entre un groupe et l'échantillon total) sont significatives. Les analyses utilisées pour tester ces différences dépendent de la nature des données et de leur distribution (cf. Tableau n°I-2).

Tableau n°I-2: tests statistiques utilisés en fonction de la nature et de la distribution des données

Nature des données	Distribution	Tests associés
Quantitatives	Normale	Student (2 groupes) ou ANOVA (>2 groupes)
	Pas Normale	Mann-witney (2 groupes) ou Kruskal-wallis (>2 groupes)
Qualitatives	X	Chi-2

Afin de rendre compte des particularités de chacun des groupes et de l'échantillon total on fait un tableau récapitulatif des paramètres statistiques et des résultats des tests statistiques.

Tableau n°I-3: Tableau récapitulatif des statistiques et des résultats des tests entre groupes

			Statistiques			
Groupes	**Variables**	**Nombre d'individus**	Quantitative	Qualitative	**G1**	**G2**
Groupe 1	V1	n_1	Paramètres de position: moyenne, médiane. Paramètres de dispersion: écart-type, coefficient de variation.	Nom de la modalité principale et pourcentage de celle-ci dans le groupe		
	V2					
Groupe 2	V1	n_2				
	V2					

Dans le tableau I-3, G1 et G2 correspondent respectivement au groupe 1 et au groupe 2, les "n" sont le nombre d'individus dans chaque groupe, tandis que V1 et V2 correspondent à deux variables d'un même thème. Les cases noires indiquent que la variable considérée présente une différence significative pour un risque alpha donné entre deux groupes (e.g. La variable 1 est significativement différente entre le groupe 1 et le groupe 2).

Une fois obtenu *k* groupes dans un thème donné, on cherche à hiérarchiser les variables du thème qui discriminent le mieux ces groupes. On réalise alors un arbre de reclassification (arbre de régression dont les groupes sont connus) monothétique. Le but de l'arbre de segmentation est de prédire l'appartenance de chacun des individus à un groupe en fonction des variables ayant servi à la construction de la typologie considérée. Il s'agit d'une segmentation répétée sur l'échantillon total où la formation d'une nouvelle branche

segmentant l'échantillon en deux dépend d'une seule variable à laquelle est associée une valeur seuil. Le choix de la variable et de la valeur seuil se fait de manière à maximiser la pureté des groupes formés pour la variable considérée (variable que l'on cherche à prédire, ici les groupes de chaque thème). Autrement dit la formation de nouvelles branches diminue au maximum le rapport variance intragroupe sur la variance intergroupe. Si les variables utilisées pour la segmentation sont de nature qualitative nominale, la formation de nouvelles branches n'est pas associée à des valeurs seuils mais à un ensemble de modalités. On utilise cette méthode avec les variables ayant servi à la formation de la typologie en tant que variables explicatives et les groupes issus du thème en tant que variable qualitative à expliquer (appartenance ou non à une classe). Ainsi la variable servant au premier niveau de la segmentation est celle qui discrimine le mieux les groupes issus de la partition non hiérarchique. Il s'agit dans cette étape d'identifier une hiérarchie dans les variables d'un même thème et donc d'aider au choix d'une métrique pour la construction d'un indicateur. Le caractère monothétique de cette méthode (une seule variable pour chaque partition) permet d'identifier des combinaisons de variables qui discriminent le mieux les groupes d'un même thème. Cependant cette combinaison de variables est forcément de nature qualitative.

La formation de l'arbre de reclassification se fait en deux étapes. La première consiste à construire l'arbre complet en minimisant le plus possible les erreurs de classification. Ensuite, il s'agit de raccourcir ce dernier de manière à ce que le R^2 (prédiction de l'arbre) n'augmente plus de manière significative avec l'ajout de nouveaux noeuds terminaux. De cette manière l'arbre est le plus simple possible tout
en minimisant les erreurs de classification.

C. Test de l'indicateur

Cette étape vise à mettre en évidence, les relations entre des variables d'intérêt et la métrique sélectionnée pour justifier les groupes d'individus. Les effets de ces variables d'intérêt permettront a) de tester l'influence de l'AMP sur l'activité de pêche b) de faire des recommandations sur la stratégie d'échantillonnage a adopter. L'hypothèse de départ est que la variabilité de la population d'usagers, décrite sous un angle de vue (un des 4 thèmes du questionnaire PAMPA), est le reflet de l'influence des variables d'intérêt telle que l'effet AMP (mise en place de mesures de régulation, augmentation de la biomasse dans l'AMP, …).

À partir de la typologie de chaque thème et grâce à la méthode de segmentation, on a identifié une ou plusieurs métriques qui discriminent le mieux les groupes issus des

typologies de chaque thème. Ces métriques se présentent sous une forme quantitative (thème pression) ou qualitative (thèmes pratique et perception). Elles peuvent correspondre soit à une variable issue des observations sur le terrain, soit à une combinaison de variables selon l'arbre de segmentation. Dans le cas où une métrique seule se révèle être un mauvais indicateur, on utilisera la combinaison de plusieurs variables. Plus le nombre de variables à combiner sera important, plus la métrique se rapprochera des groupes de typologie eux-mêmes (cas où la structure de groupes est floue)

De nombreuses variables (externe à la typologie) peuvent expliquer la variabilité des données et donc avoir une influence sur une métrique donnée. Le choix de facteurs explicatifs doit se faire de telle façon à éviter les redondances d'une part, et la présence dans le modèle statistiques de variables "inutiles" d'autre part. On entend par "inutile", une variable qui n'a pas d'effet sur la métrique. On teste alors l'effet de chacune de ces variables séparément sur la métrique à l'aide de tests statistiques (cf. Tableau n°I-2). L'ensemble des variables ayant un effet significatif sur la métrique, est injecté dans le modèle complet. Ces variables d'intérêt peuvent être regroupées en catégories. Ces catégories regroupent : i) les variables temporelles, relatives au moment de pêche, ii) les variables spatiales, indiquant les lieux de pêche, iii) les variables d'impact, iv) les variables liées à l'activité de pêche, v) les variables météorologiques et vi) les typologies construites par la méthode présentée en amont.

Tableau n°I-4: catégories, niveau de détail et présence dans les jeux de données des variables d'intérêt

Catégories	+ détaillées → - détaillées	Nouvelle-Calédonie	Banyuls
Temporelle	Mois → Saison	X	X
	Année		X
	Type de jour	X	X
Spatiale	Zone → Statut PAMPA → Statut	X	X
	Part de pêche dans l'AMP		X
Impact	Catégorie de capture		X
Activité	Catégorie de bateau	X	X
	Type de bateau	X	X
	Ancienneté		X
	Activité de pêche habituelle	X	X
	Engin utilisé → groupe d'engins		X
Météorologie	Météo		X
	État de la mer		X
	Direction du vent		X
	Force du vent		X
Typologie	Pression	X	X
	Pratique	X	X
	Impact	X	
	Perception		X
	Socio-Économie	X	X

Le facteur « statut PAMPA » correspond à un zonage distinguant les zones internes et externes à l'AMP et qui prend en compte les zones d'influence de chacune de ces zones.

Selon la nature de la métrique, le modèle statistique utilisé sera différent. En effet pour une métrique de nature quantitative, on utilisera un modèle linéaire général, tandis que pour une variable de nature qualitative, on utilisera une régression logistique multinomiale.

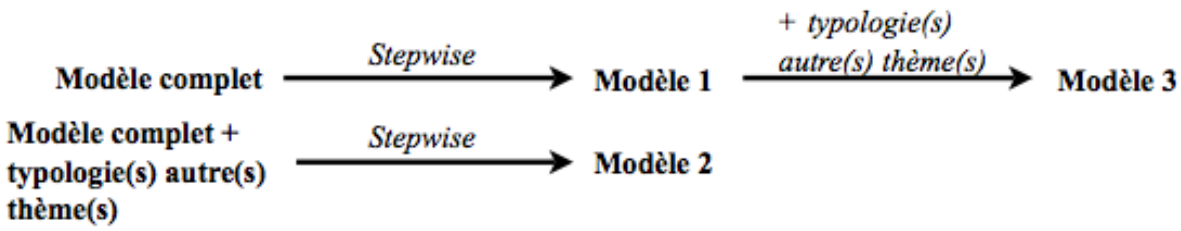


Figure n°I-2: Schéma présentant la méthode de construction des modèles

On commence par construire deux modèles complets, l'un avec toutes les variables d'intérêt qui ont un effet sur la métrique sans prendre en compte les typologies, et un autre en y ajoutant les typologies. À partir de ces modèles complets est appliquée une méthode de stepwise en se basant sur le critère d'information d'Akaike (AIC, 1971). Cette méthode consiste, à partir du modèle complet, à calculer l'AIC de tous les P-1 sous-modèles (avec P le nombre de facteurs) et de choisir le sous modèle pour lequel l'AIC est le plus faible. Il s'agit d'une méthode itérative se déroulant en étapes successives:

1. Calcul de l'AIC du modèle
2. Calcul de l'AIC pour chacun des P-1 sous-modèle
3. retrait de la variable dont le retrait du modèle diminue le plus l'AIC
4. Sous-modèle sélectionné et retour à l'étape 1

La méthode s'arrête une fois que le retrait d'un facteur ne diminue plus le critère d'AIC. Le modèle final contient uniquement des facteurs qui apportent une réelle information dans la prédiction de la variable à expliquer. Le critère d'AIC s'écrit sous la formule:

$$AIC = 2P + n\ [\ ln(SCR)]$$

Où P correspond au nombre de facteurs explicatifs, n au nombre d'observations et SCR à la somme des carrés des résidus du modèle.

Certaines conditions doivent être respectées pour qu'un modèle soit fiable. En effet les résidus doivent être:

- distribués selon la loi Normale

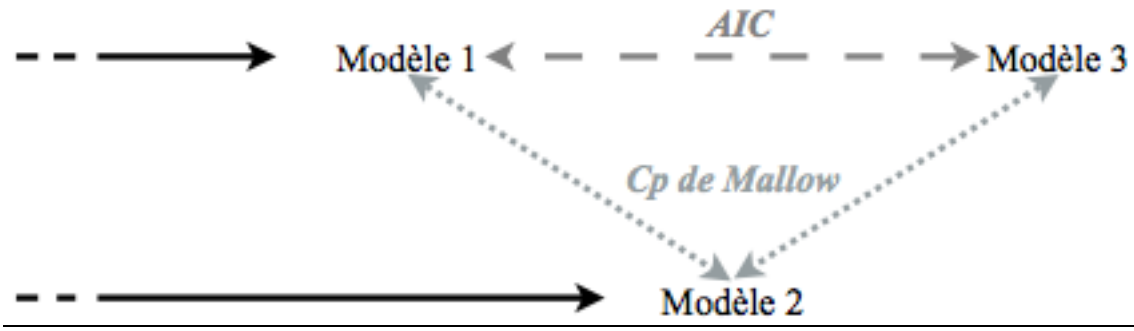
- indépendants des valeurs prédites (la corrélation entre les résidus et les valeurs prédites du modèle doit être proche de zéro)
- homoscédastiques

À la fin de la méthode, on obtient 3 modèles. Pour identifier le meilleur modèle entre le modèle 1 et le modèle 3, on utilise le critère d'AIC (puisque le modèle 1 est un sous-modèle de dimension P-1 du modèle 3). Pour identifier le meilleur modèle entre le modèle 2 et les modèles 1 et 3 on utilisera le Cp de Mallow (cf. Fig. n°2). Le Cp de Mallow, contrairement à l'AIC, est un critère statistique permettant de comparer des modèles ayant des facteurs explicatifs différents. Il dépend de la somme des carrés des résidus, du nombre de facteurs explicatifs et de la variance des observations:

$$Cp = SCR + 2 \times \sigma^2 \times P$$

Où SCR correspond à la somme des carrés des résidus, et σ^2 est la variance des observations généralement estimée à partir de la variance des résidus du modèle (Ferraris et al. 2005).

Figure n°I-3: Schéma



prés

entant les tests de comparaison entre les modèles

Dans le cas où le modèle 1 est retenu, alors la typologie n'apporte aucune information pour expliquer la métrique. Les thèmes (métrique et typologie(s) en facteur) ne présentent alors pas de corrélation. Dans le cas où le modèle 2 est retenu, la/les typologie(s) en facteur(s) permet(tent) d'expliquer une partie de la variance de la métrique. Il peut exister alors une corrélation entre les thèmes considérés (métrique et typologie(s) en facteur). Dans le cas où le modèle 3 est sélectionné la typologie a un effet similaire que dans le modèle 2, cependant une corrélation avec au moins un autre facteur du modèle l'aurait éliminé du modèle lors de la phase de stepwise.

Une fois le modèle sélectionné, on porte notre attention au R^2 qui est une statistique qui rend compte de la part de variance expliquée par le modèle. Le modèle retenu n'a pas forcément le R^2 le plus élevé puisque cette statistique présente le désavantage (contrairement

à l'AIC et au Cp de Mallow) d'être dépendant du nombre de facteurs. De plus pour ce modèle on réalise une ANOVA de type 3 afin de tester l'effet de chacun des facteurs dans le modèle.

Enfin on calcule la puissance du test qui nous renseignera sur la probabilité de rejeter l'hypothèse nulle à raison (1-β). De ce fait, que l'effet soit détecté ou non, si la puissance du test est élevée, le résultat sera fiable. La puissance d'un test dépend de a) la taille de l'échantillon (nombre d'observations par modalités), b) la taille de l'effet, c) la valeur seuil de rejet de l'hypothèse nulle (α) et d) du nombre de modalités du facteur considéré (dans le cas d'une ANOVA). Compte tenu de la nature du problème, on considère que la puissance du test est acceptable à partir de 0,8 (S. Champely, 2006). Ce dernier propose d'évaluer la taille de l'effet d'un facteur par la formule:

$$f = \sqrt{\frac{R^2}{1-R^2}}$$ R^2: part de variance expliquée par le facteur considéré

La suite de ce document présente l'application de la méthode présentée aux cas d'étude des réserves naturelle marine de Cerbère-Banyuls, de la Réunion et de la Nouvelle-Calédonie et les principaux résultats obtenus. Cependant, la partie de la méthode « test de l'indicateur » concernera uniquement le thème pression pour les cas de Cerbère-Banyuls et de la Nouvelle-Calédonie et le thème impact du cas d'étude de la Nouvelle-Calédonie.

II. Cas de Cerbère-Banyuls

La réserve naturelle marine de Cerbère-Banyuls (RNMCB) est l'unique réserve exclusivement marine française et a été créée en 1974. Elle s'étend sur 6,5 km entre Banyuls-sur-mer et Cerbère. Elle comprend une zone de protection partielle faisant 650 ha dont 65 ha en zone de protection intégrale.

Les données disponibles au format PAMPA pour la réserve de Cerbère-Banyuls comprennent un nombre d'unités d'échantillonnage s'élevant à 933 pêcheurs plaisanciers. La période de collecte de données a débuté en mars 2008 et fini en février 2010. Les données ont été obtenues à partir de questionnaires réalisés en mer et aux bureaux de la reserve marine.

A. Identification des thèmes

La sélection des données (nombre de variables et nombre d'individus) est présentée pour chaque thème dans le tableau n°II-1. Le détail présentant les variables de chaque thème est présenté en annexe 2.

Tableau n°II-1: Tableau récapitulatif du nombre d'individus et du nombre de variables sélectionnées pour la construction du plan factoriel de chaque thème

	Cerbère-Banyuls	
thèmes	Nombre d'individus	Nombre de variables
Pression	729	4
Pratique	188	5
Perception	352	10
Socio-économie	344	4

B. Thème Pression

Étant donné la présence de valeurs extrêmes pour une ou plusieurs des quatre variables du thème pression, 8 individus sont écartés pour la construction du plan factoriel de ce thème.

Typologie

Lors de la construction du plan de l'ACP pour le thème pression, les deux premiers axes rassemblent 90% de l'inertie du nuage de points. De plus ils permettent de garantir une réelle signification de ces axes avec des corrélations supérieures à 0,5 pour au moins une variable. Le premier axe factoriel est négativement corrélé à plus de 50% avec la totalité des variables. L'axe 2 est corrélé positivement et à plus de 50% avec le nombre de pêcheurs et le nombre d'engins tandis qu'il est corrélé négativement à presque 50% avec la durée de sortie et de pêche. L'axe 1 résume l'intensité de la pression de pêche tandis que l'axe 2 permet de qualifier cette pression (figure II-1).

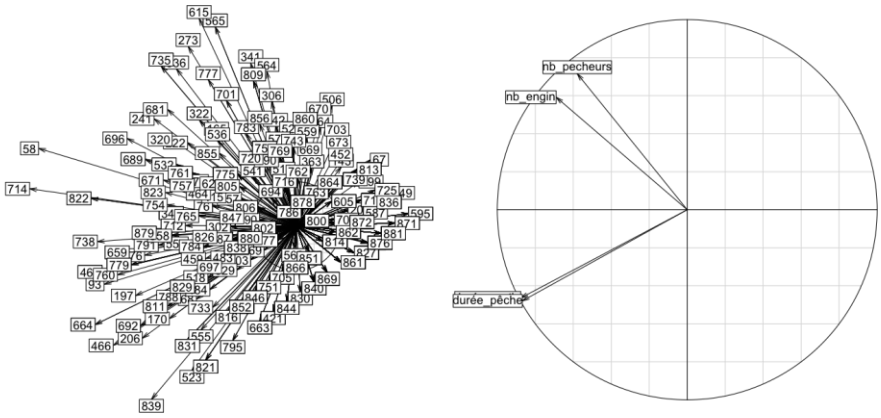


Figure n°II-1: projection des individus sur le plan (gauche) et cercle de corrélation (droite) de l'ACP du thème pression des données de la RNMCB

Sur la figure II-2 est représentée l'évolution des valeurs de l'indice Δi de robustesse des résultats en fonction du nombre d'individus de l'échantillon du thème pression. On observe une stabilité de la moyenne de Δi entre 250 et 700 individus. Il est donc nécessaire d'échantillonner au moins 250 pêcheurs pour avoir une représentation fiable des données sur le plan factoriel et garantir sa structure.

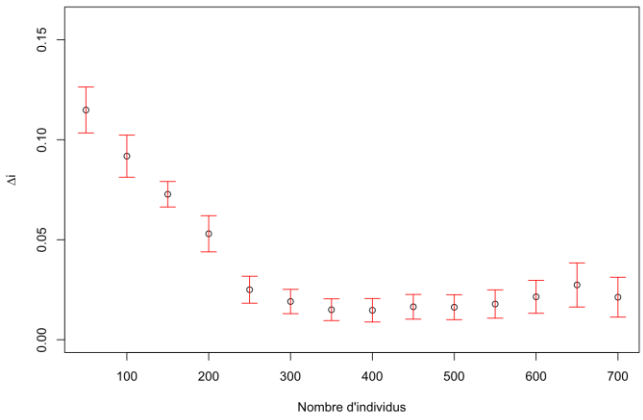


Figure n°II-2: Évolution de la différence moyenne des matrices de corrélation issues des ACP du thème pression de l'échantillon total et des sous échantillons tirés aléatoirement en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (590,83). On choisit donc de former 5 groupes de pression parmi l'échantillon de 729 individus.

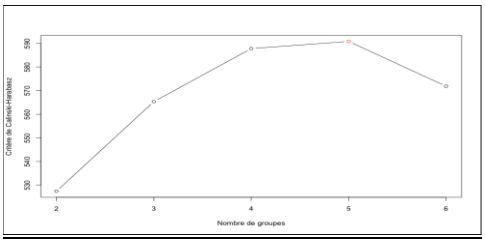


Figure n°II-3: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

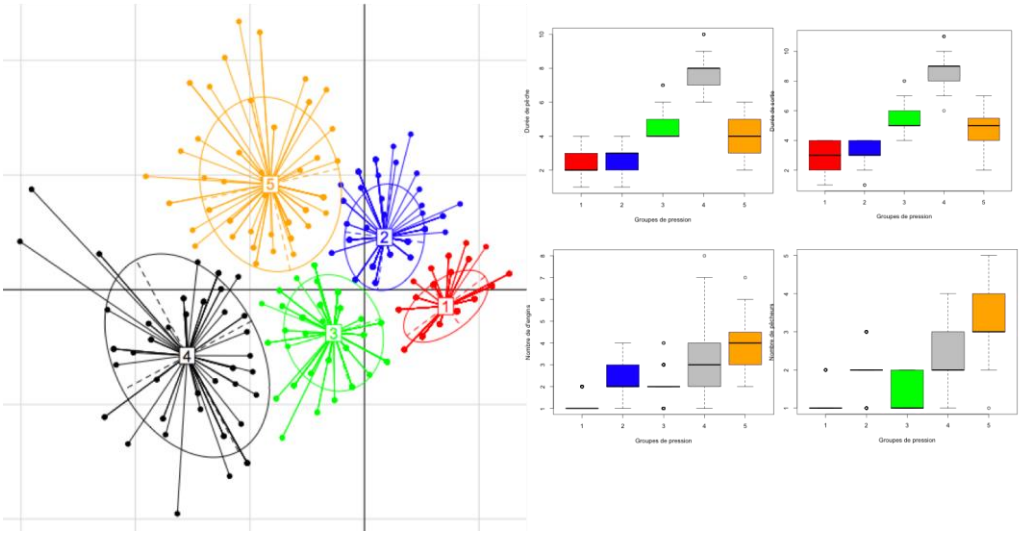


Figure n°II-4: Projection des individus par groupe de pression (points) sur le plan factoriel et représentation des diagrammes de dispersions associés pour chacune des variables du même thème

Le groupe 1 rassemble les individus qui ont la plus faible intensité de pression de pêche (figure II-4). Les groupes 4 et 5 réunissent les individus qui ont l'intensité de pression de pêche la plus élevée respectivement pour la durée de pêche et le nombre de pêcheurs. Les groupes 2 et 3 comptent dans leurs effectifs des individus qui ont une intensité de pression de pêche intermédiaire.

Tableau n°II-2: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de pression

	Description de l'échantillon et des groupes						Groupes				
Groupes	***Nombre d'individus***	***Variables***	***moy***	***med***	***sd***	***C.V.***	***1***	***2***	***3***	***4***	***5***
Échantillon	***729***	***Nombre de pêcheurs***	1.65	1	0.84	50.54					
		Nombre d'engins	2.10	2	1.09	51.83					
		Durée de sortie	4.30	4	1.91	44.52					
		Durée de pêche	3.56	3	1.85	52.14					
Groupe 1: "Nombre de pêcheurs et durée de sortie faible"	***262***	***Nombre de pêcheurs***	**1.04**	**1**	0.20	19.28					
		Nombre d'engins	**1.24**	**1**	0.43	34.60					
		Durée de sortie	**2.96**	**3**	0.81	27.22					
		Durée de pêche	**2.29**	**2**	0.78	34.17					
Groupe 2: "Durée de sortie faible"	***149***	***Nombre de pêcheurs***	**2.05**	**2**	0.5	24.32					
		Nombre d'engins	**2.42**	2	0.61	25.05					
		Durée de sortie	**3.22**	**3**	0.78	24.17					
		Durée de pêche	**2.48**	3	0.72	29.16					
Groupe 3: "Nombre de pêcheurs faible et durée de sortie élevée"	***174***	***Nombre de pêcheurs***	**1.41**	1	0.49	35					
		Nombre d'engins	1.97	2	0.68	34.68					
		Durée de sortie	**5.47**	**5**	0.80	14.68					
		Durée de pêche	**4.69**	**4**	0.82	17.40					
Groupe 4: "Durée de sortie élevée"	***65***	***Nombre de pêcheurs***	**2.05**	**2**	0.79	38.37					
		Nombre d'engins	**2.88**	**3**	1.13	39.44					
		Durée de sortie	**8.64**	**9**	1.21	14.04					
		Durée de pêche	**7.77**	**8**	1.23	15.85					
Groupe 5: "Nombre de pêcheurs élevé"	***79***	***Nombre de pêcheurs***	**3.15**	**3**	0.87	27.71					
		Nombre d'engins	**3.94**	**4**	0.99	25.02					
		Durée de sortie	**4.70**	**5**	1.17	24.97					
		Durée de pêche	**3.88**	**4**	1.12	28.85					

La totalité des groupes présente, entre eux et avec l'échantillon total, des différences significatives pour toutes les variables ou presque (tableau II-2). La structure de groupes mise en évidence est donc justifiée. Néanmoins, d'une part les groupes 2 et 4 ne présentent pas un nombre de pêcheurs significativement différent et d'autre part le groupe 3 et l'échantillon total n'ont pas un nombre d'engins significativement différent. Les libellés des groupes sont rajoutés à la vue des résultats en se basant sur les deux variables les plus indépendantes (durée de sortie et nombre de pêcheurs).

Sélection des métriques

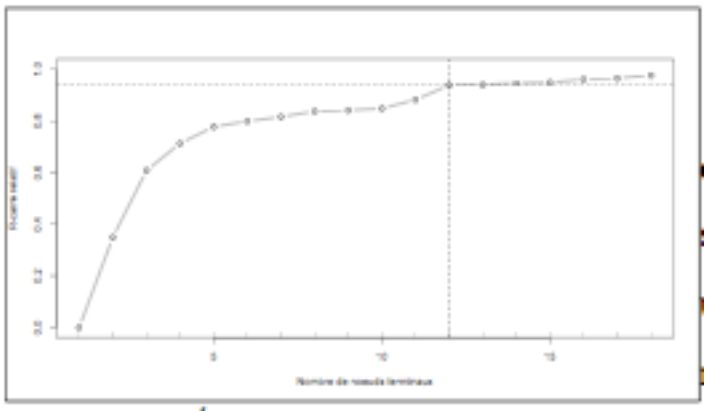


Figure n°II-5: Évolution de la différence moyenne des matrices de corrélation issues des ACP du thème pression de l'échantillon total et des sous échantillons en fonction du nombre d'individus

On n'observe qu'une très légère augmentation du R^2 relatif pour un nombre de noeuds terminaux supérieur à 12. On choisit donc de conserver uniquement 12 noeuds terminaux pour l'arbre de reclassification du thème pression.

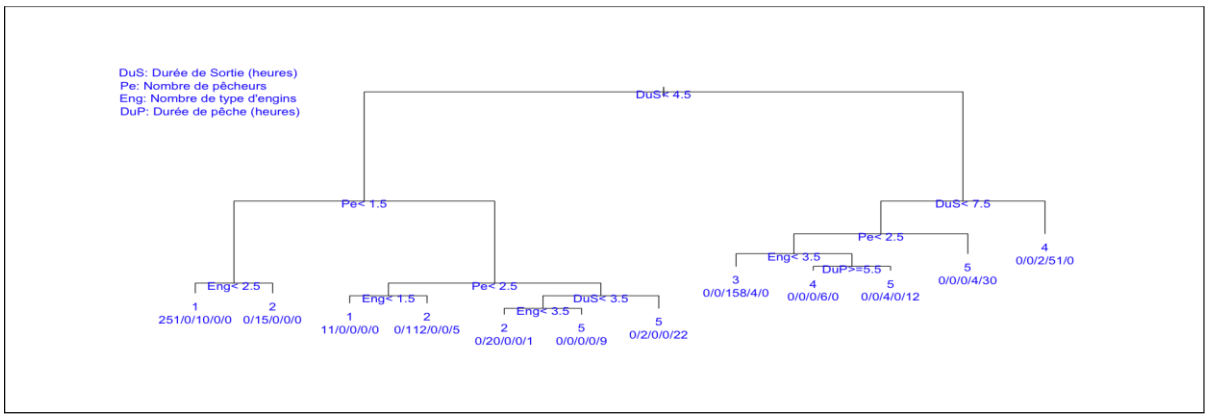


Figure n°II-6: Arbre de reclassification pour les groupes de pression en fonction des variables du même thème

Sur la figure II-6, sont indiquées, au niveau des noeuds de l'arbre, les variables et valeurs seuils associées pour la partition de l'échantillon tandis qu'au niveau des noeuds terminaux est représenté le numéro de groupe principal associé. En dessous de ce numéro est indiqué le nombre d'individus de chaque groupe qui compose le noeud terminal (nombre d'individus du groupe n°1 / nombre d'individus du groupe n°2 / … / nombre d'individus du groupe n°5). Cette indication permet de nuancer les résultats de la segmentation pour des noeuds terminaux dont la composition est diverse.

Les résultats obtenus sont cohérents avec le descriptif des groupes dans le tableau n°II-2. Le groupe 1 présente les valeurs les plus faibles pour chacune des variables, tandis que les groupes 4 et 5 présentent les valeurs les plus élevées respectivement pour la durée de sortie et le nombre de pêcheurs/engins. La durée de sortie est la variable qui discrimine le mieux les groupes de pression (figure II-6). Il s'agit donc de la première métrique à tester comme indicateur potentiel d'un effet de variables d'intérêt sur l'existence de groupes de pression. Si il est nécessaire de considérer une combinaison de variables, on choisira la combinaison des variables "durée de pêche" et "nombre de pêcheurs", en prenant comme borne des classes les premières valeurs seuils générées par l'arbre de reclassification.

Mise en évidence d'effet

Dans cette section, sont présentés les résultats obtenus pour le thème pression en choisissant comme indicateur la durée de sortie. Avec l'utilisation de tests non paramétriques (tableau I-2), on identifie les variables d'intérêt parmi le tableau n°I.4 qui sont corrélé avec la durée de sortie. Hormis la "part de pêche dans l'AMP", l' "ancienneté" et la "direction du vent", toutes les variables ont un effet significatif. Cependant le facteur "année" n'est pas conservé car le plan d'échantillonnage ne s'est pas effectué durant les mêmes mois entre les années et puisque le mois a un effet sur la durée de sortie il est impossible d'intégrer ce facteur au modèle. La typologie de pression n'est pas non plus conservée pour la construction du modèle puisqu'elle est directement dépendante de l'indicateur considéré (la durée de sortie). Enfin les facteurs « statut PAMPA » et « statut » ont été préférés au facteur « zone » car il est préférable que le niveau de détail des facteurs explicatifs soit le plus faible possible. De plus la totalité des modèles construits avec ce facteur ont des résidus qui ne sont pas distribués selon la loi Normale. On construit donc le modèle complet avec la totalité des facteurs explicatifs du tableau n°4, hormis les variables "zone", "année", "pression", "part de pêche dans l'AMP", "ancienneté" et "direction du vent".

La variable à expliquer est le log de la durée de sortie à laquelle on a ajouté 1. Cette transformation en log +1 est effectuée afin que les résidus des modèles suivent une loi Normale puisque sans cette transformation, la totalité des modèles est rejetée.

Tableau n°II-3: détails des facteurs sélectionnés et des paramètres statistiques de chacun des modèles construits

Modèles	Facteurs explicatifs	Paramètres	
Modèle 1	**Mer**; **Force du vent**; **mer**; **Catégorie de bateau**; **Type de bateau**; **Engin**; **Mois**; Statut PAMPA	Nombre d'individus	621
		Erreur résiduel standard	0,3153
		R^2	0,2185
		AIC	-1407,212
		Cp de Mallow	56,8187
Modèle 2	Type de jour; Type de bateau; **Activité de pêche**; Météo; **Mois**; Statut PAMPA; **Typologie de pratique**	Nombre d'individus	163
		Erreur résiduel standard	0,3249
		R^2	0,3846
		AIC	-323,4746
		Cp de Mallow	26,4057
Modèle 3	Mer; Force du vent; mer; Catégorie de bateau; **Type de bateau**; Engin; **Mois**; Statut PAMPA; **typologie de pratique**	Nombre d'individus	163
		Erreur résiduel standard	0,3317
		R^2	0,3782
		AIC	-352,1909
		Cp de Mallow	30,1432

Les effets des facteurs sont mis en évidence avec une ANOVA de type III. Les **facteurs en gras** ont un effet significatif en dessous du seuil de 0,05, les facteurs soulignés ont un effet significatif en dessous du seuil de 0,1.

Le modèle 1 est rejeté car, contrairement aux modèles 2 et 3, la distribution des résidus ne suit pas une loi Normale. Le modèle 2 a le Cp de Mallow le plus faible et est donc le modèle conservé. On constate aussi que pour ce modèle le R^2 est plus élevé et l'erreur résiduelle standard plus faible par rapport au modèle 3. Enfin on constate que l'effet spatial n'apparaît que dans le modèle 2. Aucune interaction n'a pu être mise en évidence entre les mois et le statut PAMPA pour ces trois modèles. L'effet du statut PAMPA sur la durée de sortie est donc le même quel que soit le mois. L'injection de la typologie de pratique dans les facteurs explicatifs permet de valider les conditions d'utilisation du modèle en normalisant la distribution des résidus. De plus cette typologie augmente le R^2 et diminue le Cp de Mallow dans le modèle 3 par rapport au modèle 1.

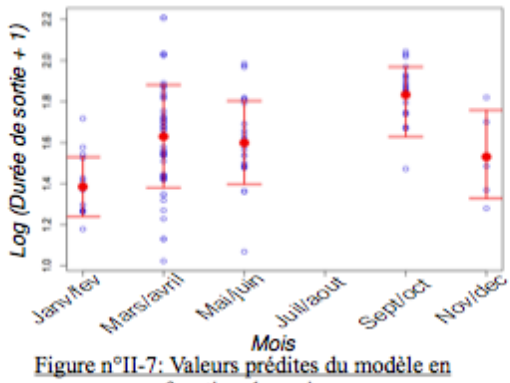


Figure n°II-7: Valeurs prédites du modèle en fonction du mois

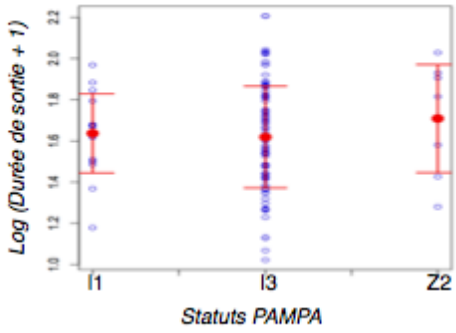


Figure n°II-8: Valeurs prédites du modèle en fonction du statut PAMPA

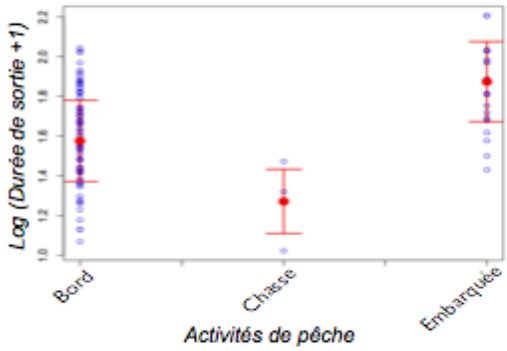


Figure n°II-9: Valeurs prédites du modèle en fonction de l'activité de pêche

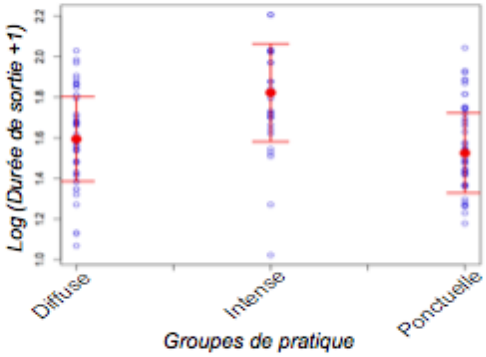


Figure n°II-10: Valeurs prédites du modèle en fonction des groupes de pratique

Les figures II-7 à II-10 représentent le log de la durée de sortie + 1 en fonction des facteurs qui ont un effet significatif (mois, statut PAMPA, activité de pêche et groupes de pratique). Les points bleus représentent chacune des observations tandis que les points rouges et les barres rouges représentent respectivement la moyenne et l'écart type pour chaque modalité.

Les mois de mars à octobre (figure II-7) ont les valeurs prédites en moyenne les plus élevées (de 4 à 5,25 heures). Néanmoins on remarque l'absence d'observations durant les mois de juillet et d'août, ce qui est dû à l'orientation du travail des agents de terrain vers la surveillance plutôt que vers le recueil de données. Les mois de novembre à février ont les valeurs en moyenne les plus faibles (de 3 à 3,6 heures). Enfin ces valeurs sont les plus variables pour les mois de mars et d'avril, avec un écart-type de 1,3 heures et les valeurs prédites les plus élevées (8 heures) ainsi que les plus faibles (1,8 heures).

L'activité correspondant aux valeurs prédites du modèle les plus élevées en moyenne (figure II-9), est la pêche embarquée (5,5 heures) tandis que la chasse est l'activité qui présente la moyenne la plus faible (2,5 heures).

Les usagers faisant partie du groupe de pratique "intense" (figure II-10) ont les valeurs

prédites qui en moyenne sont les plus élevées (5,2 heures). Au contraire, les pêcheurs du groupe de pratique “ponctuelle” présentent les valeurs prédites qui sont en moyenne les plus faibles (3,6 heures).

Les valeurs prédites du modèle, sur la figure n°8, montrent une moyenne très légèrement supérieure pour les zones “I1” (4,2 heures) et “Z2” (4,5 heures) par rapport à la zone “I3” (4 heures). De plus il y a une grande variation du nombre d’observations entre les zones I1 (n=17), Z2 (n=8) et I3 (n=138).

Les puissances statistiques des ANOVA de type III pour les facteurs activité de pêche et mois sont égales à 1. On peut donc se fier aux résultats du modèle et affirmer qu’il y a bien une différence significative de la durée de pêche en fonction du mois d’une part, et en fonction de l’activité de pêche d’autre part. Pour que la puissance statistique de cette ANOVA soit d’au moins 0,8 il suffit que le nombre d’observations par groupe de mois soit de 10 individus pour une même taille d’effet et un même niveau de significativité. En appliquant le même principe à l’activité de pêche, on obtient un nombre d’observations par activité de 9 individus au minimum pour que la puissance du test soit supérieure à 0.8. La puissance de l’ANOVA de type III pour le statut PAMPA, en considérant un seuil de significativité α de 0.1, est de 0,22. La faible puissance statistique du test ne permet pas de conclure sur une réelle différence de la durée de pêche entre les zones, ce qui est à la fois dû à la taille de l’effet statut PAMPA faible (8%) et au seuil de significativité α élevé (0.1). En considérant un seuil de significativité α, une taille d’échantillon et des zones identiques, une taille d’effet d’au moins 22% est nécessaire pour avoir une puissance statistique de 0,8 ou plus. C’est à dire que la durée de sortie dans la réserve (zone I1 et/ou Z2) doit être 22% plus importante que hors de la réserve, soit approximativement une moyenne de 6,1 heures dans la réserve contre 5 heures en moyenne hors de la réserve.

Conclusion

On a mis en évidence 5 groupes de pression parmi l’échantillon en fonction de la durée de pêche, de la durée de sortie, du nombre de pêcheur et du nombre d’engin. La durée de pêche est la variable qui reflète le mieux les groupes de pression. Elle s’avère être donc l’indicateur le plus pertinent de la période de pêche et du type d’activité. Il faudrait une différence de la durée de pêche dans et hors la réserve pour mettre en évidence un effet significatif du zonage. La taille de l’échantillon a été réduite à 163 personnes enquêtées, il est donc possible que l’effet de zonage aurait pu être mis en évidence en considérant la totalité de l’échantillon.

C. Thème Pratique

Pour les variables du thème pratique, le nombre de sorties est codé en 4 modalités en prenant comme bornes des classes, le premier et troisième quartile ainsi que la médiane. Les variables correspondant au planning de l'activité dans la journée et de la semaine on été codées en 2 modalités: "IND" dans le cas où il y a indifférence de cette planification ou "PAS IND" dans le cas contraire. La planification à l'année comporte 3 modalités: "année", "été", "hiver" selon que l'usager pêche toute l'année, uniquement en été ou uniquement en hiver.

Typologie

Pour l'ACM du thème pratique, la décroissance des valeurs propres entre le 2e et le 3e axe est de 7%, on conserve donc uniquement les deux premiers axes. Néanmoins même si ces 2 premiers axes ne sont pas corrélés à plus de 0,5 avec la totalité des variables, conserver des axes supplémentaires ne permet pas de répondre à la règle selon les corrélations.

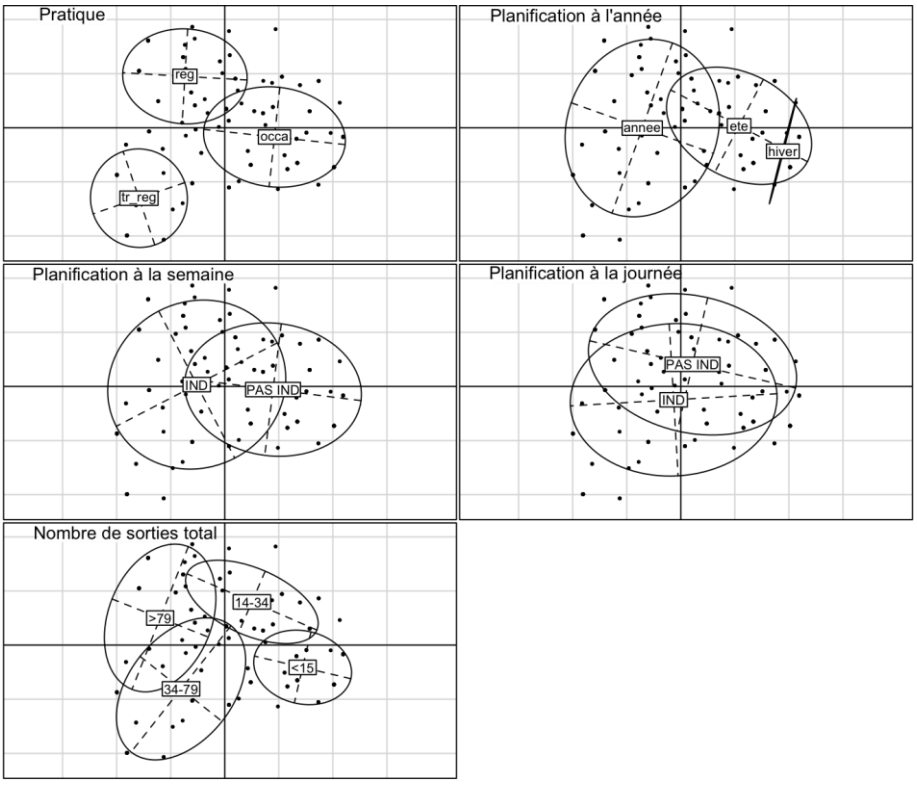


Figure n°II-11: Projection des individus et des modalités pour chaque variable de pratique sur le plan factoriel de l'ACM

Sur la figure II-11 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

De manière générale, la modalité de pratique "très régulière" s'apparie favorablement avec les modalités représentant un nombre de sorties à l'année important et une "indifférence" à la planification de l'activité de pêche. D'un autre côté, la modalité de pratique occasionnelle s'apparie favorablement avec les modalités correspondant à un faible nombre de sorties à l'année et une planification de l'activité "pas indifférente" au type de jour de la semaine et à l'heure de la journée. Ainsi les usagers pêchant beaucoup (pratique ou nombre de sorties) pratiquent l'activité, en général, quelle que soit la période de l'année, le type de jour ou l'heure de la journée.

Sur la figure II-12 est représentée l'évolution des valeurs de Δi en fonction du nombre d'individus de l'échantillon du thème pratique. On observe une stabilité de la moyenne de Δi entre 75 et 175 individus. Il est donc nécessaire d'échantillonner au moins 75 pêcheurs pour avoir une représentation fiable des individus sur le plan factoriel et garantir sa structure.

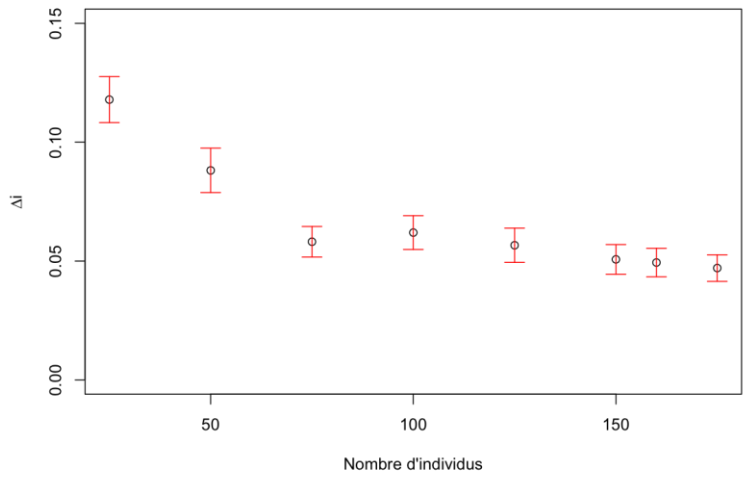


Figure n°II-12: Évolution de la différence moyenne des matrices de corrélation issues des ACM du thème pratique de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 53groupes (244.04). On choisit donc de former 3 groupes de pratique parmi l'échantillon de 188 individus.

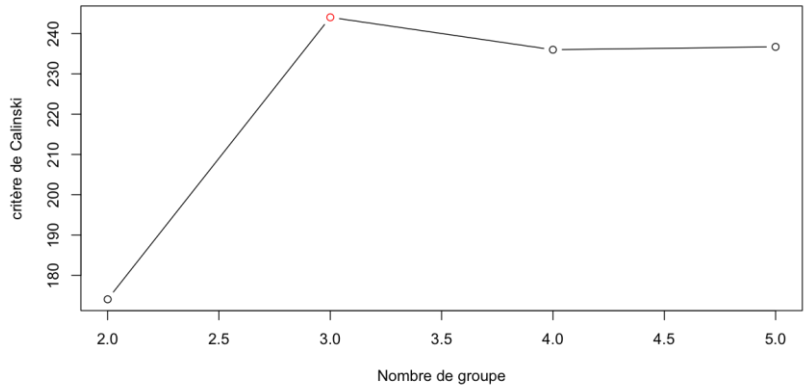


Figure n°II-13: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

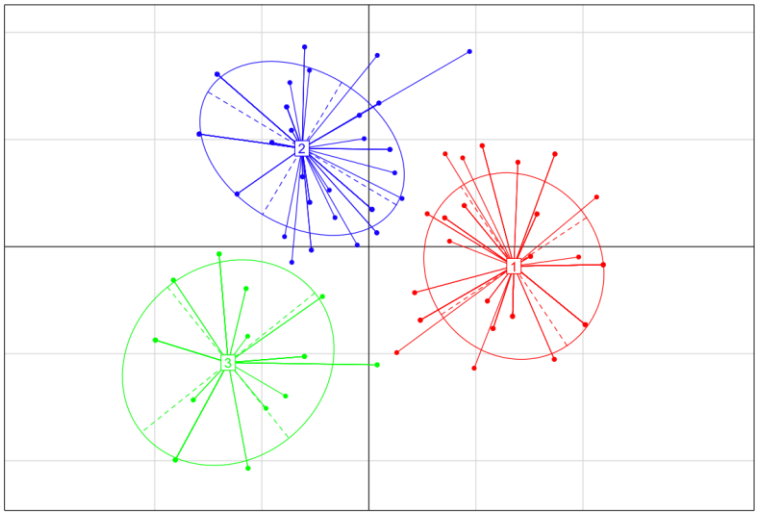


Figure n°II-14: Projection des individus par groupe de pratique (points) sur le plan factoriel de l'ACM

Le groupe 1 rassemble les individus qui pratiquent avec la plus faible fréquence leur activité de pêche. De plus ces individus ne sont pas, pour la plupart, indifférents à la période de l'année, au type de jour ou au moment de la journée. Les individus des groupes 2 et 3 sont , eux, indifférents au moment de la journée, de l'année ou au type de jour. De plus ces individus et particulièrement les individus du groupe 3, pratiquent leur activité avec une plus grande fréquence (tableau II-4).

Tableau n°II-4: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de pratique

	Nombre d'ind.	Variables	proportion	1	2	3
Echantillon	**188**	**pratique**	Occasionnelle: 53%			
		Planification à l'année	Année: 61%			
		Planification à la semaine	Indifférent: 63%			
		Planification à la journée	Indifférent: 62%			
		Nombre de sorties total	<15: 27%			
Groupe 1: *"pratique ponctuelle"*	**75**	**pratique**	Occasionnelle: 92%			
		Planification à l'année	Été: 71%			
		Planification à la semaine	Pas indifférent: 60%			
		Planification à la journée	Indifférent: 59%			
		Nombre de sorties total	<15: 65%			
Groupe 2: *"pratique diffuse"*	**68**	**pratique**	Régulière: 72%			
		Planification à l'année	Année: 79%			
		Planification à la semaine	Indifférent: 78%			
		Planification à la journée	Indifférent: 56%			
		Nombre de sorties total	>79: 47%			
Groupe 3: *"pratique diffuse et intense"*	**45**	**pratique**	Très régulière: 64%			
		Planification à l'année	Année: 96%			
		Planification à la semaine	Indifférent: 80%			
		Planification à la journée	Indifférent: 78%			
		Nombre de sorties total	34-79: 69%			

La variable "planification à la journée" est celle qui discrimine le moins les groupes de pratique entre eux et avec l'échantillon total. Hormis celle-ci les groupes de pratique présentent de nombreuses différences significative entre eux pour chacune des variables du même thème. La structure de groupe mise en évidence est donc pertinente.

Sélection des métriques

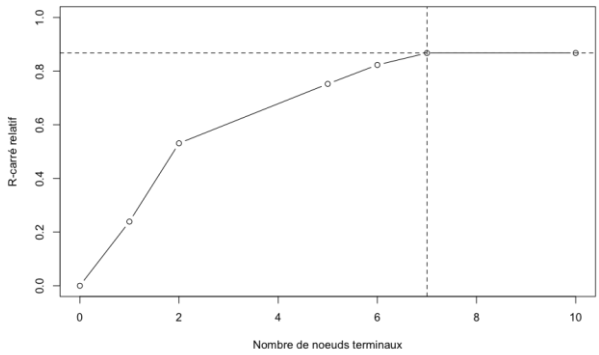


Figure n°II-15: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de noeud terminaux

À partir de 7 partitions, le R-carré relatif de l'arbre de segmentation atteint son maximum (figure II-15). Il suffit donc de 8 noeuds terminaux pour construire l'arbre de segmentation.

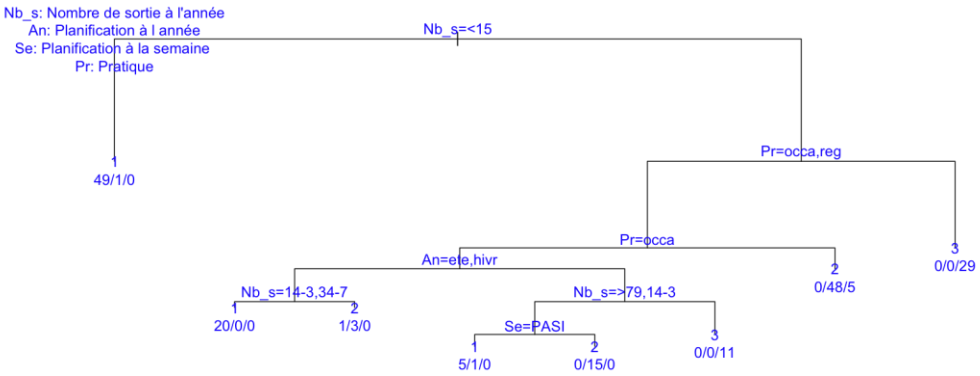


Figure n°II-16: Arbre de reclassification pour les groupes de pratique en fonction des variables du même thème

Les résultats obtenus sont cohérents avec le descriptif des groupes dans le tableau II-4. Le groupe 1 présente le nombre de sortie le plus faible sur l'année et le groupe 3 se distingue du groupe 2 par une pratique très régulière de l'activité. Le nombre de sortie à l'année est la variable qui discrimine le mieux les groupes de pratique (figure II-16). Il s'agit donc de la première métrique à tester comme indicateur potentiel d'effets de variables d'intérêt. Si il est nécessaire de considérer une combinaison de variables, on choisira la combinaison des variables "Nombre de sortie à l'année" et "pratique", en prenant comme bornes des classes les premières valeurs seuils générées par l'arbre de reclassification.

Conclusion

On a mis en évidence 3 groupes de pratique parmi l'échantillon en fonction de la fréquence de l'activité (pratique, nombre de sorties à l'année) et de la planification de l'activité (planification à a journée, à la semaine, à l'année). Le nombre de sorties à l'année est la variable qui reflète le mieux les groupes de pratique.

C. Thème Perception

Tableau n°II-5: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
D1_existence_amp	Oui	Oui	326
	Non	Non	26
D1_amp_choix_lieu	Modéré	plutot_oui	160
	Décisif		
	Faible_nul	Faible_nul	189
	Nsp	Nsp	3
D1_suffi_info	Oui	Oui	231
	Non	Non	67
	Nsp	Nsp	54
D1_adapt_reg	Bien	Bien	194
	Insu	Pas_bien	66
	Trop		
	Nsp	Nsp	92
D1_respect_reg	Oui	Oui	113
	Non	Non	94
	Nsp	Nsp	145
D1_ass_process_dec	Pas_du_tout	Pas	44
	Plutot_pas		
	Tout_a_fait	Dac	170
	Plutot_dac		
	Nsp	nsp	138
D2_effet_ecosyst	Tres_positif	Positif	321
	Plutot_positif		
	Tres_negatif	Negatif	0 (pas pris en compte dans l'analyse)
	Plutot_negatif		
	Neutre	Neutre	13
	Nsp	Nsp	18
D2_effet_econo	Tres_positif	Positif	227
	Plutot_positif		
	Tres_negatif	Negatif	4
	Plutot_negatif		
	Neutre	Neutre	50
	Nsp	Nsp	71
D2_effet_peche	Tres_positif	Positif	186
	Plutot_positif		
	Tres_negatif	Negatif	7
	Plutot_negatif		
	Neutre	Neutre	99
	Nsp	Nsp	60
nb_conflits	quantitatif	≥2	113
		1	64
		0	175

Pour chacune des variables utilisés, une recodage des données est effectué de manière à

avoir un nombre de modalité par variable compris entre 2 et 4. Ce recodage doit permettre de conserver une cohérence au niveau de l'information conservée afin d'être pertinent en temps qu'indicateur de perception.

Typologie

Pour l'ACM du thème perception, la première décroissance de valeurs propre apparaît entre le 3e et le 4e axe (5% de décroissance), cependant seul le premier axe présente des corrélation de plus de 50% avec au mois une variable. De ce fait on ne conserve que les deux premiers axes pour construire l'ACM de perception.

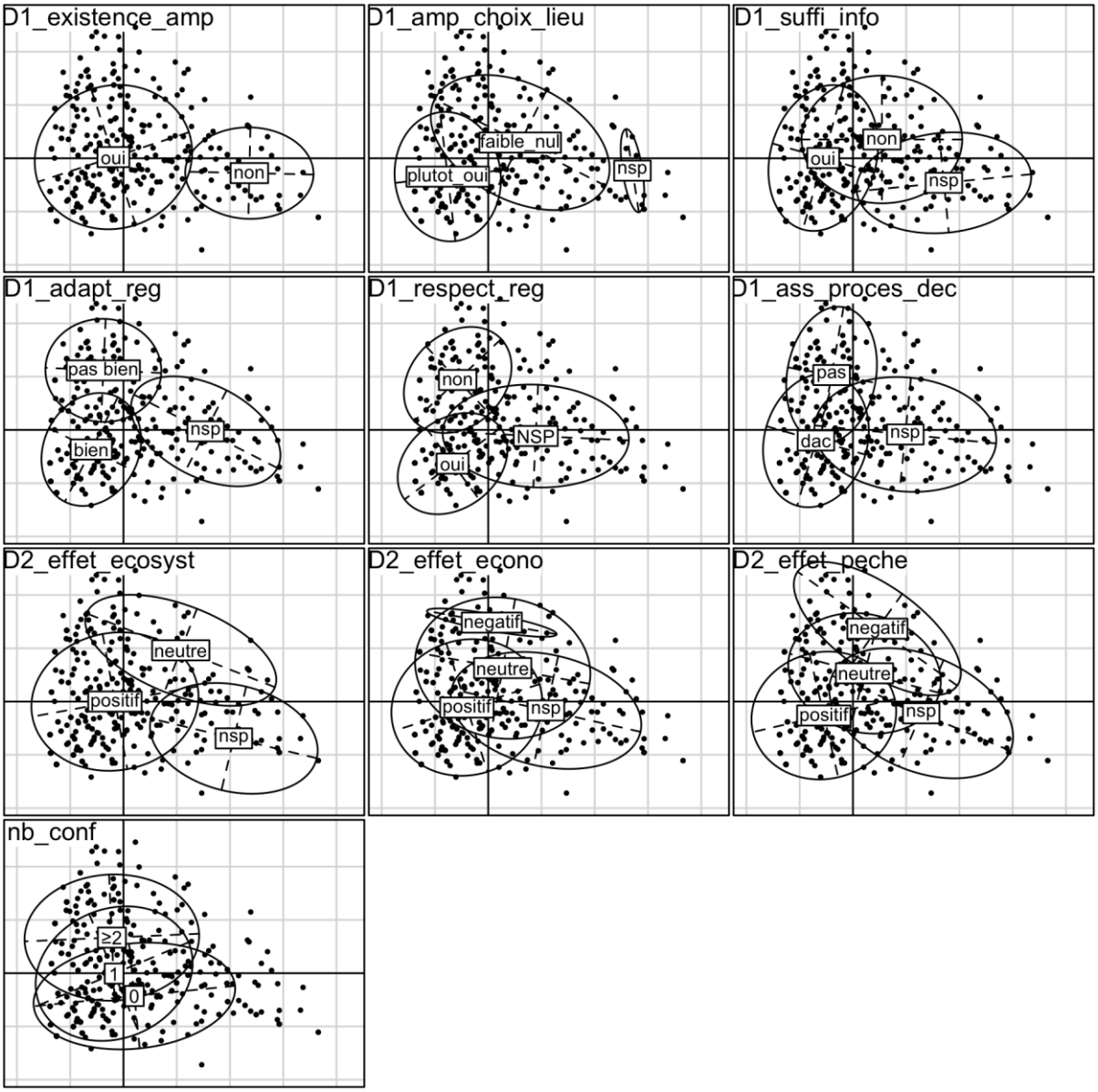


Figure n°II-17: Projection des individus et des modalités pour chaque variable de perception sur le plan factoriel de l'ACM

Sur la figure II-17 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

Sur la figure II-18 est représentée l'évolution des valeurs de Δi en fonction du nombre d'individus de l'échantillon du thème perception. Le calcul de Δi ne peut être fait qu'à partir de 175 individus pour garantir la représentativité de chacune des modalités par au moins 1 individus. On observe une stabilité de la moyenne de Δi entre 175 et 325 individus. Il est donc suffisant d'échantillonner 175 pêcheurs pour avoir une représentation fiable des individus sur le plan factoriel et garantir sa structure.

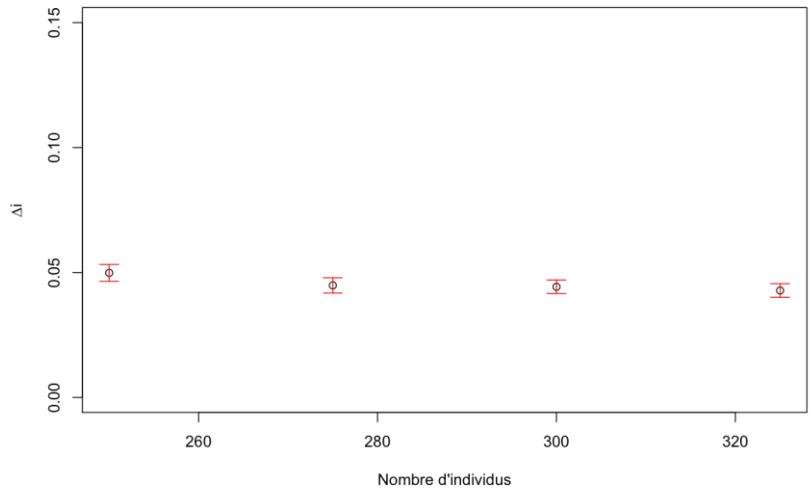


Figure n°II-18: Évolution de la différence moyenne des matrices de corrélation issues des ACM du thème perception de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 3 groupes (399.57). On choisit donc de former 3 groupes de pratique parmi l'échantillon de 352 individus.

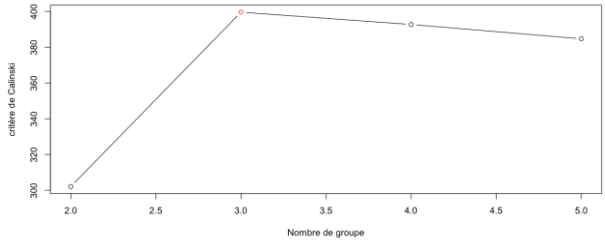


Figure n°II-19: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

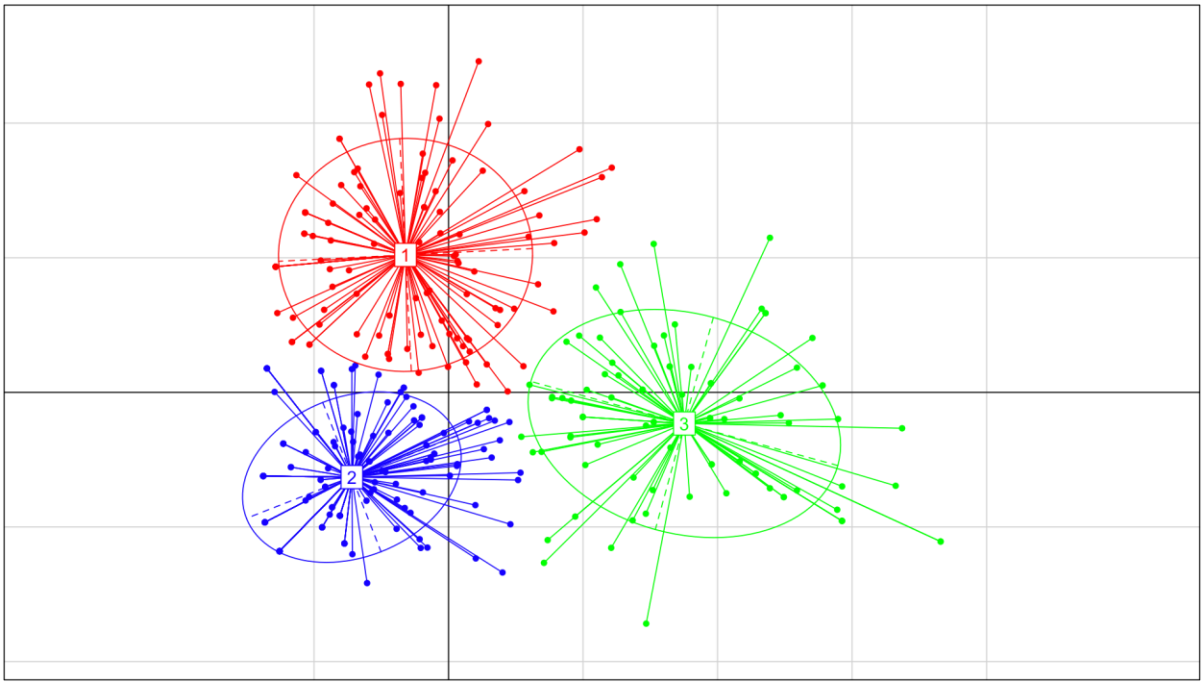


Figure n°II-20: Projection des individus par groupe de perception (points) sur le plan factoriel de l'ACM

Le groupe 1 rassemble les individus qui ont un avis globalement négatif de l'AMP, les individus du groupe 2 en ont un avis plutôt positif tandis que les individus du groupe 3 ne savent pas répondre à la plupart des questions du thème perception (tableau II-6). Sur ce même tableau, on peut voir que la variable "effet de l'AMP sur l'ecosystème" est celle qui discrimine le moins les groupes de perception entre eux et avec l'échantillon total. En effet, la quasi-totalité des individus pensent que l'AMP a un effet positif sur l'écosystème.

Tableau n°II-6: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de perception

Groupes	Variables	Modalités principales	1	2	3
Échantillon	Adaptation de la réglementation	Bien: 55%			
	Association au processus de décision	Daccord: 48%			
	Rôle de l'amp dans le choix du lieu	Faible_mil: 54%			
	Effet sur l'économie	Positif: 64%			
	Effet sur l'écosystème	Positif: 91%			
	Effet dur l'activité de peche	Positif: 53%			
	Connaissance de l'existence de l'amp	Oui: 93%			
	Nombre de conflits	0: 50%			
	Respect de la réglementation	Nsp: 41%			
	Suffisance de l'information	Oui: 66%			
Groupe 1: Avis plutot négatif de l'amp	Adaptation de la réglementation	Pas bien: 51%			
	Association au processus de décision	Daccord: 45%			
	Rôle de l'amp dans le choix du lieu	Faible_mil: 65%			
	Effet sur l'économie	positif: 61%			
	Effet sur l'écosystème	positif: 92%			
	Effet dur l'activité de peche	Neutre: 45%			
	Connaissance de l'existence de l'amp	Oui: 100%			
	Nombre de conflits	≥2: 61%			
	Respect de la réglementation	Non: 65%			
	Suffisance de l'information	Oui: 74%			
Groupe 2: avis plutot positif de l'amp	Adaptation de la réglementation	bien: 95%			
	Association au processus de décision	Daccord: 71%			
	Rôle de l'amp dans le choix du lieu	Plutot_oui: 73%			
	Effet sur l'économie	Positif 82%			
	Effet sur l'écosystème	Positif: 99%			
	Effet dur l'activité de peche	Positif: 76%			
	Connaissance de l'existence de l'amp	Oui: 99%			
	Nombre de conflits	0: 63%			
	Respect de la réglementation	Oui: 60%			
	Suffisance de l'information	Oui: 86%			
Groupe 3: avis plutot indécis de l'amp	Adaptation de la réglementation	Nsp: 84%			
	Association au processus de décision	Nsp: 86%			
	Rôle de l'amp dans le choix du lieu	Faible_mil: 86%			
	Effet sur l'économie	Nsp:48%			
	Effet sur l'écosystème	Positif: 75%			
	Effet dur l'activité de peche	Nsp: 46%			
	Connaissance de l'existence de l'amp	Oui: 70%			
	Nombre de conflits	0: 64%			
	Respect de la réglementation	Nsp: 86%			
	Suffisance de l'information	Nsp: 51%			

Sélection des métriques

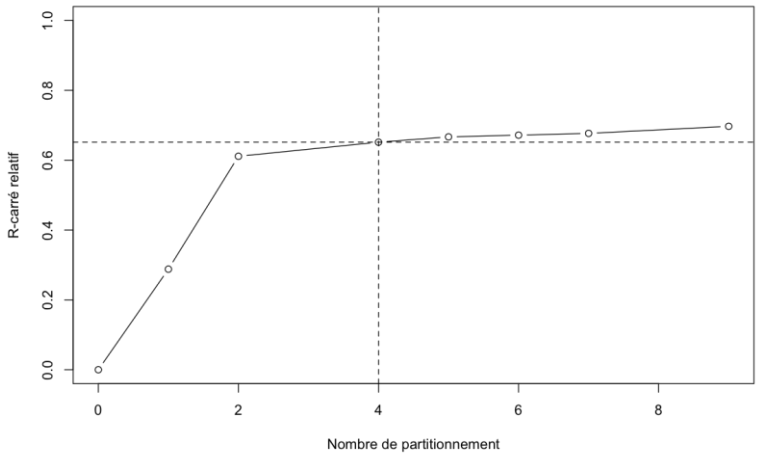


Figure n°II-21: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de noeud terminaux

À partir de 4 partitions de l'échantillon, le R-carré relatif de l'arbre de segmentation atteint pratiquement son maximum. Il suffit donc de 6 noeuds terminaux pour construire l'arbre de segmentation pour prédire l'appartenance des individus aux groupes de pratique à partir des variables du même thème.

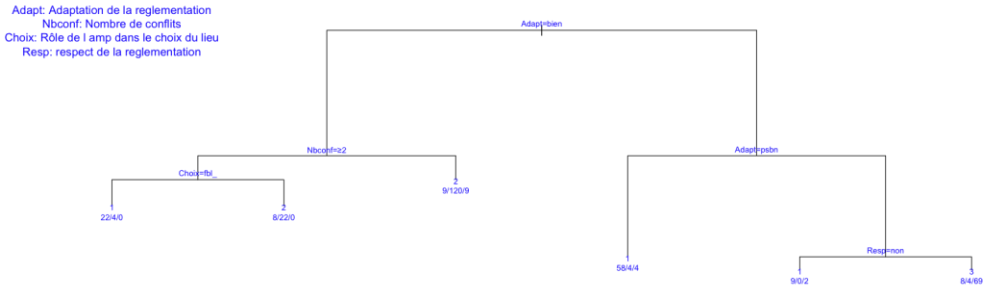


Figure n°II-22: Arbre de reclassification pour les groupes de perception en fonction des variables du même thème

Les résultats obtenus sont cohérents avec le descriptif des groupes dans le tableau n°II-6. Le groupe 2 présente les individus qui pensent principalement que la réglementation est adaptée, les individus du groupe 1 la trouve pour l'essentiel inadapté et les individus du

groupe 3 ne se prononcent pas sur cette question.

L'avis du pêcheur sur l'adaptation de la réglementation est la métrique prioritaire qui serait à analyser comme indicateur potentiel d'effets de variables d'intérêt sur les groupes de perception.

Conclusion

On a mis en évidence 3 groupes de perception parmi l'échantillon. L'avis de l'usager sur l'adaptation de la réglementation est la variable qui reflète le mieux les groupes de perception.

D. Thème Socio-économie

Tableau n°II-7: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
E3_CSP	Actif	actif	178
	Agriculteur		
	Cadre		
	Artisan		
	Employé		
	Intermediaire		
	liberal		
	ouvrier		
	Étudiant	Pas actif	166
	retraité		
	Pas actif		
E2_an_nais	Quantitatif	≤1960	174
		>1960	170
E4_resident	Oui	Oui	108
	Non	Non	236
budget	quantitatif	>200	150
		≤200	194

Pour chacune des variables utilisés, une recodage des données est effectué de manière à avoir un nombre de modalité par variable compris entre 2 et 4. Ce recodage doit permettre de conserver une cohérence de l'information afin d'être pertinent en temps qu'indicateur de socio-économie.

Typologie

Pour l'ACM du thème coscio-économie, la première décroissance de valeurs propre apparaît entre le 2e et le 3e axe (2% de décroissance), De plus les deux premier axes présentent des corrélations de plus de 50% avec au mois une variable. De ce fait on ne conserve que les deux premiers axes pour construire l'ACM de socio-économie.

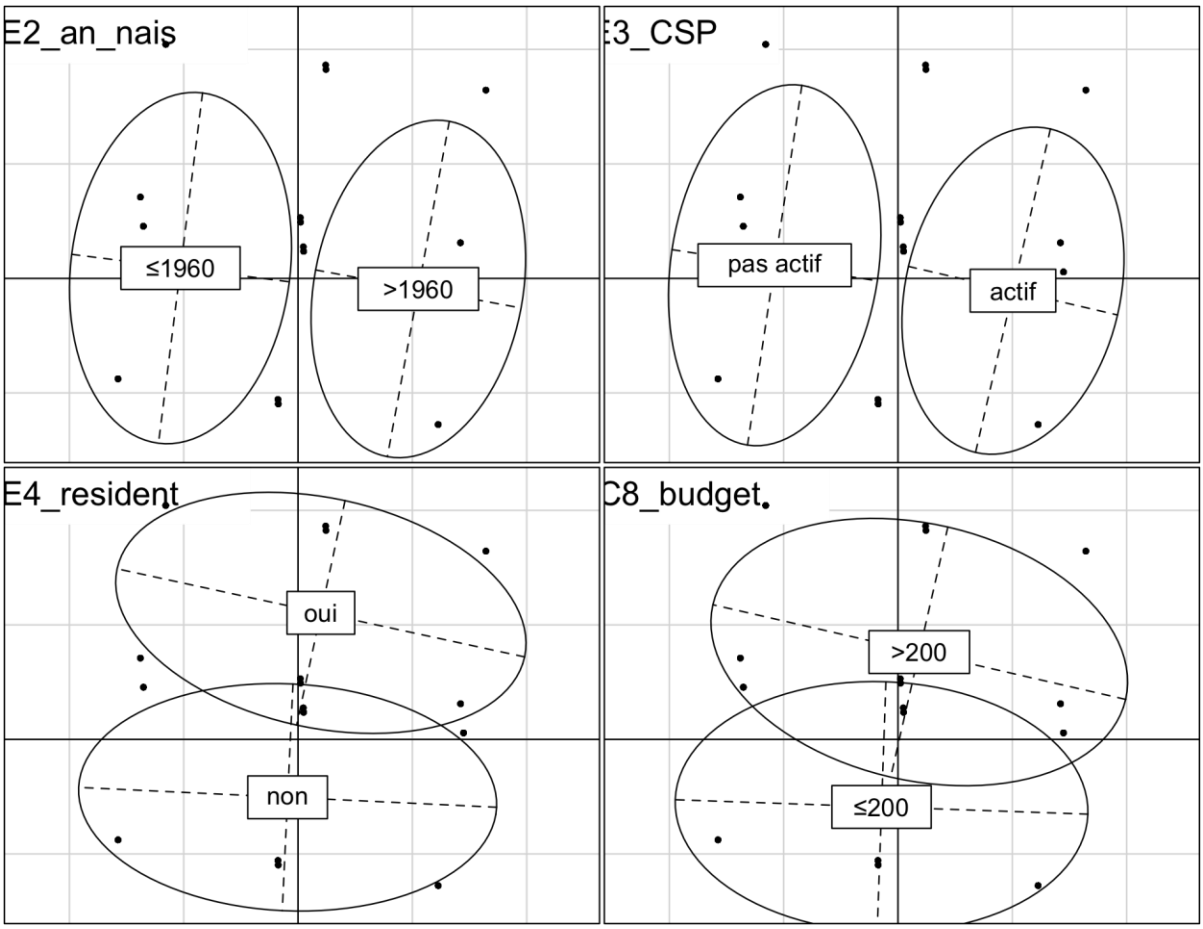


Figure n°II-23: Projection des individus et des modalités pour chaque variable de socio-économie sur le plan factoriel de l'ACM

Sur la figure II-23 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

Les individus résident (commune d'habitation à moins de 40 km de l'AMP) ont tendance à avoir un budget liée à leur activité plus important que les non-résident. La classe d'âge de plus de 50 ans (qui comporte tous les individus retraités) comporte principalement des personnes non-active.

Sur la figure II-24 est représentée l'évolution des valeurs de Δi en fonction du nombre d'individus de l'échantillon du thème perception. On observe une stabilité de la moyenne de Δi entre 250 et 300 individus. Il est donc suffisant d'échantillonner 250 pêcheurs pour avoir une représentation fiable des individus sur le plan factoriel et garantir sa structure.

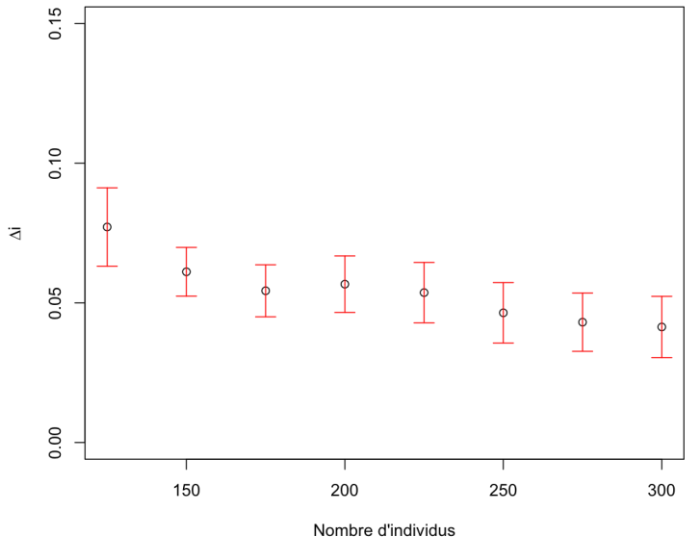


Figure n°II-24: Évolution de la différence moyenne des matrices de corrélation issues des ACM du thème socio-économie de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (488,68). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 344 individus.

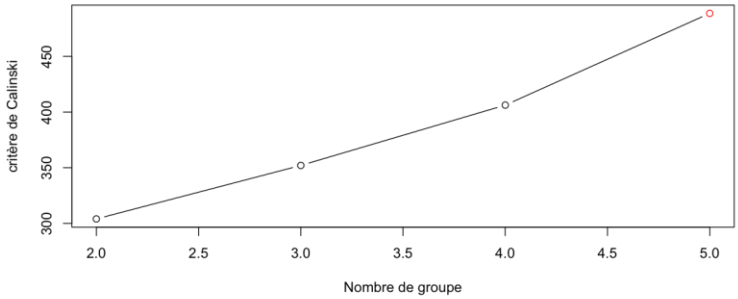


Figure n°II-25: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

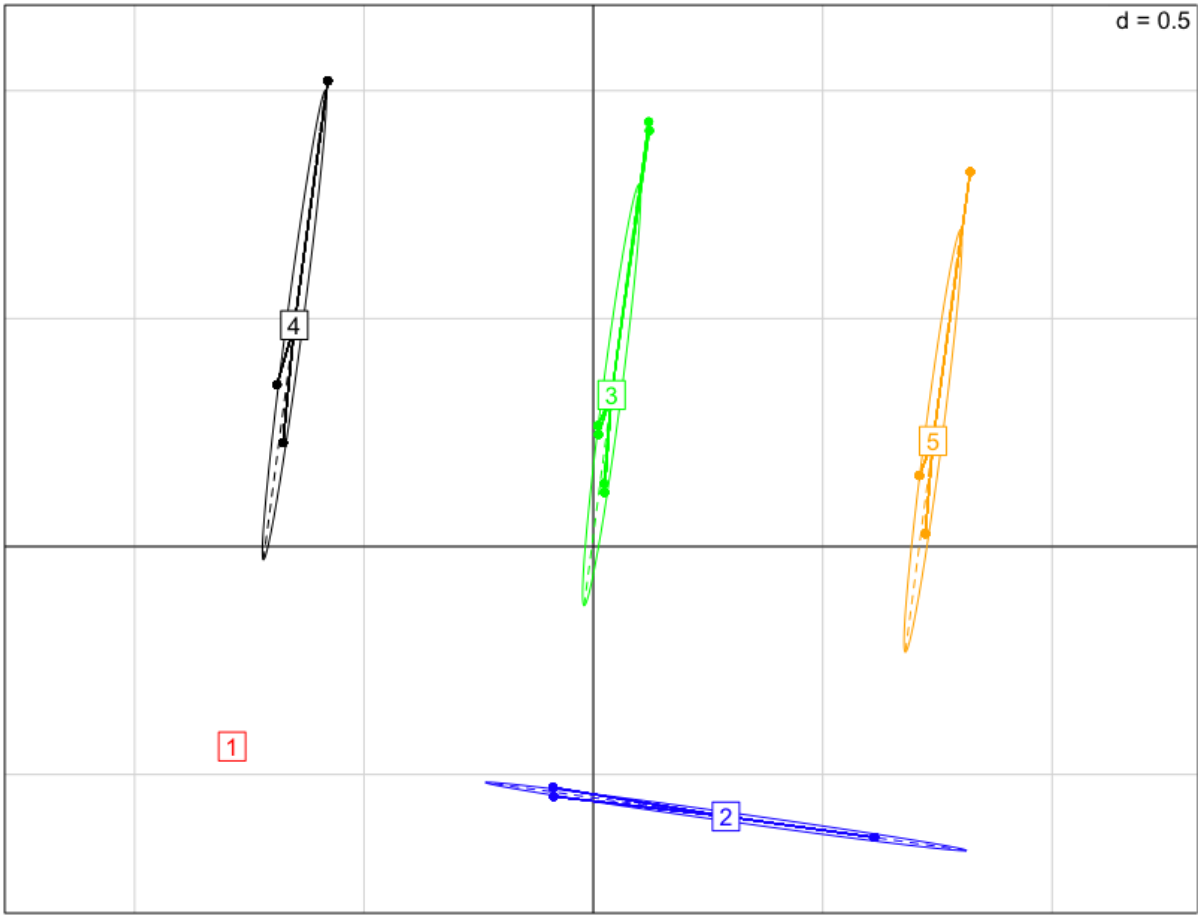


Figure n°II-26: Projection des individus par groupe de socio-économie (points) sur le plan factoriel de l'ACM

Le groupe 1 rassemble les individus de plus de 50 ans et non-actifs, qui sont non résidents et qui consacrent moins de 200 euros à leur activité de pêche (tableau II-8). Les individus du groupe 2 sont eux aussi non résident et consacrent moins de 200 euros à leur activité, cependant ils sont principalement actifs et ont moins de 50 ans. Le groupe 3 est le groupe le plus éclectique. Cependant, la plupart sont actifs de plus de 50 ans et consacrent

plus de 200 euros dans leur activité. Les individus du groupe 4 sont des non actif avec un budget de plus de 200 euros. Enfin les individus du groupe 5 sont des actifs de moins de 50 ans dont le budget alloué à la pêche est de plus de 200 euros.

Tableau n°II-8: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de socio-économie

Groupes	Nombre d'individus	Variables	Modalité principale	1	2	3	4	5
Échantillon	344	CSP	Actif: 52%					
		Année de naissance	≤1960: 50%					
		Résident	Non: 69%					
		budget	≤200: 56%					
1: "non actif et non résident"	55	CSP	Non actif: 100%					
		Année de naissance	≤1960: 100%					
		Résident	Non: 100%					
		budget	≤200: 100%					
2: "actif non résident"	80	CSP	Actif: 78%					
		Année de naissance	>1960: 76%					
		Résident	Non: 100%					
		budget	≤200: 100%					
3: "actif de plus de 50 ans avec un budget de plus de 200 euros"	53	CSP	Actif: 57%					
		Année de naissance	<1960: 57%					
		Résident	Non: 53%					
		budget	>200: 68%					
4; "non actif avec un budget de >200 euros"	70	CSP	Non actif: 100%					
		Année de naissance	≤1960: 100%					
		Résident	Oui: 53%					
		budget	>200: 76%					
5:"actif avec un budget de plus de 200 euros"	86	CSP	Actif: 100%					
		Année de naissance	>1960: 100%%					
		Résident	Oui et non: 50%					
		budget	>200: 71%					

Sélection des métriques

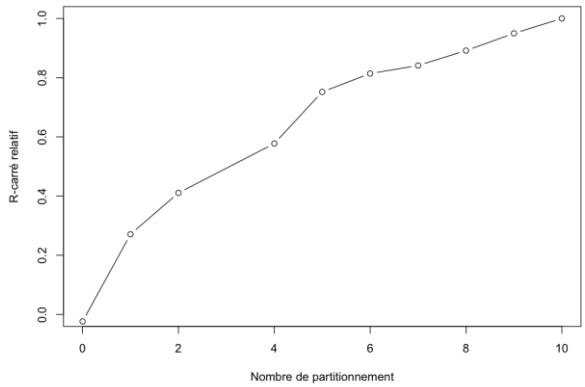


Figure n°II-27: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de noeud terminaux

Le R-carré relatif de l'arbre de segmentation ne cesse d'augmenter avec le nombre de partition. Il est donc nécessaire de conserver la totalité des noeuds terminaux.

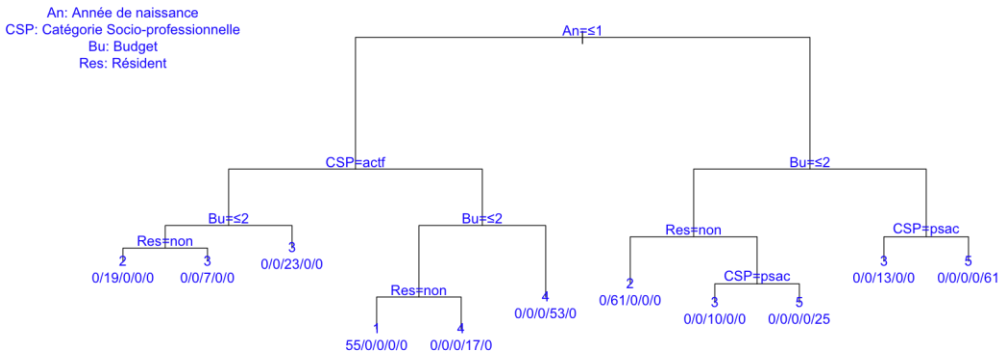


Figure n°II-28: Arbre de reclassification pour les groupes de pression en fonction des variables du même thème

L'année de naissance apparaît par la méthode de segmentation comme la métrique prioritaire à analyser comme indicateur potentiel (figure II-28). La deuxième métrique prioritaire a analyser est une combinaison des variable "année de naissance", "budget" et "CSP" en prenant comme borne des classes les valeurs seuils indiquées par les 3 premières partitions de l'arbre.

Conclusion

On a mis en évidence 5 groupes de socio-économie parmi l'échantillon en fonction de

l'année de naissance, de la catégorie socio-professionnelle, de la qualité de « résident » de l'usager et du budget pour l'activité de l'usager. L'année de naissance est la variable qui reflète le mieux les groupes de socio-économie.

Conclusion du cas d'étude

Pour chacun des thèmes étudiés, un indicateur potentiel a pu être mis en évidence via la construction de typologies. Pour le thème pression, la durée de pêche a été identifiée comme la métrique la plus pertinente a étudier comme indicateur potentiel, pour le thème pratique, il s'agit du nombre de sorties à l'année, pour le thème perception, de l'avis du pêcheur sur l'adaptation de la réglementation et enfin, pour le thème socio-économie, l'année de naissance.

La typologie de pratique est un facteur explicatif qui a un effet significatif sur l'indicateur de pression. Ainsi la durée de pêche d'un usager, et plus généralement la pression qu'il exerce sur le milieu dépend de son appartenance à un groupe de pratique donné. L'injection de certaines typologies d'autres thèmes, en tant que facteur explicatif pour prédire un indicateur potentiel donnée, dans un modèle statistique permet :

- d'augmenter la prédiction du modèle
- de normaliser la distribution des résidus du modèle
- mettre en évidence des effets d'intérêt

III. Cas de la Nouvelle-Calédonie

La totalité de l'aire marine protégées de la Nouvelle-Calédonie, qui couvre plus de 457 km^2 (i.e. plus de 2% de la surface du lagon). Au sein de cette AMP sont appliquées, des restrictions sur les prises (i.e. taille, quota, espèces) et sur les engins utilisé depuis 1967, 1974 et 1981. Cette AMP comprend 2 sortes de zone de régulation spécifique. On y retrouve une zone de protection intégrale (172 km^2) à l'intérieure de laquelle aucune entrée ou activité est autorisée depuis 1970. De plus 18 zones de protection (280 km^2), établies entre 1981 et 2006, correspondent à des fermetures saisonnières et des régulations de pêche renforcées. Il est important de noter la surveillance et l'application de la réglementation est très faible.

Les données issues de ce cas d'étude proviennent de questionnaires sur 841 pêcheurs embarqués à leur retour de sortie entre octobre 2004 et octobre 2005.

A. Identification des thèmes

La sélection des données (nombre de variables et nombre d'individus) est présentée pour chaque thème dans le tableau III-1.

Tableau n°III-1: Tableau récapitulatif du nombre d'individus et du nombre de variables sélectionnées pour la construction du plan factoriel de chaque thème

	Nouvelle-Calédonie	
thèmes	Nombre d'individus	Nombre de variables
Pression	396	4
Pratique	81	5
Impact	267	4
Socio-économie	704	3

B. Thème Pression

Étant donné la présence de valeurs extrêmes pour une ou plusieurs des quatre variables du thème pression, 22 individus sont écartés pour la construction du plan factoriel de ce thème.

Typologie

Lors de la construction du plan de l'ACP pour le thème pression, les deux premiers axes rassemblent 77% de l'inertie du nuage de points. De plus ils permettent de garantir une réelle signification de ces axes avec des corrélations supérieures à 0,5 pour au moins une variable. D'une part L'axe 1 est corrélé négativement avec le nombre d'engins par pêcheur et positivement avec le nombre de pêcheurs. L'axe 1 reflète donc la pression de pêche en termes de moyens déployés. D'autre part l'axe 2 est corrélé positivement avec le rapport de la durée de pêche sur la durée de sortie et négativement avec la durée de sortie. Ce deuxième axe reflète donc la pression de pêche en termes de temps consacré à l'activité (figure III-1).

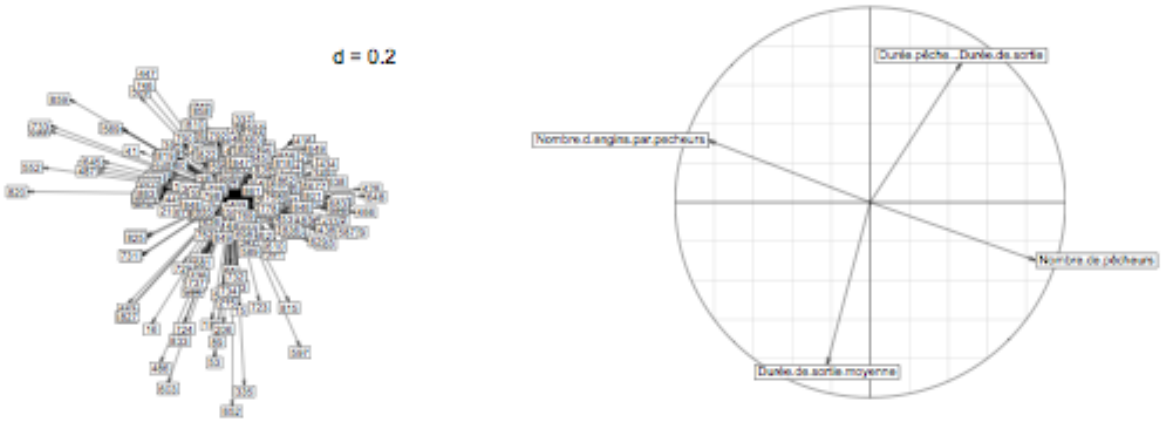


Figure n°III-1: projection des individus sur le plan (gauche) et cercle de corrélation (droite) de l'ACP du thème pression des données de la Nouvelle-Calédonie

Sur la figure III-2 est représentée l'évolution des valeurs de Δi en fonction du nombre d'individus de l'échantillon du thème pression. On observe une stabilité de la moyenne de Δi entre 150 et 350 individus. Il est donc nécessaire d'échantillonner au moins 150 pêcheurs pour avoir une représentation fiable des individus sur le plan factoriel et garantir sa structure.

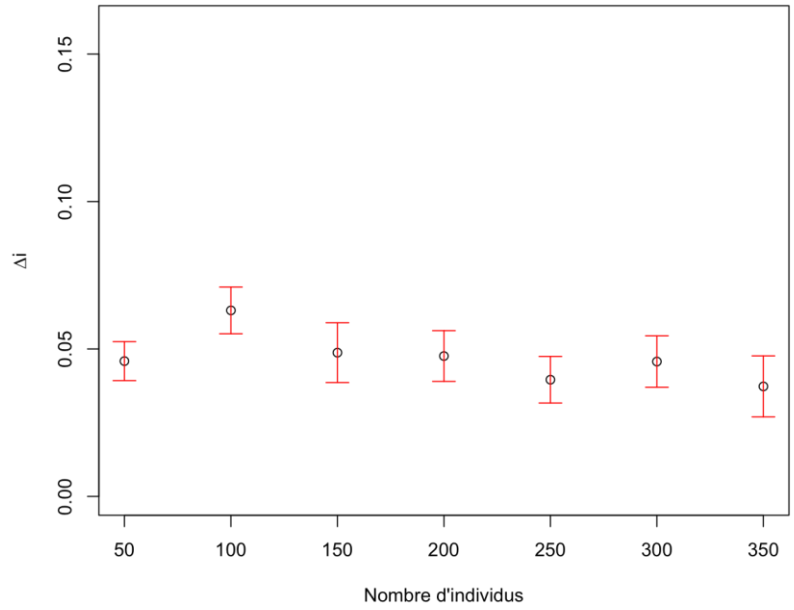


Figure n°III-2: Évolution de la différence moyenne des matrices de corrélation issues des ACP du thème pression de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (289,49). On choisit donc de former 5 groupes de pression parmi l'échantillon de 396 individus.

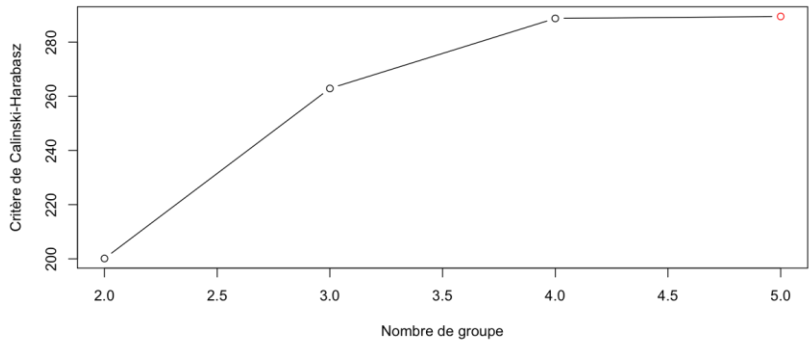


Figure n°III-3: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

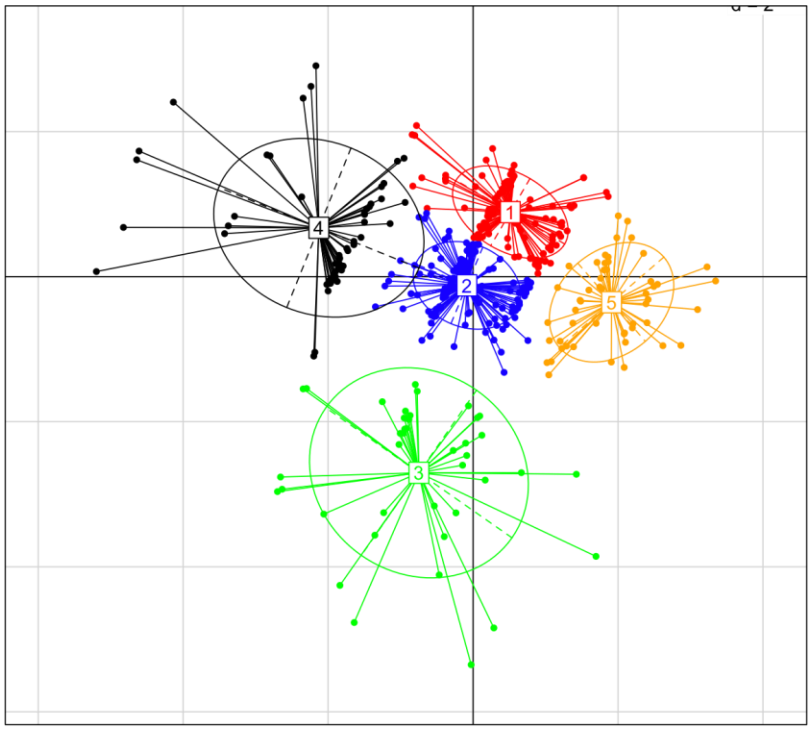


Figure n°III-4: Projection des individus par groupe de pression (points) sur le plan factoriel et représentation des diagrammes de dispersions associés pour chacune des variables du même thème

Le groupe 3 se distingue par une durée de sortie importante, le groupe 4 par un nombre d'engin par pêcheur élevée et le groupe 5 par un grand nombre de pêcheurs (figure III-4). Enfin le groupe 2 est celui qui se distingue le moins de l'échantillon total.

Tableau n°III-2: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de pression

	Description de l'échantillon et des groupes						Groupes				
Groupes	*Nombre d'individus*	*Variables*	*moy.*	*med.*	*sd*	*C.V.*	*1*	*2*	*3*	*4*	*5*
Échantillon	396	*Nombre de pêcheurs*	2.45	2	1.07	44					
		Nombre d'engins / pêcheur	0.59	0.5	0.43	73					
		Durée de sortie	10.31	7	11.21	109					
		Durée de pêche / durée de sortie	0.47	0.46	0.28	60					
Groupe 1: *"Nombre d'engins faible et durée de pêche élevée"*	118	*Nombre de pêcheurs*	2.38	2	0.55	23					
		Nombre d'engins / pêcheur	0.52	0.5	0.19	37					
		Durée de sortie	5.77	5.75	2.25	39					
		Durée de pêche / durée de sortie	0.77	0.78	0.13	17					
Groupe 2: *"Faible pression en général"*	132	*Nombre de pêcheurs*	2.30	2	0.46	20					
		Nombre d'engins / pêcheur	0.51	0.50	0.19	37					
		Durée de sortie	7.58	7.5	2.98	39					
		Durée de pêche / durée de sortie	0.34	0.35	0.15	45					
Groupe 3: *"Nombreux engins et un seul type d'engin"*	40	*Nombre de pêcheurs*	2.53	2	1.06	42					
		Nombre d'engins / pêcheur	0.51	0.5	0.25	49					
		Durée de sortie	39.17	32.25	14.79	38					
		Durée de pêche / durée de sortie	0.09	0.09	0.06	64					
Groupe 4: *"Durée de sortie élevée et nombre de pêcheurs faible"*	55	*Nombre de pêcheurs*	1.06	1	0.23	22					
		Nombre d'engins / pêcheur	1.36	1	0.64	47					
		Durée de sortie	7.46	6.5	5.23	70					
		Durée de pêche / durée de sortie	0.33	0.24	0.26	81					
Groupe 5: *"Durée de sortie élevée et nombre de pêcheurs important"*	51	*Nombre de pêcheurs*	4.45	4	0.70	16					
		Nombre d'engins / pêcheur	0.24	0.25	0.08	34					
		Durée de sortie	8.29	8	3.32	40					
		Durée de pêche / durée de sortie	0.57	0.58	0.23	40					

Les groupes 4 et 5 sont ceux qui se distinguent le mieux de l'échantillon total. Seule la variable durée de pêche / durée de sortie permet de distinguer la totalité des groupes vis à visde l'échantillon total.

Sélection des métriques

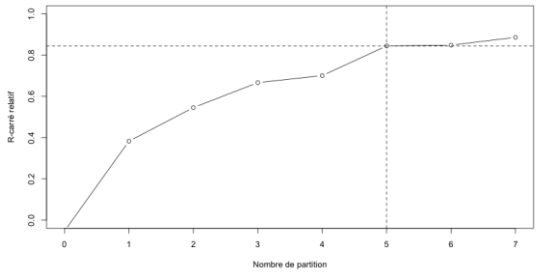


Figure n°III-5: Évolution du R-carré relatif de l'arbre de segmentation en fonction du nombre de partition

On n'observe qu'une très légère augmentation du R^2 relatif pour un nombre partition de l'échantillon supérieur à 5. On choisit donc de conserver uniquement 5 partitions de l'échantillon pour l'arbre de reclassification du thème pression.

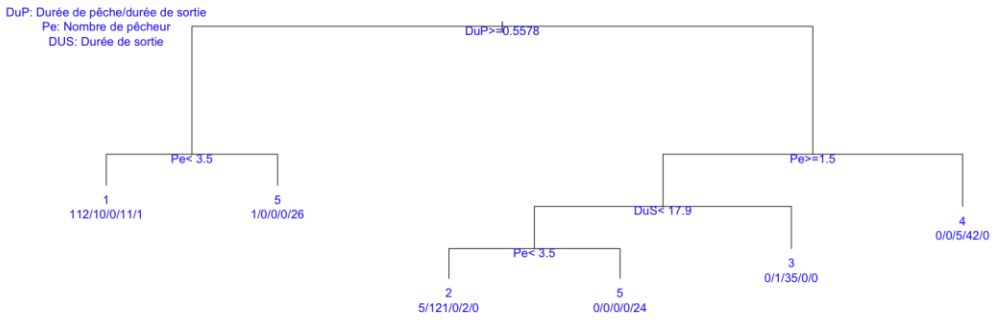


Figure n°III-6: Arbre de reclassification pour les groupes de pression en fonction des variables du même thème

Sur la figure III-6, sont indiquées, au niveau des noeuds de l'arbre, les variables et valeurs seuils associées pour la partition de l'échantillon tandis qu'au niveau des noeuds terminaux est représenté le numéro de groupe associé. En-dessous de ce numéro est indiqué le nombre d'individus de chaque groupe qui compose le noeud terminal (nombre d'individus du groupe n°1 / nombre d'individus du groupe n°2 / … / nombre d'individus du groupe n°5). Cette indication permet de nuancer les résultats de la segmentation pour des noeuds terminaux dont la composition est diverse.

Les résultats obtenus sont cohérents avec le descriptif des groupes dans le tableau III-2. Les individus des groupes 1 et 5 consacrent moins de 56% du temps de leur sortie à la pêche. Les individus du groupe 4 consacre plus de la moitié de leur temps de sortie à la pêche et ne pêchent pas seuls. Le rapport Durée de pêche / durée de sortie est la métrique prioritaire à analyser comme indicateur potentiel.

Mise en évidence d'effet

La variable à expliquer est le log de durée de pêche / durée de sortie à laquelle on a ajouté 10e-05. Cette transformation en log +10e-05 est effectuée afin que les résidus des modèles suivent une loi Normale puisque sans cette transformation, la totalité des modèles est rejetée.

Tableau n°III-3: détails des facteurs sélectionnés et des paramètres statistiques de chacun des modèles construits

Modèles	Facteurs explicatifs	Paramètres	
Modèle 1	Saison; **type de jour; taille de bateau; nombre de sortie sur l'année; groupe de lagon**	Nombre d'individus	160
		Erreur résiduelle standard	0.273 / 146df
		R2	0.2682
		AIC	-402.1354
		Cp de Mallow	18.65
Modèle 2	Saison; **type de jour; taille de bateau; nombre de sortie sur l'année; groupe de lagon**; groupe de socio-économie	Nombre d'individus	159
		Erreur résiduelle standard	0.275 / 144df
		R2	0.269
		AIC	-396.6572
		Cp de Mallow	19.56
Modèle 3	Saison; **type de jour; taille de bateau; nombre de sortie sur l'année; groupe de lagon**	Nombre d'individus	160
		Erreur résiduelle standard	0.273 / 146df
		R2	0.2682
		AIC	-402.1354
		Cp de Mallow	18.65

Les effets des facteurs sont mis en évidence avec une ANOVA de type III. Les **facteurs en gras** ont un effet significatif en dessous du seuil de 0,05, les facteurs soulignés ont un effet significatif en dessous du seuil de 0,1.

.

Les Modèles 1 et 3, qui sont identique, sont les modèles conservés par la méthode car ils ont l'AIC et le Cp de Mallow les plus faible. On constate une variabilité temporelle (saison) et spatiale (groupe de lagon) de la pression de pêche. Cependant la totalité des tests mettant en évidence l'effet des facteurs explicatifs ont une puissance inférieure à 80% et ne permettent pas de conclure avec certitude. Pour mettre en évidence ces effets sur la pression de pêche il faudrait augmenter le nombre d'observations ou que la taille de l'effet soit plus importante (cf. Fiche descriptive (1)).

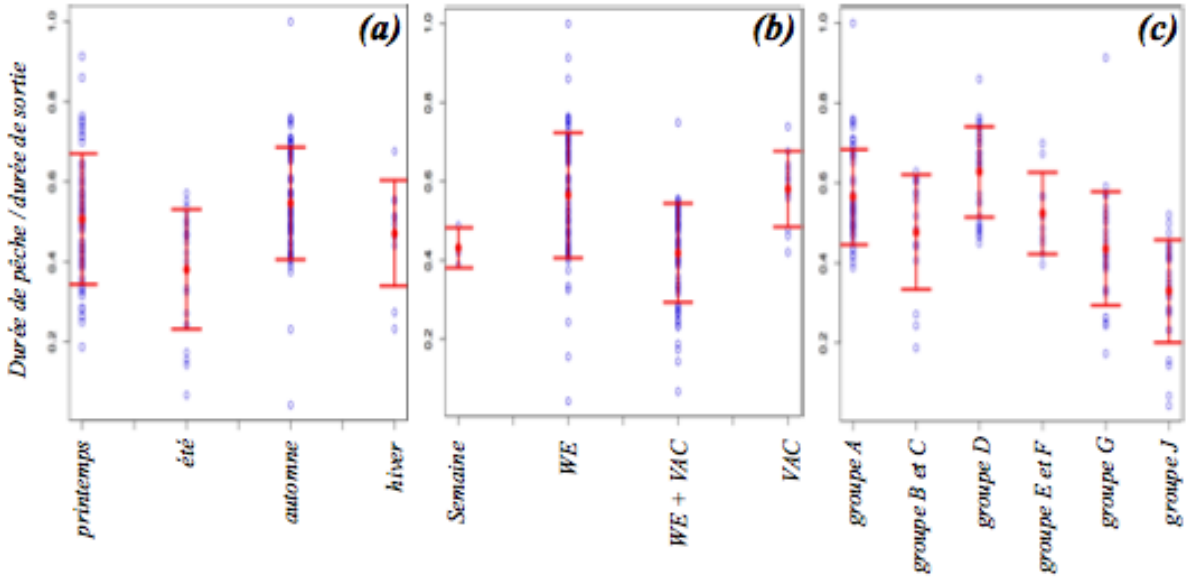


*Figure n°III-7: Valeurs prédites du modèle en fonction de la saison (**a**), du type de jour (**b**) et de la zone de pêche (**c**)*

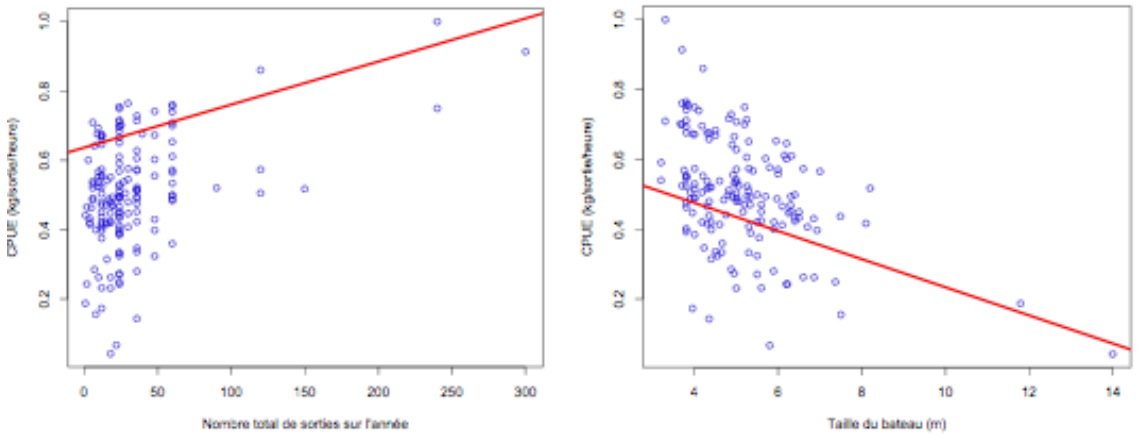


Figure n°III-8: Valeurs prédites du modèle en fonction du nombre de sortie sur l'année (gauche), de la taille du bateau (droite)

Malgré une puissance des tests statistiques faible pour mettre en évidence les effets des facteurs explicatifs sur le rapport durée de pêche / durée de sortie on peut observer des tendances dans les résultats obtenus:

C'est durant l'automne et le printemps que l'on observe les valeurs de l'indicateur les plus élevées avec plus de la moitié du temps de la sortie consacrée à la pêche (respectivement et en moyenne 0.54 et 0.51). En revanche les saison d'été et d'hiver sont celle qui présentent les valeurs de l'indicateur le plus faibles (respectivement en moyenne 0.38 et 0.47).

Les valeurs prédites de l'indicateur dans le modèle en fonction de du type de jour ne permet pas d'interprétation fiable du fait de l'hétérogénéité du nombre d'individus échantillonnée entre les modalités de cette variable (entre 3 et 81).

Les groupes de lagons de A à D (groupes de lagon les plus proches de Nouméa) présentent les valeurs prédites du modèle les plus élevées. En effet l'indicateur présente en moyenne la valeur de 0,56 pour le groupe de lagon A, 0,48 pour le groupe de lagon B et C et 0,63 pour le groupe de lagon D. Pour le reste, plus le groupe de lagon est éloignée de Nouméa, plus la valeur prédite de l'indicateur sera faible pour atteindre 0,32 en moyenne pour le groupe de lagon J.

Le rapport durée de pêche / durée de sortie prédit par le modèle est corrélé positivement au nombre de sorties et négativement à la longueur du bateau. Ainsi les pêcheurs sortant en mer le plus souvent sont aussi ceux qui consacre relativement le plus de temps à la pêche à chaque sortie (coefficient = 0.00124). À l'inverse les pêcheurs possédant les plus grands bateaux sont aussi ceux qui consacre relativement le moins de temps à la pêche durant chaque sortie (coefficient = -0.04025).

Conclusion

On a mis en évidence 5 groupes de pression parmi l'échantillon en fonction du rapport durée de pêche / durée de sortie, de la durée de sortie, du nombre de pêcheur et du rapport nombre d'engin / nombre de pêcheur. Le rapport durée de pêche / durée de sortie est la variable qui reflète le mieux les groupes de pression. Elle s'avère être donc l'indicateur le plus pertinent de la période de pêche (saison, type de jour), de la fréquence de pêche et de la

zone de pêche. Il faudrait une différence du rapport de la durée de pêche / la durée de sortie dans et hors la réserve pour mettre en évidence un effet réserve.

C. Thème Pratique

Les variables correspondant au planning de l'activité de la semaine on été codées en 2 modalités: "IND" dans le cas où il y a indifférence de cette planification ou "PAS IND" dans le cas contraire. La planification à l'année comporte 3 modalités: "année/ind", "été" selon que l'usager pêche toute l'année, uniquement en été ou uniquement en hiver. La variable "nombre de facteurs environnementaux" comprend 3 modalités: "0", "1" et "2". La variable nombre d'activités comprend 2 modalités: "1" et "2" (tableau III-4).

Tableau n°III-4: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
Activité principale	Chasse	CHAS	38
	Embarquée	EMB	43
Plannification à l'année	Année	Année/ ind	75
	Indifférent		
	Été	Été	6
Plannification à la semaine	JS	PAS IND	44
	JW		
	JW+VAC		
	VAC		
	IND	IND	37
Nombre d'activité	quantitatif	1	61
		2	20
Nombre de facteurs environnementaux	quantitatif	2	23
		1	55
		0	3

Typologie

Pour l'ACM du thème pratique, la décroissance des valeurs propres entre le 2e et le 3e axe est de 2%. L'axe 1 est corrélé à plus de 50% avec les variables liées aux activités de pêche ("activité habituelle" et "nombre d'activité"). Le deuxième axe est corrélé à plus e 50% avec la variable "planification à la semaine". Enfin le troisième axe est corrélé à plus de 50% avec la variable "nombre de facteurs environnementaux". On conserve donc les trois

premiers axes factoriels pour réaliser l'ACM de pratique.

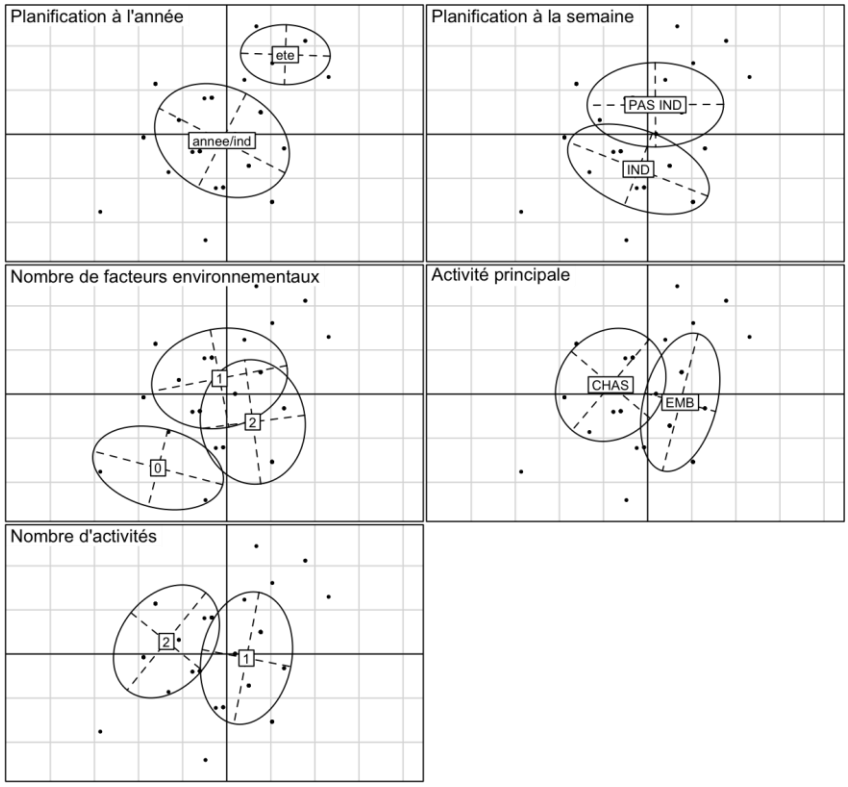


Figure n°III-9: Projection des individus et des modalités pour chaque variable de pratique sur le plan factoriel de l'ACM

Sur la figure III-9 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

En général, les modalités représentant des individus indifférents aux facteurs environnementaux, à la période de l'année ou du jour de la semaine sont appareillés et inversement. Ensuite on peut voir, que les individus pratiquant la chasse sous-marine comme activité principale ont préférentiellement plus d'une activité de pêche.

Sur la figure III-10 est représentée l'évolution des valeurs de Δi en fonction du nombre d'individus de l'échantillon du thème pratique. On observe une stabilité de la moyenne de Δi entre 200 et 250 individus. Il est donc nécessaire d'échantillonner au moins 200 pêcheurs

pour avoir une représentation fiable des individus sur le plan factoriel et garantir la structure mise en évidence.

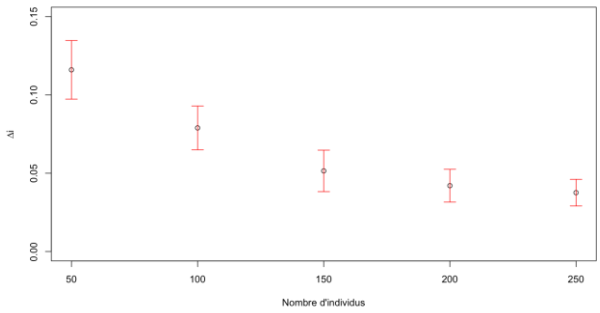


Figure n°III-10: Évolution de la différence moyenne des matrices de corrélation issues des ACM du thème pratique de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (66,71). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 81 individus.

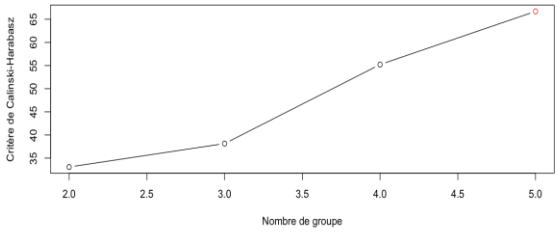


Figure n°III-11: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

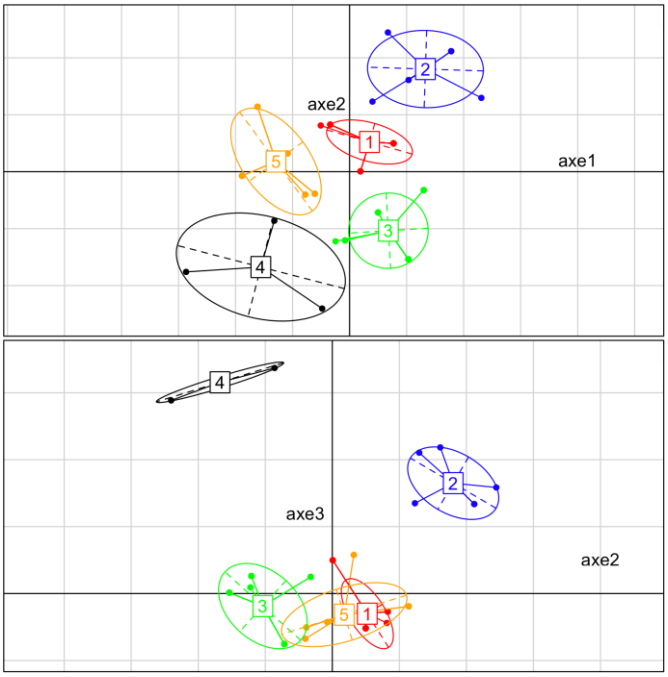


Figure n°III-12: Projection des individus par groupe de pratique (points) sur les plans factoriels de l'ACM

Le groupe 1 rassemble les individus qui ne sont ni indifférent au jour de la semaine, ni aux conditions environnementales mais qui pêchent toute l'année. Le groupe 2 contient des individus qui pêchent uniquement en été. Le groupe 3 est composé de pêcheurs planifiant leurs sorties en fonction des facteurs environnementaux. Les individus du groupe 4 sont des chasseurs exclusifs. Enfin le groupe 5 comporte des chasseurs qui pratiquent d'autres activités et qui conditionnent leurs sorties en fonction de facteurs environnementaux (figure III-12 et tableau III-5).

Il n'y a que très peu de différences significatives entre les groupes et l'échantillon total ainsi qu'entre les groupes eux-mêmes. La structure de groupe mise en évidence est donc issue d'un échantillon relativement homogène ce qui peut être dû au faible effectif de l'échantillon.

Tableau n°III-5: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de pratique

	Description de l'échantillon et des groupes			Groupes				
Groupes	*Nombre d'individus*	*Variables*	*Modalités principales*	*1*	*2*	*3*	*4*	*5*
Échantillon	*81*	*Planification à l'année*	Année / indifférent: 92%					
		Planification à la semaine	Pas indifférent: 54%					
		Nombre de facteurs environnementaux	1: 67%					
		Activité de pêche principale	Embarquée: 53%					
		Nombre d'activités de pêche	1: 75%					
Groupe 1: *"pêcheur pas indifférent au type de jour, aux conditions environnementales et qui pêche toute l'année"*	*27*	*Planification à l'année*	Année / indifférent: 100%					
		Planification à la semaine	Pas indifférent: 83%					
		Nombre de facteurs environnementaux	1: 85%					
		Activité de pêche principale	Embarquée: 59%					
		Nombre d'activités de pêche	1: 96%					
Groupe 2: *"Pêcheur estival"*	*6*	*Planification à l'année*	Ete: 100%					
		Planification à la semaine	Pas indifférent: 83%					
		Nombre de facteurs environnementaux	1: 67%					
		Activité de pêche principale	Embarquée: 67%					
		Nombre d'activités de pêche	1: 83%					
Groupe 3: *"pêcheur planifiant ses sorties en fonction des conditions environnementales"*	*23*	*Planification à l'année*	Année / indifférent: 100%					
		Planification à la semaine	Indifférent: 87%					
		Nombre de facteurs environnementaux	2: 65%					
		Activité de pêche principale	Embarquée: 86%					
		Nombre d'activités de pêche	1: 96%					
Groupe 4: *"Chasseur exclusif indifférent aux conditions environnementales"*	*3*	*Planification à l'année*	Année / indifférent: 100%					
		Planification à la semaine	Indifférent: 67%					
		Nombre de facteurs environnementaux	0: 100%					
		Activité de pêche principale	Chasse: 67%					
		Nombre d'activités de pêche	1: 67%					
Groupe 5: *"chasseur pluri-actif planifiant ses sorties en fonction d'un facteur environnemental"*	*22*	*Planification à l'année*	année / indifférent: 100%					
		Planification à la semaine	Indifférent: 64%					
		Nombre de facteurs environnementaux	1: 91%					
		Activité de pêche principale	Chasse: 91%					
		Nombre d'activités de pêche	2: 73%					

La variable "planification à la l'année" est celle qui discrimine le moins les groupes de pratique entre eux et avec l'échantillon total (sauf le groupe 2).

Sélection des métriques

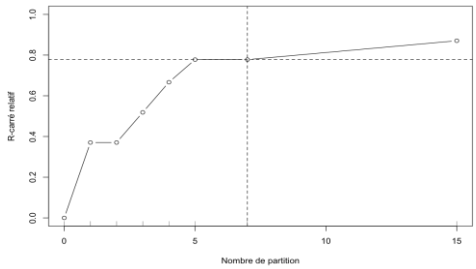


Figure n°III-13: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de noeud terminaux

À partir de 7 partitions, le R-carré relatif de l'arbre de segmentation atteint son maximum (figure III-13). Il suffit donc de 8 noeuds terminaux pour construire l'arbre de segmentation.

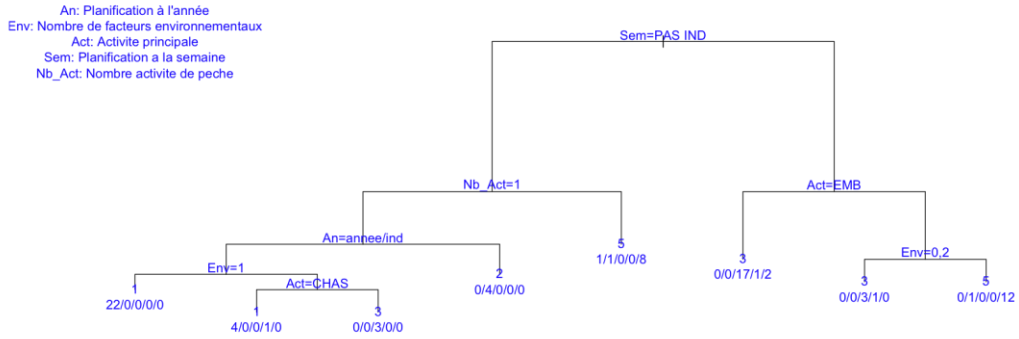


Figure n°III-14: Arbre de reclassification pour les groupes de pression en fonction des variables du même thème

La planification à la semaine est la variable qui discrimine le mieux les groupes de pratique (figure III-14). Il s'agit donc de la première métrique à tester comme indicateur potentiel. Si il est nécessaire de considérer une combinaison de variables, on choisira la combinaison des variables "Planification à la semaine", "Nombre d'activité" et "Activité principale", en prenant comme bornes des classes les premières valeurs seuils générées par l'arbre de reclassification.

Conclusion

On a mis en évidence 5 groupes de pratique parmi l'échantillon en fonction du nombre d'activités pratiquées, de l'activité principale, du nombre de facteurs environnementaux influant sur la sortie de pêche et sur la planification des sorties de pêche (à l'années, à a semaine). La planification de l'activité à la semaine est la variable qui reflète le mieux les groupes de pratique. Elle s'avère être donc la métrique le plus pertinent à tester comme indicateur potentiel.

D. Thème Impact

Tableau n°III-6: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
Nombre d'espèces	quantitatif	1	81
		[2-3]	99
		≥4	87
Catégorie de capture	0-10kg	0-50kg	79
	11-20kg		
	21-50kg		
	51-100kg	51-200kg	60
	101-200kg		
	>200kg	>200kg	128
Capture	Oui	Oui	231
	Non	Non	36
CPUE (kg/sortie/h)	Quantitatif	≤1	68
		]1-2,73]	67
		]2,73-5,97]	67
		>5,97	65

Pour chacune des variables utilisés, une recodage des données est effectué de manière à avoir un nombre de modalité par variable compris entre 2 et 4. Ce recodage doit permettre de conserver une cohérence au niveau de l'information conservée afin d'être pertinent en temps qu'indicateur d'impact.

Typologie

Pour l'ACM du thème impact, les décroissances de valeurs propre sont toutes supérieures à 10%, cependant la plus faible se trouve entre le 3e et le 4e axe (11%). Les trois premiers axes sont corrélés à plus de 50% avec au moins une variable. On conserve donc 3 axes pour réaliser l'ACM d'impact.

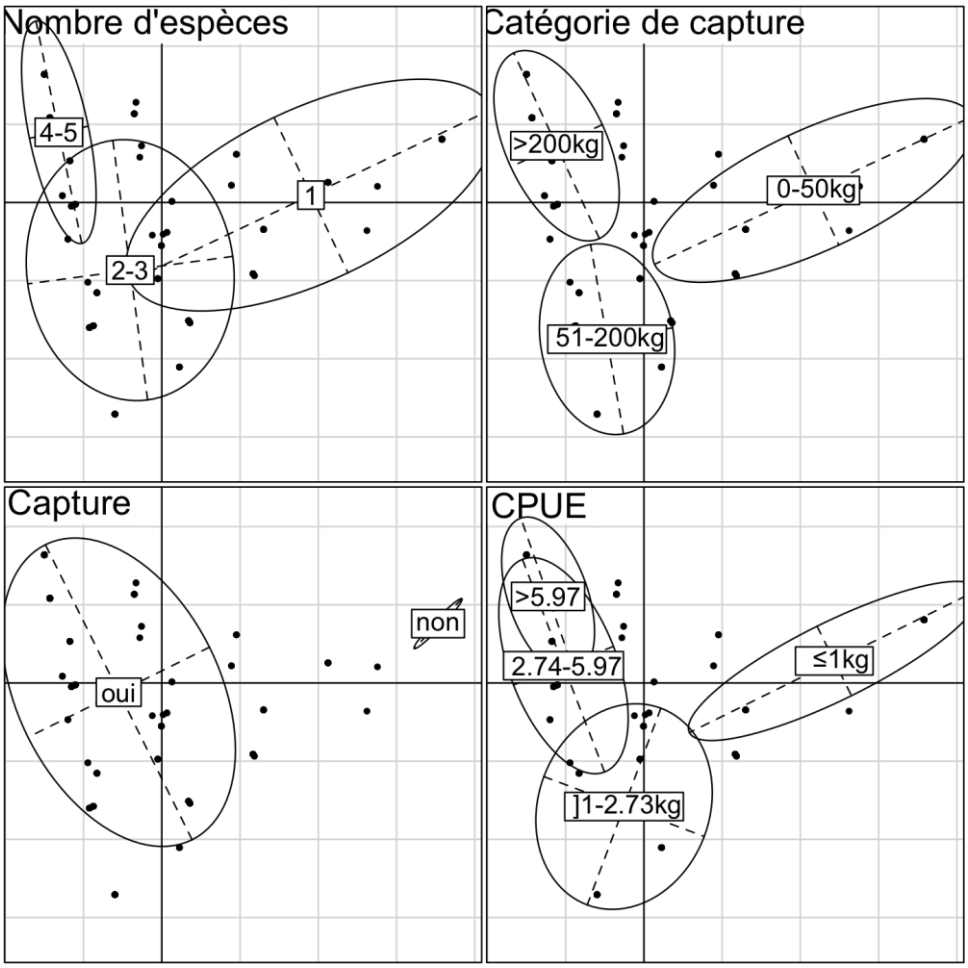


Figure n°III-15: Projection des individus et des modalités pour chaque variable de perception sur le plan factoriel de l'ACM

Sur la figure III-15 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

Une corrélation logique apparaît entre les modalités des variables d'impact. En effet plus un pêcheur aura une CPUE élevée, plus la biomasse prélevée à l'année et le nombre d'espèces capturées seront important. De plus les pêcheurs qui n'avait pas de capture le jour de l'enquête sont aussi ceux qui ont un impact faible sur la ressource (biomasse, CPUE et nombre d'espèces).

Sur la figure III-16 est représentée l'évolution des valeurs de Δi en fonction du nombre

d'individus de l'échantillon du thème perception. Le calcul de Δi ne peut être fait qu'à partir de 175 individus pour garantir la représentativité de chacune des modalités par au moins 1 individus. On observe que Δi est en constante diminution sans que l'on puisse observer de stabilité dans la structure mise en évidence. Il est donc nécessaire d'échantillonner au moins 267 pêcheurs pour avoir une représentation fiable des individus sur le plan factoriel et garantir la structure mise en évidence.

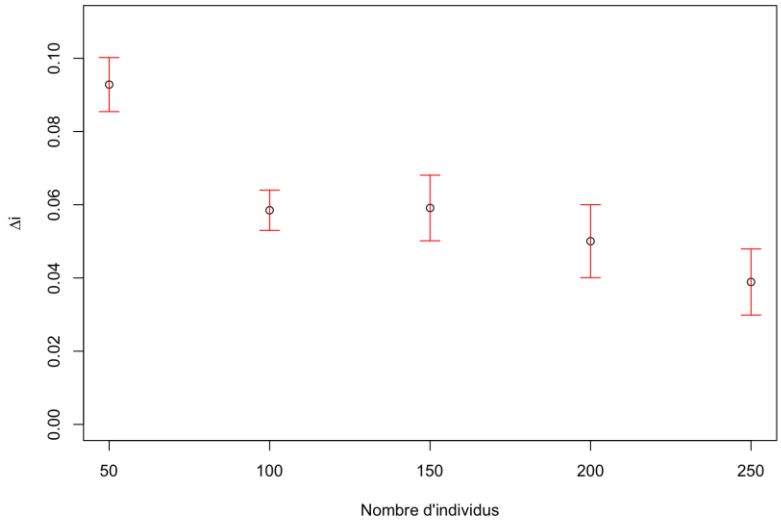


Figure n°III-16: Évolution de la différence moyenne des matrices de corrélation issues des ACM du thème perception de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (425,24). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 267 individus.

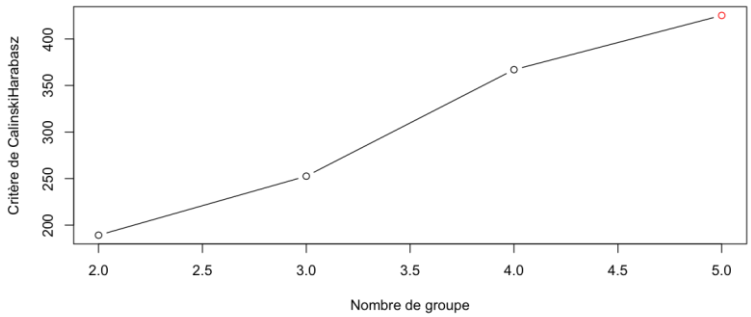


Figure n°III-17: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

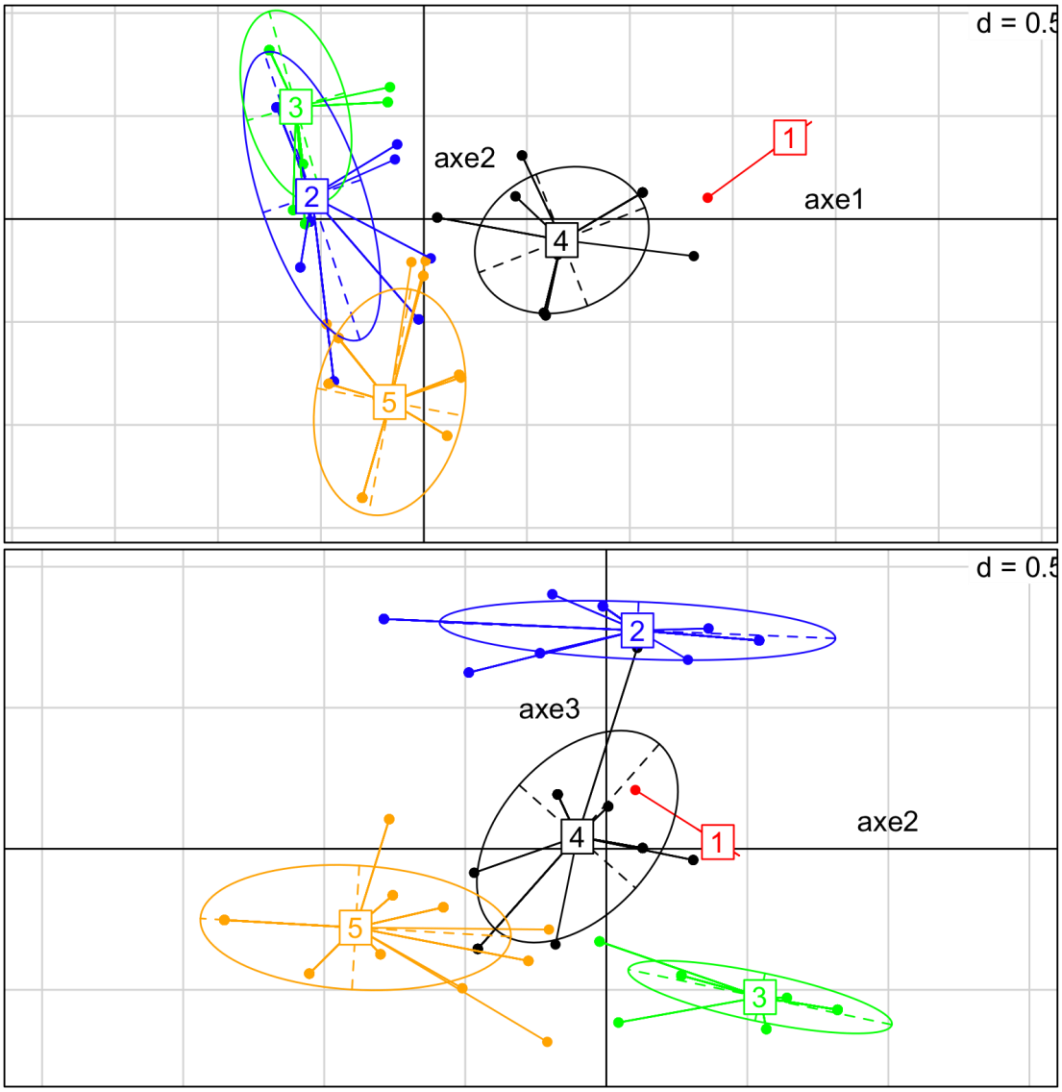


Figure n°III-18: Projection des individus par groupe de perception (points) sur les plans factoriels de l'ACM

Les groupe 1, 3 et 5 sont ceux qui se distingue le plus de l'échantillon total tandis que le groupe 4 est celui qui en est la plus semblable (figure III-18). Le groupe 2 et plus particulièrement le groupe 3 ont un impact important sur la ressource. Les groupes 1 et 4 ont un impact faible sur la ressource. La présence de capture le jour de l'enquête est la variable qui discrimine le moins les groupes d'impact entre eux et avec l'échantillon total (tableau III-7).

Tableau n°III-7: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de perception

	Description de l'échantillon et des groupes			Groupes				
Groupes	*Nombre d'individus*	*Variables*	*Modalités principales*	*1*	*2*	*3*	*4*	*5*
Échantillon	*267*	*Nombre d'espèces*	2 - 3: 37%					
		Catégorie de capture	>200kg: 48%					
		Capture	Oui: 87%					
		CPUE	≤1kg; 1 - 2.73kg; 2.74 - 5.97kg; >5.97kg: 25%					
Groupe 1: *"Très faible impact"*	*35*	*Nombre d'espèces*	1 : 97%					
		Catégorie de capture	0 - 50 kg: 100%					
		Capture	Non: 100%					
		CPUE	≤1kg: 100%					
Groupe 2: *"Fort impact"*	*65*	*Nombre d'espèces*	4 - 5: 52%					
		Catégorie de capture	>200kg: 71%					
		Capture	Oui: 100%					
		CPUE	2.74 - 5.97: 100%					
Groupe 3: *"Fort impact et pêcheurs performant"*	*68*	*Nombre d'espèces*	4 - 5: 66%					
		Catégorie de capture	>200kg: 93%					
		Capture	Oui: 100%					
		CPUE	>5.97: 91%					
Groupe 4: *"Faible impact"*	*38*	*Nombre d'espèces*	1: 58%					
		Catégorie de capture	0 - 50 kg: 79%					
		Capture	Oui: 97%					
		CPUE	≤1kg: 76%					
Groupe 5: *"impact moyen"*	*61*	*Nombre d'espèces*	2 - 3: 64%					
		Catégorie de capture	51 - 200kg: 66%					
		Capture	Oui: 100%					
		CPUE	1 - 2.73: 89%					

La présence de "capture" (le jour de l'enquête) est la variable qui discrimine la moins les groupes entre eux. De plus il s'agit de la variable qui ne discrimine pas chacun des groupes avec l'échantillon total.

Sélection des métriques

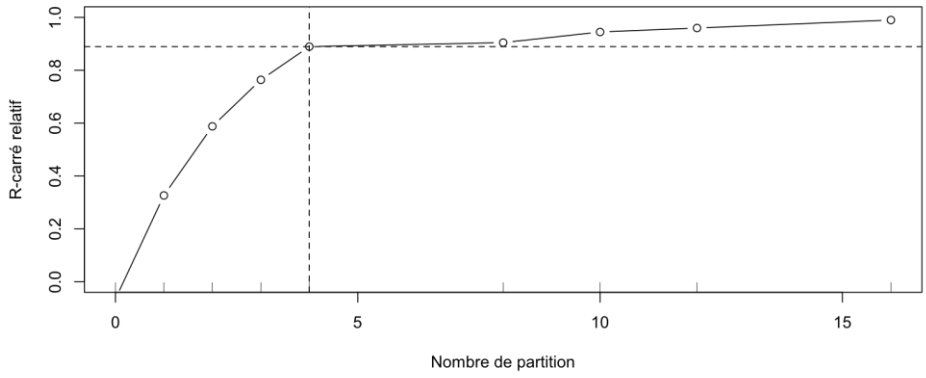


Figure n°III-19: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de partition

À partir de 4 partitions de l'échantillon, le R-carré relatif de l'arbre de segmentation atteint pratiquement son maximum. Il suffit donc de 6 noeuds terminaux pour construire

l'arbre de segmentation pour prédire l'appartenance des individus aux groupes de pratique à partir des variables du même thème.

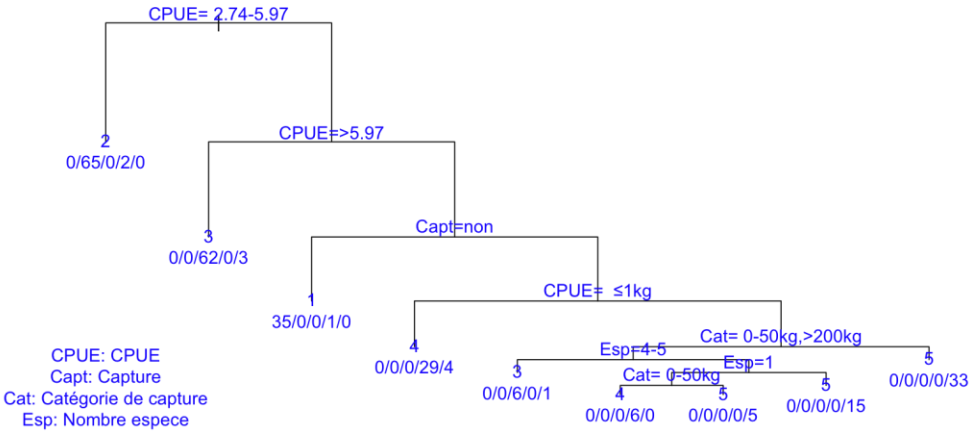


Figure n°III-20: Arbre de reclassification pour les groupes d'impact en fonction des variables du même thème

Les résultats obtenus sont cohérents avec le descriptif des groupes dans le tableau III-7. Les groupes 2 et 3 présentent les individus qui ont l'impact le plus important sur la ressource, tandis que les individus des groupes 1 et 4 ont l'impact le plus faible.

La CPUE (la biomasse moyenne capturée par heure) est la métrique prioritaire a analyser comme indicateur potentiel d'impact.

Mise en évidence d'effet

La variable à expliquer est le log de la CPUE à laquelle on a ajouté 0,275. Cette transformation en log + 0,275 est effectuée afin que les résidus des modèles suivent une loi Normale puisque sans cette transformation, la totalité des modèles est rejetée.

Tableau n°III-8: détails des facteurs sélectionnés et des paramètres statistiques de chacun des modèles construits

Modèles	Facteurs explicatifs	Paramètres	
Modèle 1	**Type de jour** (jour de l'enquête); **zone de pêche PAMPA** (n°1); **saison** (jour de l'enquête); **activité habituelle** (n°1); **groupe du lagon**	Nombre d'individus	252
		Erreur résiduelle standard	1.08 / 233df
		R2	0.2437
		AIC	57.14
		Cp de Mallow	414.99
Modèle 2	**Type de jour** (jour de l'enquête); zone de pêche PAMPA (n°1); **saison** (jour de l'enquête); **activité habituelle** (n°1); groupe du lagon; **typologie de pression**; **Typologie de socio-économie**	Nombre d'individus	230
		Erreur résiduelle standard	1.008 / 203df
		R2	0.3653
		AIC	28.95
		Cp de Mallow	313.22
Modèle 3	**Type de jour** (jour de l'enquête); zone de pêche PAMPA (n°1); **saison** (jour de l'enquête); **activité habituelle** (n°1); groupe du lagon; **typologie de pression**; **Typologie de socio-économie**	Nombre d'individus	230
		Erreur résiduelle standard	1.008 / 203df
		R2	0.3653
		AIC	28.95
		Cp de Mallow	313.22

Les effets des facteurs sont mis en évidence avec une ANOVA de type III. Les **facteurs en gras** ont un effet significatif en dessous du seuil de 0,05, les facteurs soulignés ont un effet significatif en dessous du seuil de 0,1.

Les Modèles 2 et 3 sont identique et sont les modèles conservés par la méthode car ils ont l'AIC et le Cp de Mallow les plus faible. On constate que les typologies de pression et de socio-économie on un effet sur la capture moyenne par heure des pêcheurs (p-value<0.05). De plus la métrique d'impact est principalement fonction de l'activité de pêche habituelle pratiquée (chasse sous-marine ou pêche embarquée), de la saison et du type de jour où l'enquête est réalisée (p-value<0.05). Enfin et dans une moindre mesure, on observe une spatialisation entre groupe de lagon de l'impact de la pêche sur les ressources avec un effet du groupe de lagon sur la CPUE (p-value<0.1). La zone de pêche selon le statut PAMPA (Z1,Z2, …) est conservée par le modèle sans pour autant avoir un effet sur la métrique d'impact.

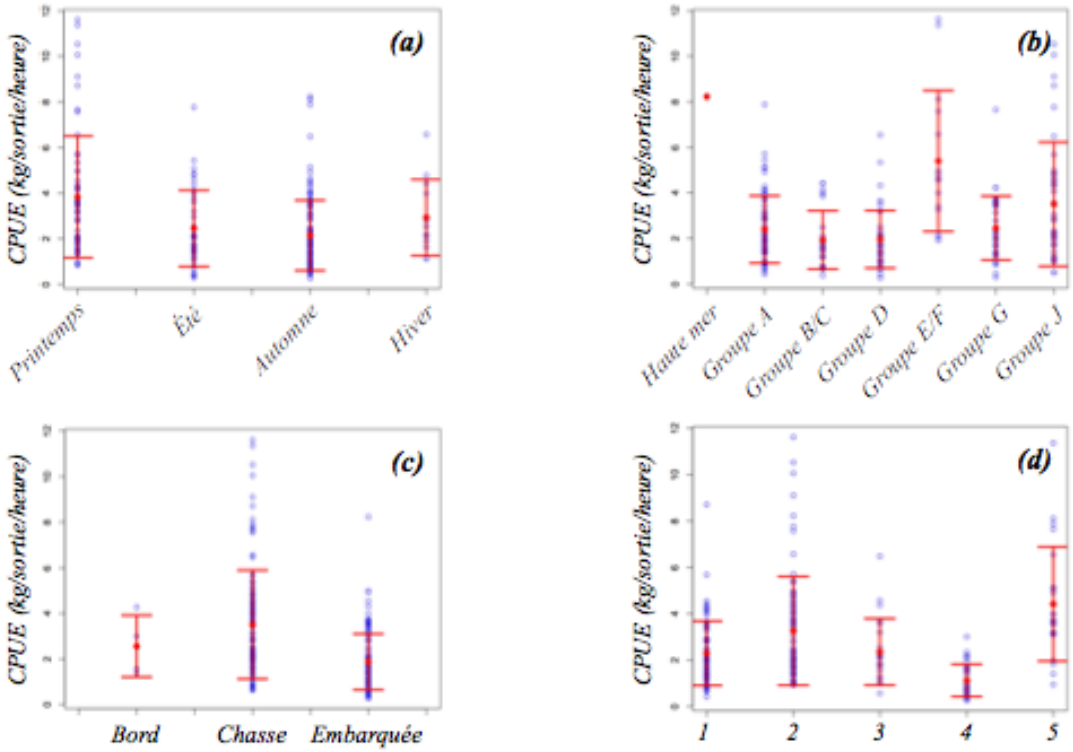


Figure n°III-21: Graphiques représentante les valeurs prédites de la CPUE, par le modèle en fonction de la saison (a), de la zone de pêche (b), de l'activité (c) et du groupe de pression (d)

Les effets mis en évidence des variables saison, zone de pêche (groupe de lagon), activité et typologie de pression ont des puissances statistiques supérieures à 0.8. On peut donc conclure, de manière fiable, à un effet de ces variables sur l'indicateur d'impact. Néanmoins l'effet spatial (par groupe de lagon) sur la CPUE présente un seuil de significativité qui est compris entre 0.05 et 0.1.

Les puissances statistiques des tests mettant en évidence l'effet des groupes de socio-économie et du type de jour sont inférieures à 0.8 ce qui ne permet pas de conclure de manière fiable à un effet sur la CPUE. Il faudrait que les tailles d'effets de ces facteurs soient égales à 23% pour les groupes de socio-économie (contre 21% actuellement) et à 22% pour les types de jours (contre 11% actuellement) pour que la puissance statistique de ces tests soit acceptable.

Conclusion

On a mis en évidence 5 groupes d'impact parmi l'échantillon en fonction du nombre d'espèces, de la catégorie de capture, de la présence de capture et de la CPUE . La CPUE est la variable qui reflète le mieux les groupes d'impact. Elle s'avère être donc l'indicateur le

plus pertinent de la période de pêche (saison et type de jour) et du type d'activité. Il faudrait une différence de la durée de pêche dans et hors la réserve pour mettre en évidence un effet significatif du zonage. La taille de l'échantillon a été réduite à 230 personnes enquêtées, il est donc possible que l'effet de zonage aurait pu être mis en évidence en considérant la totalité de l'échantillon.

E. Thème Socio-économie

Tableau n°III-9: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
E3_CSP	Actif	actif	641
	Agriculteur		
	Cadre		
	Artisan		
	Employé		
	Intermediaire		
	liberal		
	ouvrier		
	Étudiant	Pas actif	63
	retraité		
	Pas actif		
E2_an_nais	Quantitatif	≤1956	187
		1957-1972	356
		>1972	161
E4_commune	Détail des commune	Nouméa	442
		Autre	262

Pour chacune des variables utilisés, une recodage des données est effectué de manière à avoir un nombre de modalité par variable compris entre 2 et 4. Ce recodage doit permettre de conserver une cohérence de l'information afin d'être pertinent en temps qu'indicateur de socio-économie.

Typologie

Pour l'ACM du thème socio-économie, la première décroissance de valeurs propre apparaît entre le 2e et le 3e axe (4% de décroissance), De plus les deux premier axes présentent des corrélations de plus de 50% avec au mois une variable. De ce fait on ne

conserve que les deux premiers axes pour construire l'ACM de socio-économie.

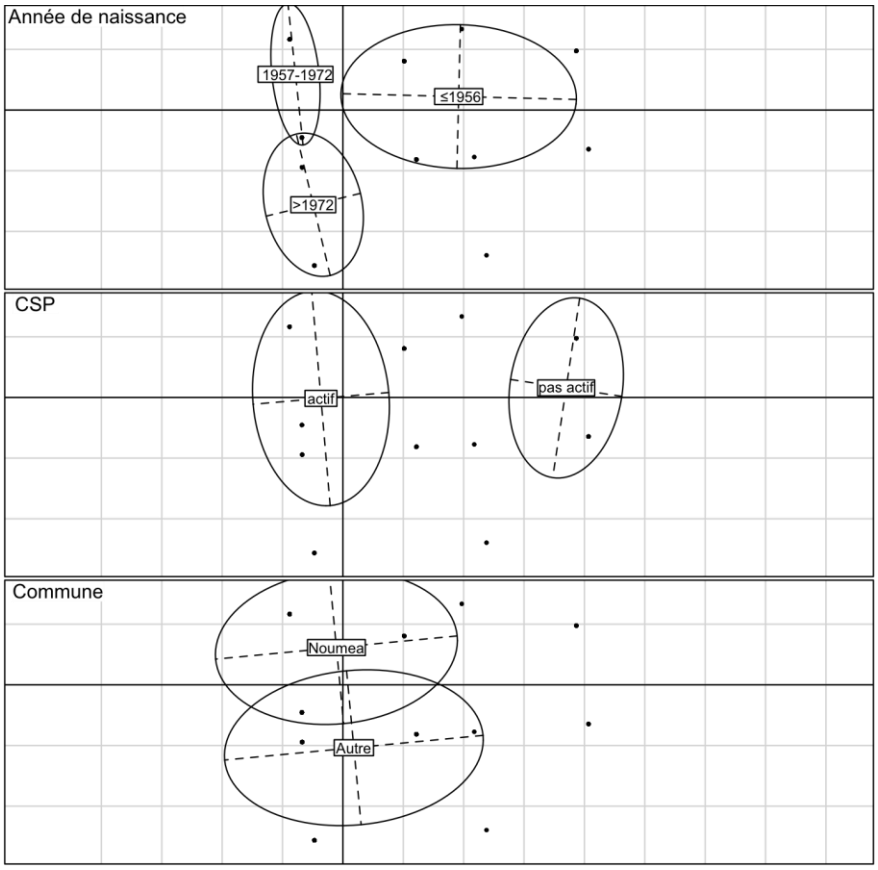


Figure n°III-22: Projection des individus et des modalités pour chaque variable de socio-économie sur le plan factoriel de l'ACM

Sur la figure III-22 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

Les individus résident à Nouméa ont tendance à avoir plus de 35 ans tandis que les moins de 35 ans ont tendance à habiter dans d'autres communes. La classe d'âge de plus de 50 ans (qui comporte tous les individus retraités) comporte principalement des personnes non-active .

Sur la figure III-23 est représentée l'évolution des valeurs de Δi en fonction du nombre d'individus de l'échantillon du thème perception. On observe une stabilité de la moyenne de Δi entre 350 et 700 individus. Il est donc suffisant d'échantillonner 350 pêcheurs pour avoir

une représentation fiable des individus sur le plan factoriel et garantir sa structure.

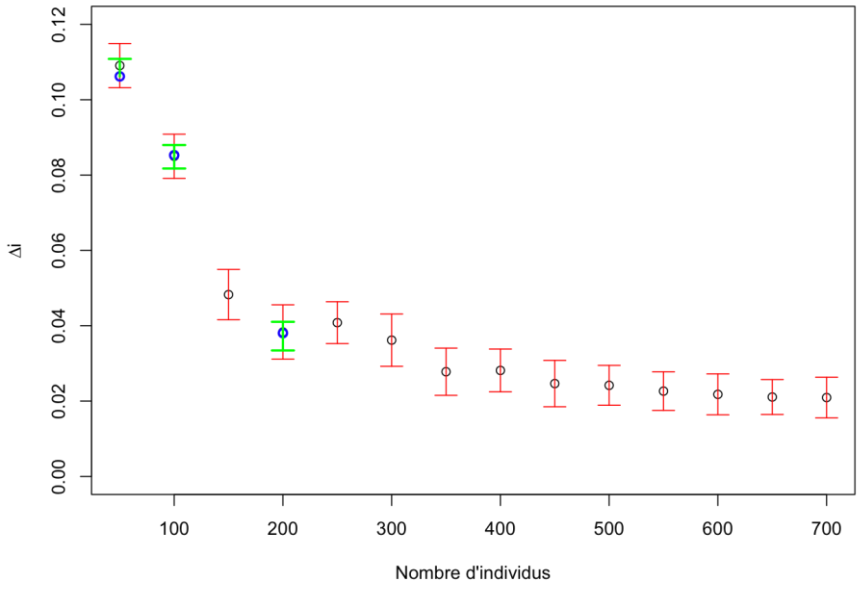


Figure n°III-23: Évolution de la différence moyenne des matrices de corrélation issues des ACM du thème socio-économie de l'échantillon total et des sous échantillons en fonction du nombre d'individus

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (2225,72). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 704 individus.

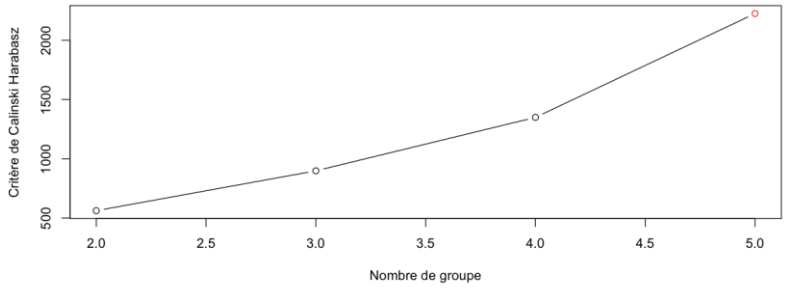


Figure n°III-24: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

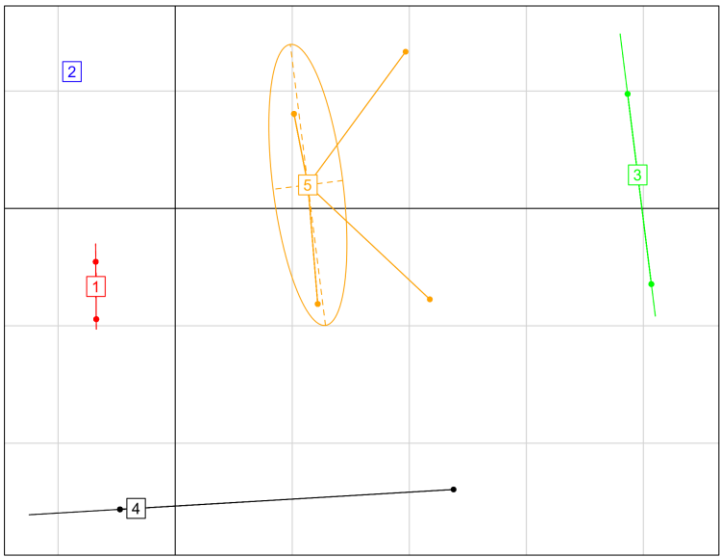


Figure n°III-24: Projection des individus par groupe de socio-économie (points) sur le plan factoriel de l'ACM

Les groupes 3 et 4 sont ceux qui se distinguent le plus de l'échantillon total (figure III-24). Le groupe 3 est le seul qui se caractérise réellement vis à vis de la CSP. Une combinaison entre l'année de naissance et la commune d'habitation permet de discriminer les groupes 1, 2 et 5.

L'année de naissance est la variable qui discrimine le mieux les groupes entre eux et les groupes avec l'échantillon total ce qui peut être dû à un nombre de modalités plus important que pour les autres variables (tableau III-8).

Tableau n°III-8: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de socio-économie

	Description de l'échantillon et des groupes			Groupes				
Groupes	*Nombre d'individus*	*Variables*	*Modalités principales*	*1*	*2*	*3*	*4*	*5*
Échantillon	*704*	*Année de naissance*	1957 - 1972: 51%					
		CSP	Actif: 91%					
		Commune	Nouméa: 63%					
Groupe 1: *"Habitant d'une autre commune que Nouméa âgé entre 48 et 33 ans"*	*223*	*Année de naissance*	1957 - 1972: 57%					
		CSP	Actif: 100%					
		Commune	Autre: 57%					
Groupe 2: *"Habitant de Nouméa âgé entre 48 et 33 ans"*	*226*	*Année de naissance*	1957 - 1972: 100%					
		CSP	Actif: 100%					
		Commune	Nouméa: 100%					
Groupe 3: *"Non actif"*	*54*	*Année de naissance*	≤1956: 100%					
		CSP	Pas actif: 100%					
		Commune	Nouméa: 57%					
Groupe 4: *"Pêcheur de moins de 33 ans"*	*62*	*Année de naissance*	>1972: 100%					
		CSP	Actif: 95%					
		Commune	Autre: 100%					
Groupe 5: *"Habitant de Nouméa de plus de 48 ans"*	*139*	*Année de naissance*	≤1956: 96%					
		CSP	Actif: 96%					
		Commune	Noumea: 57%					

Sélection des métriques

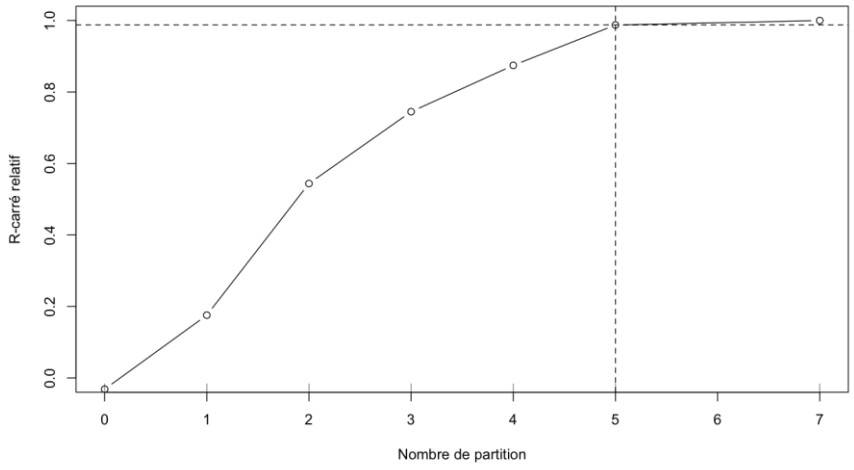


Figure n°III-25: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de noeud terminaux

À partir de 5 partitions de l'échantillon, le R-carré relatif de l'arbre de segmentation atteint pratiquement son maximum (figure III-25). Il suffit donc de 6 noeuds terminaux pour construire l'arbre de segmentation pour prédire l'appartenance des individus aux groupes de

socio-économie à partir des variables du même thème.

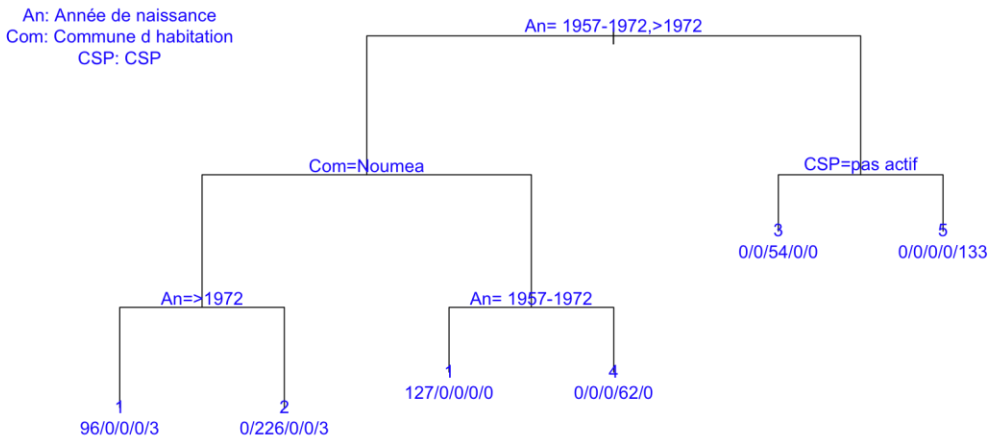


Figure n°III-26: Arbre de reclassification pour les groupes de pression en fonction des variables du même thème

L'année de naissance apparaît par la méthode de segmentation comme la métrique prioritaire à analyser comme indicateur potentiel. La deuxième métrique prioritaire a analyser est une combinaison des variable "année de naissance", "commune" et "CSP" en prenant comme borne des classes les valeurs seuils indiquées par les 3 premières partitions de l'arbre.

Conclusion

On a mis en évidence 5 groupes de socio économie parmi l'échantillon en fonction de l'année de naissance, de la catégorie socio professionnelle, et de la commune d'habitation du pêcheur. L'année de naissance du pêcheur est la variable qui reflète le mieux les groupes de pression. Elle s'avère être donc la métrique le plus pertinente à tester en tant qu'indicateur potentiel.

Conclusion du cas d'étude

Pour chacun des thèmes étudiés, un indicateur potentiel a pu être mis en évidence via la construction de typologies. Pour le thème pression, le rapport durée de pêche / durée de sortie a été identifié comme la métrique la plus pertinente a étudier comme indicateur potentiel, pour le thème pratique, il s'agit de la planification à la semaine, pour le thème impact, de la

biomasse prélevée par heure et enfin, pour le thème socio-économie, l'année de naissance.

La typologie de pression et de socio économie est un facteur explicatif qui a un effet significatif sur l'indicateur d'impact. Ainsi l'efficacité du pêcheur en terme de biomasse prélevée par heure, et plus généralement son impact sur le milieu dépend de son appartenance à un groupe de pression et un groupe de socio économie donné, autrement dit de son effort de pêche et de ses caractéristiques socio-économiques. L'injection de certaines typologies d'autres thèmes, en tant que facteur explicatif pour prédire un indicateur potentiel donnée, dans un modèle statistique permet :

- d'augmenter la prédiction du modèle
- de normaliser la distribution des résidus du modèle
- mettre en évidence des effets d'intérêt

IV. Cas de la Réunion

La réserve naturelle marine de la Réunion (RNMR) a été créée fin 2007 sur le côté ouest de l'île et remplace l'ancien parc marin créé en 1997. Elle s'étend sur un linéaire de côte de 40 km et sur une surface de 3500 ha environ (Annexe 1). Au sein de la réserve sont définis 3 niveaux de protection.

Les données traitées sur la RNMR sont issues d'un travail de thèse réalisé par A. Thomassin concernant les usages à la Réunion. Elles ont été recueillies à partir d'entretiens semi-directifs réalisés en 2009 auprès de pêcheurs récréatifs ayant une embarcation dans l'un des 3 ports situés au sein de la RNMR (Étangs-Salés, Saint-Gilles et Saint-Leu). La population étudiée ne comporte donc que des individus pratiquant la pêche plaisancière embarquée.

L'échantillon comporte 50 pêcheurs embarqués qui ont été enquêtés entre le 15/04/09 et le 15/06/09. seulement 26% des variables conformes au format PAMPA (36 variables) ont pu être renseignées. Ces données ont servi principalement à mettre au point: 1) la méthode de sélection des données pour décrire l'échantillon avec des typologies d'usagers; 2) la méthode de sélection des métriques à tester en tant qu'indicateurs d'un effet AMP.

A. Identification des thèmes

La sélection des données (nombre de variables et nombre d'individus) est présentée pour chaque thème dans le tableau IV-1.

Tableau n°IV-1: Tableau récapitulatif du nombre d'individus et du nombre de variables sélectionnées pour la construction du plan factoriel de chaque thème

Thèmes	Nombre d'individus	Nombre de variables
Pression	**50**	**3**
Pratique	**50**	**4**
Perception	**50**	**5**
Socio-économie	**50**	**3**

B. Thème Pression

Étant donné la présence de valeurs extrêmes pour une ou plusieurs des quatre variables du thème pression, 8 individus sont écartés pour la construction du plan factoriel de ce thème.

Typologie

Lors de la construction du plan de l'ACP pour le thème pression, les deux premiers axes rassemblent 78% de l'inertie du nuage de points. De plus ils permettent de garantir une réelle signification de ces axes avec des corrélations supérieures à 50% avec au moins une variable. Le premier axe factoriel est négativement corrélé à plus de 50% avec le nombre d'engins et positivement avec le nombre de pêcheurs. L'axe 2 est corrélé positivement et à plus de 50% avec le la durée de pêche. L'axe 1 résume l'intensité de la pression de pêche en termes de moyens déployés tandis que l'axe 2 reflète la pression de pêche par en termes de temps consacré à l'activité (figure IV-1).

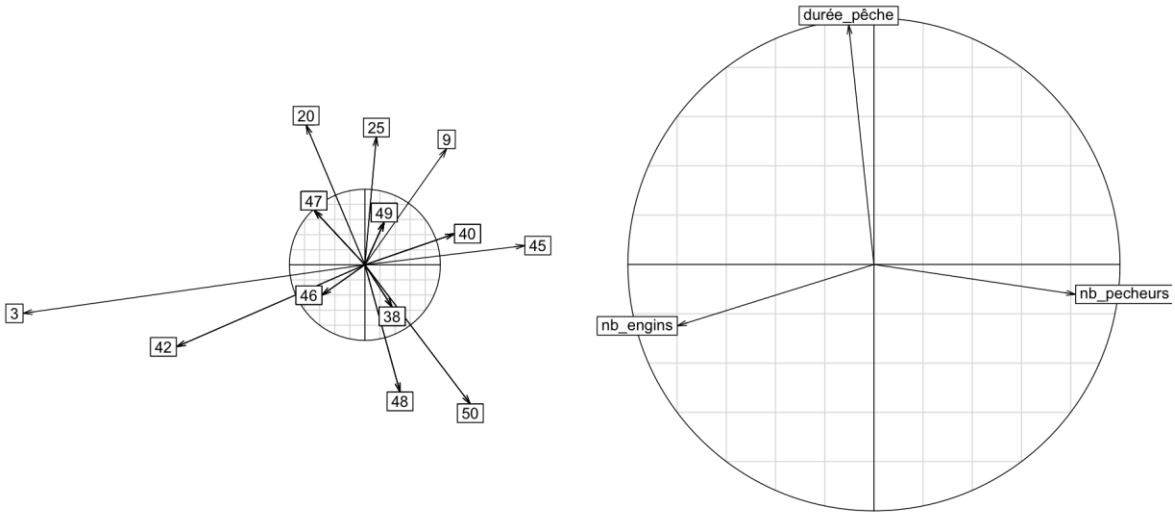


Figure n°IV-1: projection des individus sur le plan (gauche) et cercle de corrélation (droite) de l'ACP du thème pression

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (63,82). On choisit donc de former 5 groupes de pression parmi l'échantillon de 50 individus.

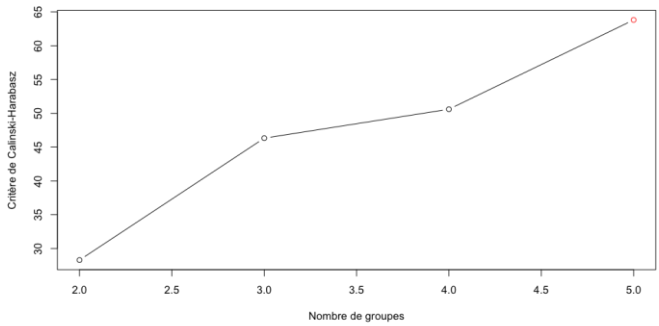


Figure n°IV-2: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

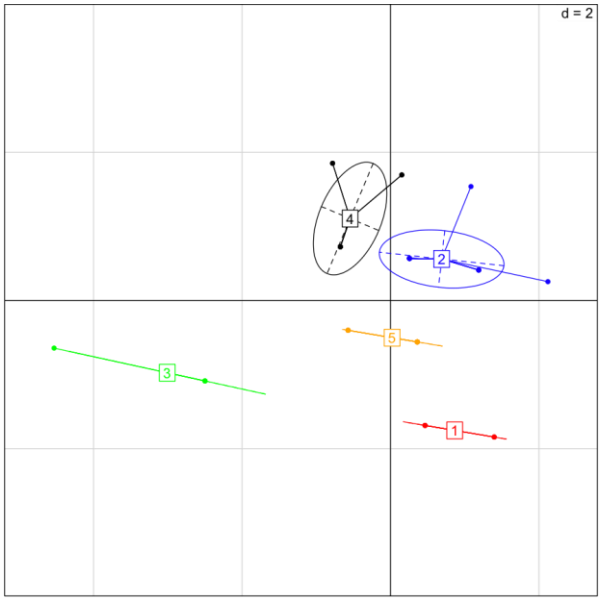


Figure n°IV-3: Projection des individus par groupe de pression (points) sur le plan factoriel

Le groupe 1 et 5 rassemble les individus qui ont la plus faible durée de pêche (figure IV-3). Les groupes 2 et 4 réunissent les individus qui présentent une durée de pêche élevée, néanmoins ils se distinguent par leur nombre de pêcheurs. Le groupe 3 compte dans son effectif des individus qui utilisent un nombre d'engins élevé (figure IV-3 et tableau IV-2).

Tableau n°IV-2: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de pression

	Description de l'échantillon et des groupes						Groupes				
Groupes	*Nombre d'individus*	*Variables*	*moy.*	*med.*	*sd*	*C.V.*	*1*	*2*	*3*	*4*	*5*
Échantillon	*50*	*Nombre de pêcheurs*	1.88	2	0.77	41.12					
		Nombre d'engins	1.1	1	0.36	33.11					
		Durée de pêche	5.54	6	0.86	15.56					
Groupe 1: *"Durée de pêche très faible"*	*7*	*Nombre de pêcheurs*	2.43	2	0.53	22					
		Nombre d'engins	1	1	0	0					
		Durée de pêche	**4**	4	0	0					
Groupe 2: *"Nombre de pêcheurs et durée de pêche élevée"*	*17*	*Nombre de pêcheurs*	**2.47**	2	0.62	25.27					
		Nombre d'engins	1	1	0	0					
		Durée de pêche	**6.06**	6	0.24	4					
Groupe 3: *"Grand nombre de types d'engin"*	*4*	*Nombre de pêcheurs*	**1**	1	0	0					
		Nombre d'engins	**2.25**	2	0.5	22.22					
		Durée de pêche	5.25	5	0.5	9.52					
Groupe 4: *"Peu de pêcheurs et durée de pêche élevée"*	*11*	*Nombre de pêcheurs*	**1.18**	1	0.40	34.22					
		Nombre d'engins	1	1	0	0					
		Durée de pêche	**6.36**	6	0.5	7.92					
Groupe 5: *"Durée de pêche faible"*	*11*	*Nombre de pêcheurs*	1.64	2	0.5	30.83					
		Nombre d'engins	1	1	0	0					
		Durée de pêche	**5**	5	0	0					

Les groupe 1 et 5 sont ceux qui se distinguent le moins de l'échantillon total. Le nombre d'engins est la variable qui discrimine le moins les groupes de pression entre eux ainsi qu'entre les groupes et l'échantillon total.

Sélection des métriques

On n'observe pas d'augmentation du R^2 relatif pour un nombre de noeuds terminaux supérieur à 5. On choisit donc de conserver uniquement 5 noeuds terminaux pour l'arbre de reclassification du thème pression.

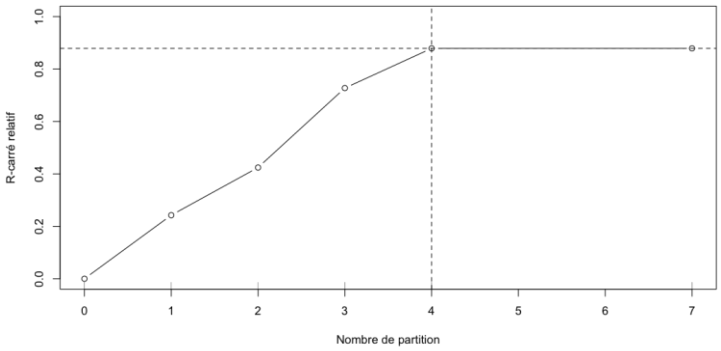


Figure n°IV-4: Évolution de la différence moyenne des matrices de corrélation issues des ACP du thème pression de l'échantillon total et des sous échantillons en fonction du nombre d'individus

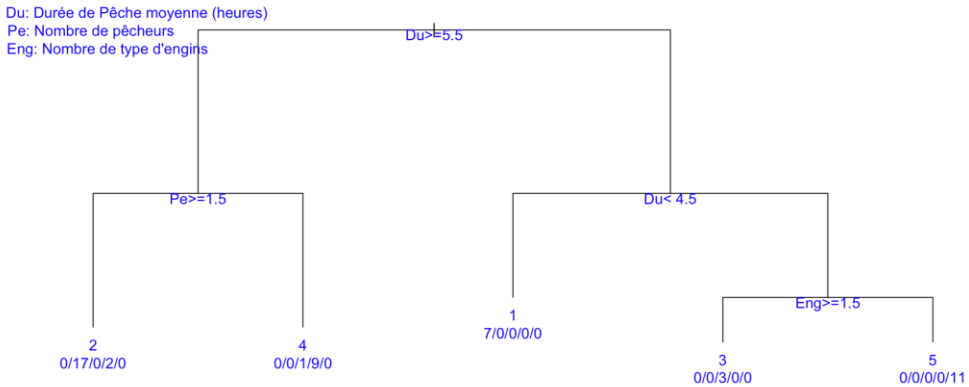


Figure n°IV-5: Arbre de reclassification pour les groupes de pression en fonction des variables du même thème

Sur la figure IV-5, sont indiquées, au niveau des noeuds de l'arbre, les variables et valeurs seuils associées pour la partition de l'échantillon tandis qu'au niveau des noeuds terminaux est représenté le numéro de groupe associé. En-dessous de ce numéro est indiqué

le nombre d'individus de chaque groupe qui compose le noeud terminal (nombre d'individus du groupe n°1 / nombre d'individus du groupe n°2 / … / nombre d'individus du groupe n°5). Cette indication permet de nuancer les résultats de la segmentation pour des noeuds terminaux dont la composition est diverse.

La durée de pêche est la variable qui discrimine le mieux les groupes de pression (figure IV-5). Il s'agit donc de la première métrique à tester comme indicateur potentiel. Si il est nécessaire de considérer une combinaison de variables, on choisira la combinaison des variables "durée de pêche" et "nombre de pêcheurs", en prenant comme borne des classes les premières valeurs seuils générées par l'arbre de reclassification.

Conclusion

On a mis en évidence 5 groupes de pression parmi l'échantillon en fonction du nombre de pêcheurs, du nombre d'engins et de la durée de pêche. La durée de pêche est la variable qui reflète le mieux les groupes de pression. Elle s'avère être donc la métrique le plus pertinente à tester en tant qu'indicateur potentiel.

C. Thème Pratique

Tableau n°IV-3: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
C4_nb_sortie_tot	Quantitatif	≤36	17
		[37-72]	17
		≥73	16
C2_plannif_an	Annee	Annee	26
	Été	Été	24
C2_plannif_sem	JS	PAS IND	22
	JW		
	JW+VAC		
	VAC		
	IND	IND	28
C2_plannif_jour	AM	AM	34
	PM	PM/SOIR	3
	SOIR		
	IND	IND	13

Les variables de pratiques sont codées pour obtenir un nombre de modalités par variable compris entre 2 et 4 avec une représentativité des modalités le plus homogène possible. Cependant il est important de conserver un niveau d'information pertinente afin de

permettre à chaque variable d'être un indicateur de pratique pertinent.

Typologie

Pour l'ACM du thème pratique, la décroissance des valeurs propres entre le 5e et le 6e axe est de 6%, on conserve donc les cinq premiers axes. Néanmoins l'axe 2 et l'axe 4 présentent de corrélations de plus de 50% avec aucune des variables.

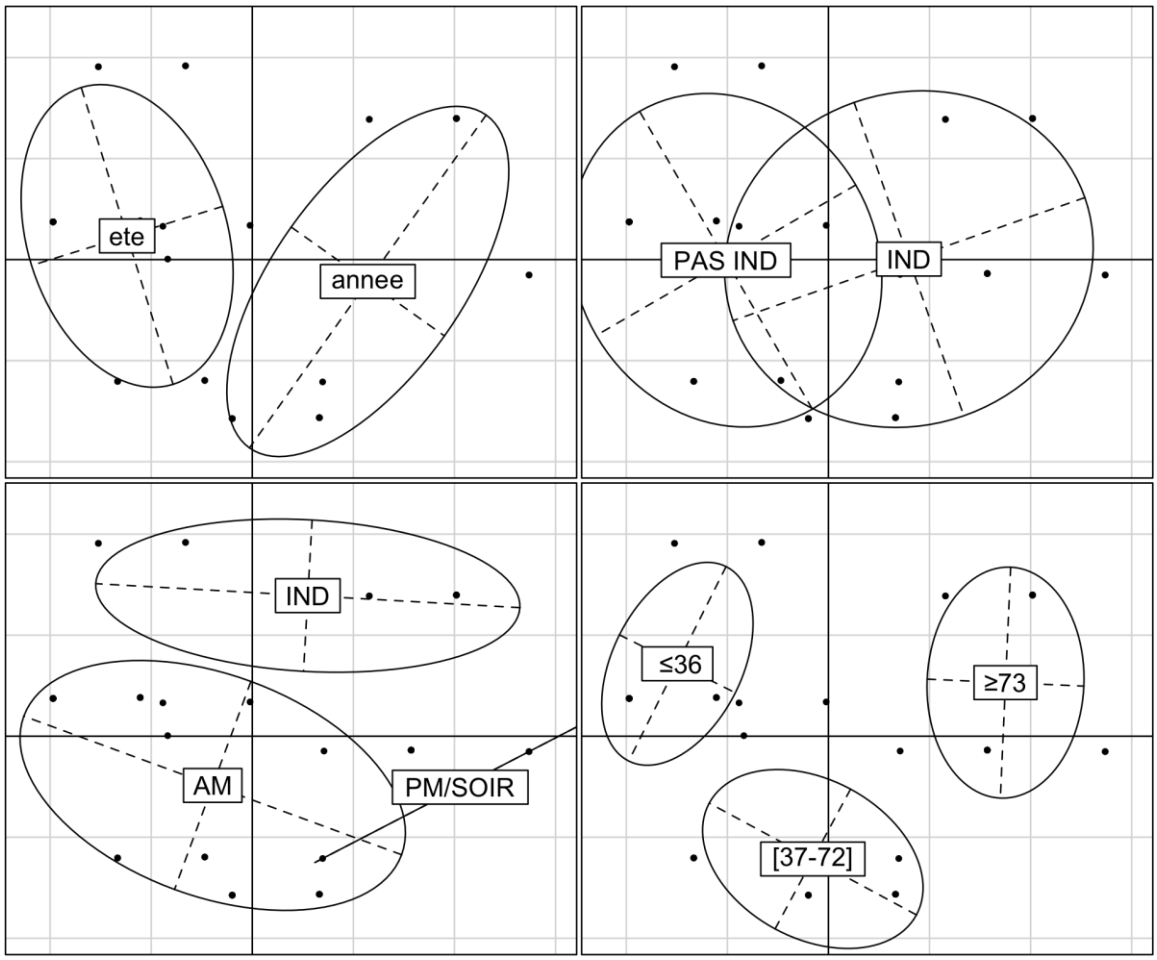


Figure n°IV-6: Projection des individus et des modalités pour chaque variable de pratique sur le plan factoriel de l'ACM

Sur la figure IV-6 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

De manière générale, les pêcheurs qui sortent le plus souvent à l'année sont indifférent au type de jour et à la période de l'année. La planification à la journée ne semble pas être corrélée au nombre de sortie à l'année. Les pêcheurs qui sortent durant la matinée, on une fréquence de sortie à l'année faible à moyenne, tandis que les pêcheurs indifférents à la planification de la sortie de pêche à la journée sortent à l'année à la fois avec une très forte et une très faible fréquence à l'année.

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (28,24). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 50 individus.

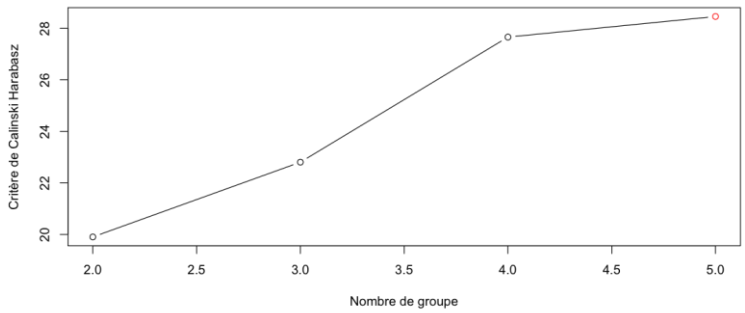


Figure n°IV-7: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

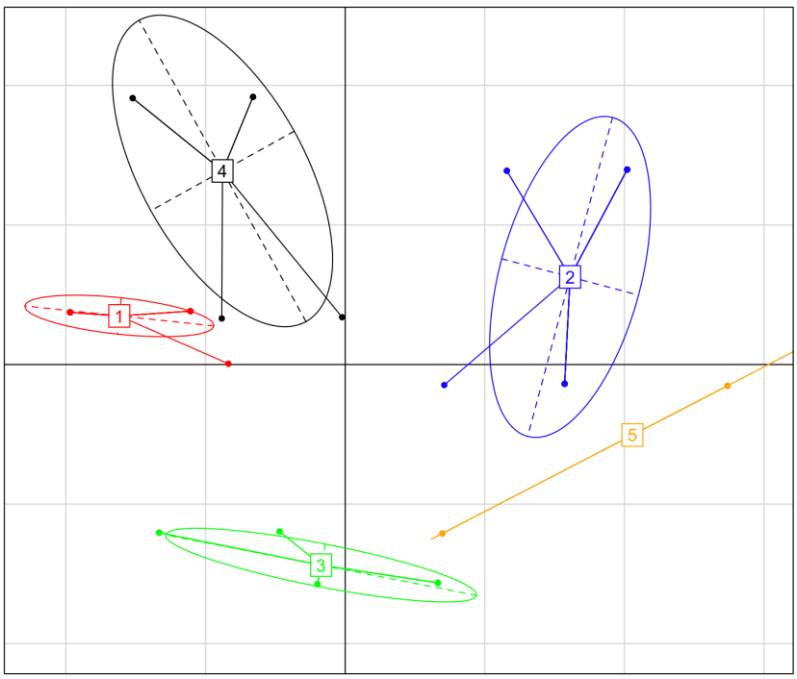


Figure n°IV-8: Projection des individus par groupe de pratique (points) sur le plan factoriel de l'ACM

Les groupes 1 et 4 présentent une pratique faible, le groupe 3 une pratique modérée et les groupes 2 et 5 une pratique intense (figure IV-8). Le nombre de sorties totale à l'année et la planification à l'année sont les variables qui discriminent le mieux les groupes de pratique avec l'échantillon total. Le nombre de sortie à l'année et la planification à la journée sont les variables qui discriminent le plus les groupes entre eux (tableau IV-4).

Tableau n°IV-4: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de pratique

	Description de l'échantillon est des groupes			Groupes				
Groupes	*Nb d'ind.*	*Variables*	*Proportion des modalités*	*1*	*2*	*3*	*4*	*5*
Échantillon	*50*	*Planification à l'année*	Année: 52%					
		Planification à la semaine	Indifférent: 56%					
		Planification à la journée	Matin: 68%					
		Nombre de sortie total	Chacune des modalités: 33%					
Groupe 1: *"Pratique faible ponctuelle"*	*13*	*Planification à l'année*	Été: 92%					
		Planification à la semaine	Pas indifférent: 69%					
		Planification à la journée	Matin: 100%					
		Nombre de sortie total	≤36: 100%					
Groupe 2: *"Pratique intense"*	*14*	*Planification à l'année*	Année: 100%					
		Planification à la semaine	Indifférent: 79%					
		Planification à la journée	Indifférent: 50%					
		Nombre de sortie total	≥73: 100%					
Groupe 3: *"Pratique modérée et matinale"*	*14*	*Planification à l'année*	Année: 64%					
		Planification à la semaine	Indifférent: 50%					
		Planification à la journée	Matin: 100%					
		Nombre de sortie total	37-72: 100%					
Groupe 4: *"Pratique faible et diffuse"*	*6*	*Planification à l'année*	Été: 100%					
		Planification à la semaine	Indifférent: 50%					
		Planification à la journée	Indifférent: 100%					
		Nombre de sortie total	≤36: 67%					
Groupe 5: *"Pratique intense et après-midi"*	*3*	*Planification à l'année*	année: 67%					
		Planification à la semaine	Indifférent: 100%					
		Planification à la journée	Après-midi / soir: 100%					
		Nombre de sortie total	≥73: 67%					

Sélection des métriques

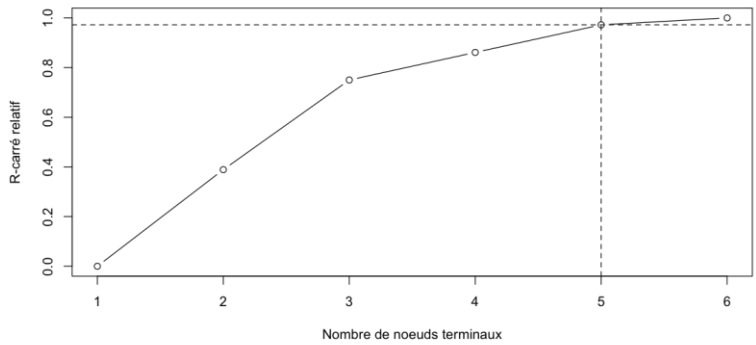


Figure n°IV-9: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de partition

À partir de 5 partitions, le R-carré relatif de l'arbre de segmentation atteint son maximum (figure IV-9). Il suffit donc de 6 noeuds terminaux pour construire l'arbre de segmentation.

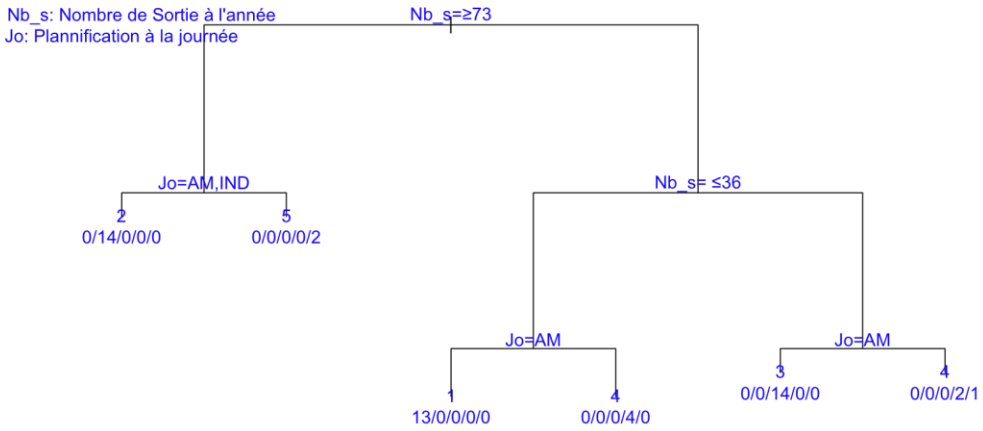


Figure n°IV-10: Arbre de reclassification pour les groupes de pratique en fonction des variables du même thème

Le nombre de sortie à l'année est la variable qui discrimine le mieux les groupes de pratique (figure IV-10). Il s'agit donc de la première métrique à tester comme indicateur potentiel. Si il est nécessaire de considérer une combinaison de variables, on choisira la combinaison des variables "Nombre de sortie à l'année" et "planification à la journée", en prenant comme bornes des classes les premières valeurs seuils générées par l'arbre de reclassification.

Conclusion

On a mis en évidence 5 groupes de pratique parmi l'échantillon en fonction du nombre de sortie à l'année et de la planification de l'activité (à l'année, à la semaine). Le nombre de sorties à l'année est la variable qui reflète le mieux les groupes de pratique. Elle s'avère être donc la métrique le plus pertinente à tester en tant qu'indicateur potentiel.

D. Thème Perception

Tableau n°IV-5: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
D1_connaiss_regR	Oui	Oui	24
	Non	Non	26
	Nsp	Nsp	0 (pas pris en compte dans l'analyse)
D1_respect_reg	Oui	Oui	11
	Non	Non	39
	Nsp	Nsp	0 (pas pris en compte dans l'analyse)
D2_effet_ecosyst	Tres_positif	Positif	33
	Plutot_positif		
	Tres_negatif	Negatif	11
	Plutot_negatif		
	Neutre	Neutre	0 (pas pris en compte dans l'analyse)
	Nsp	Nsp	6
D2_effet_peche	Tres_positif	Positif	6
	Plutot_positif		
	Tres_negatif	Negatif	9
	Plutot_negatif		
	Neutre	Neutre	16
	Nsp	Nsp	19
nb_conflits	quantitatif	≥2	4
		1	22
		0	24

Pour chacune des variables utilisés, une recodage des données est effectué de manière à avoir un nombre de modalité par variable compris entre 2 et 4. Ce recodage doit permettre de conserver une cohérence au niveau de l'information conservée afin d'être pertinent en temps qu'indicateur de perception.

Typologie

Pour l'ACM du thème perception, la première décroissance de valeurs propre apparaît entre le 5e et le 6e axe (9,7% de décroissance), de plus les quatre premiers axes présentent des corrélations de plus de 50% avec au mois une variable. Le cinquième axe est corrélé à presque 50% avec l'avis sur l'effet de l'amp sur la pêche. De ce fait on conserve les cinq premiers axes pour construire l'ACM de perception.

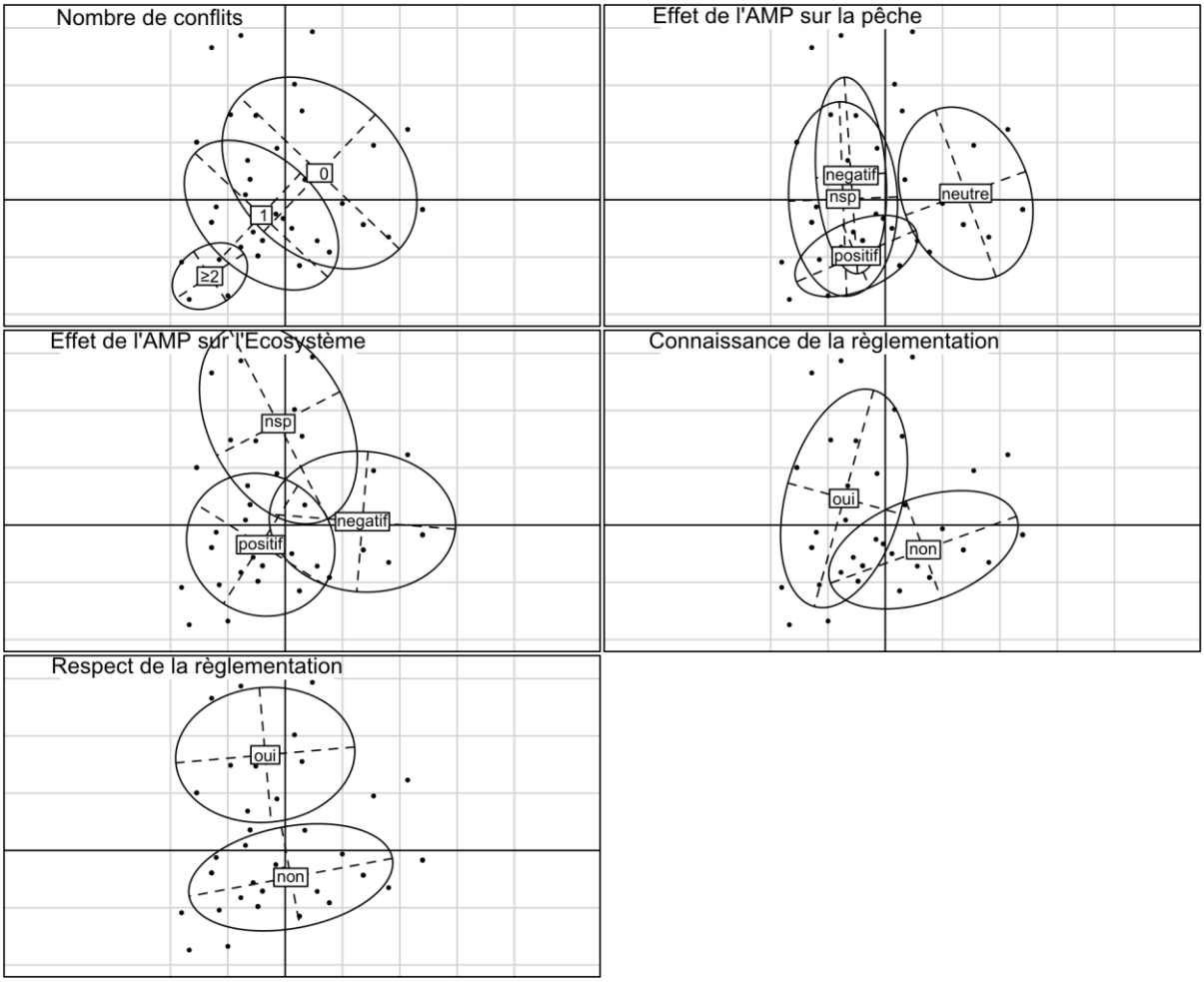


Figure n°IV-11: Projection des individus et des modalités pour chaque variable de perception sur le plan factoriel de l'ACM

Sur la figure IV-11 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (17,65). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 50 individus.

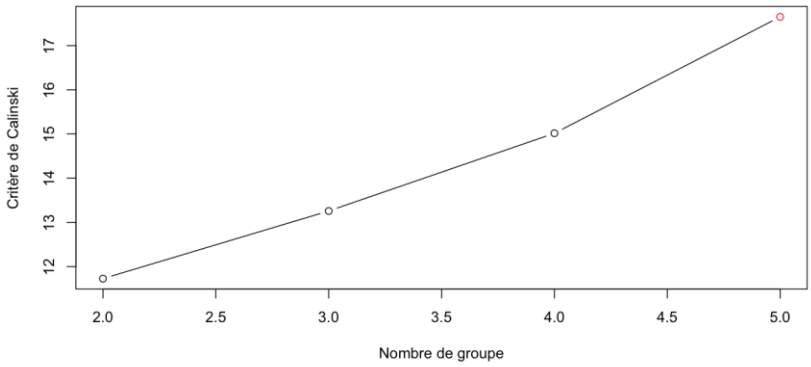


Figure n°IV-12: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

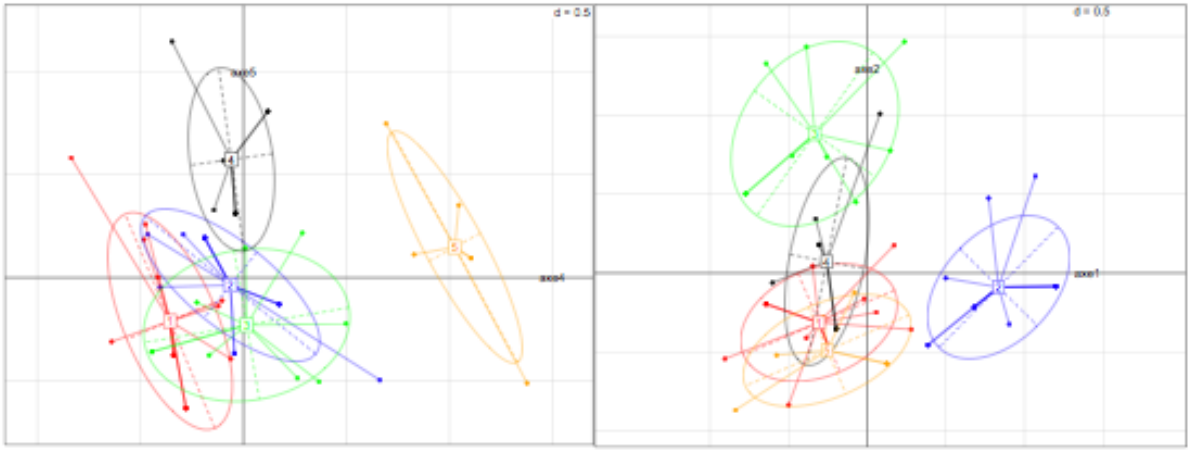


Figure n°IV-13: Projection des individus par groupe de perception (points) sur les plans factoriels de l'ACM

Le groupe 1 est le groupe qui se distingue le moins de l'échantillon total (figure IV-13). Le groupe 2 comporte des pêcheurs qui pensent que l'amp a un effet neutre. Les individus du groupe 3 pensent que la réglementation est respectée. Le groupe 4 rassemble des individus qui pensent que l'amp a un effet négatif sur la pêche contrairement aux individus du groupe 5 (tableau IV-6).

Tableau n°IV-6: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de perception

	Description de l'échantillon et des groupes			Groupes				
Groupes	***Nb d'ind.***	***Variables***	***Proportion***	***1***	***2***	***3***	***4***	***5***
Échantillon	***50***	***Nombre de conflits***	0 conflit : 48%					
		Effet de l'AMP sur la pêche	Ne sait pas: 38%					
		Effet de l'AMP sur l'ecosyst.	Positif: 66%					
		Connaissance de la regl.	Non: 52%					
		Avis sur le respect de la regl.	Non: 78%					
Groupe 1: *"Pense que la réglementation n'est pas respectée et ne se prononce pas sur l'effet de l'AMP sur la pêche"*	***14***	***Nombre de conflits***	1 conflit : 57%					
		Effet de l'AMP sur la pêche	Ne sait pas: 92%					
		Effet de l'AMP sur l'ecosyst.	Positif: 78%					
		Connaissance de la regl.	Oui: 57%					
		Avis sur le respect de la regl.	Non: 100%					
Groupe 2: *"Avis d'un effet neutre de l'AMP sur la pêche"*	***13***	***Nombre de conflits***	0 conflit: 69%					
		Effet de l'AMP sur la pêche	Neutre: 100%					
		Effet de l'AMP sur l'ecosyst.	Négatif: 53%					
		Connaissance de la regl.	Non: 92%					
		Avis sur le respect de la regl.	Non: 92%					
Groupe 3: *"Pense que la réglementation est respectée et ne se prononce pas sur l'effet de l'AMP sur la pêche"*	***9***	***Nombre de conflits***	0 conflit: 69%					
		Effet de l'AMP sur la pêche	Ne sait pas: 69%					
		Effet de l'AMP sur l'ecosyst.	Positif: 56%					
		Connaissance de la regl.	oui: 89%					
		Avis sur le respect de la regl.	oui: 89%					
Groupe 4: *"Avis négatif de l'effet de l'AMP sur la pêche"*	***8***	***Nombre de conflits***	1 conflit: 63%					
		Effet de l'AMP sur la pêche	Négatif: 100%					
		Effet de l'AMP sur l'ecosyst.	Positif: 88%					
		Connaissance de la regl.	Non: 50% ; Oui: 50%					
		Avis sur le respect de la regl.	Non: 75%					
Groupe 5: *"Avis positif de l'effet de l'AMP sur la pêche"*	***6***	***Nombre de conflits***	0 conflit: 50%					
		Effet de l'AMP sur la pêche	Positif: 100%					
		Effet de l'AMP sur l'ecosyst.	Positif: 83%					
		Connaissance de la regl.	Non: 50% ; Oui: 50%					
		Avis sur le respect de la regl.	Non: 100%					

Sélection des métriques

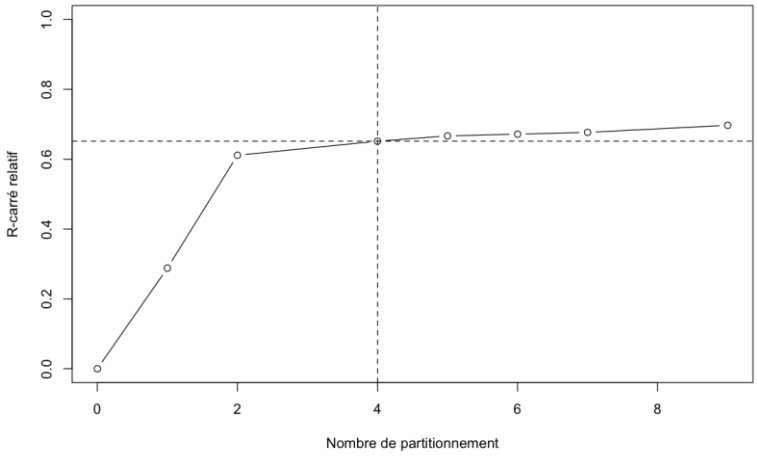


Figure n°IV-14: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de noeud terminaux

À partir de 4 partitions de l'échantillon, le R-carré relatif de l'arbre de segmentation atteint pratiquement son maximum. Il suffit donc de 6 noeuds terminaux pour construire l'arbre de segmentation pour prédire l'appartenance des individus aux groupes de pratique à partir des variables du même thème (figure IV-14).

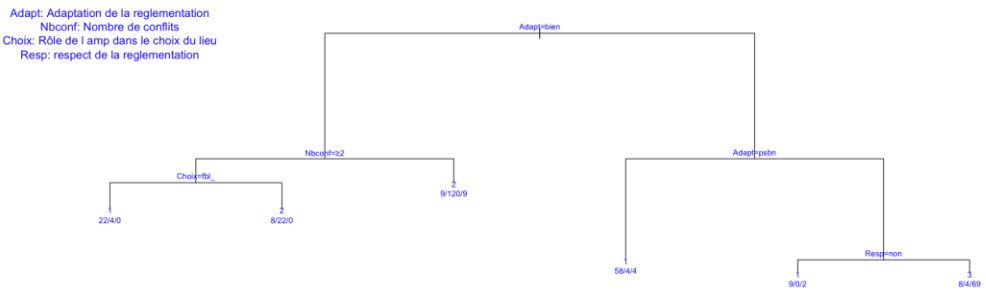


Figure n°IV-15: Arbre de reclassification pour les groupes de perception en fonction des variables du même thème

Les résultats obtenus sont cohérents avec le descriptif des groupes dans le tableau n°6. Le groupe 2 présente les individus qui pensent principalement que la réglementation est adaptée, les individus du groupe 1 la trouvent pour l'essentiel inadapté et les individus du groupe 3 ne se prononcent pas sur cette question (figure IV-15).

L'avis du pêcheur sur l'adaptation de la réglementation est la métrique prioritaire a analyser comme indicateur potentiel.

Conclusion

On a mis en évidence 5 groupes de perception parmi l'échantillon en fonction de la connaissance et de l'avis sur le respect de la réglementation, de l'avis de l'usager sur l'effet de l'AMP (sur la pêche, sur l'écosystème) et du nombre de conflits. L'avis du pêcheur sur l'effet de l'AMP sur la pêche est la variable qui reflète le mieux les groupes de perception. Elle s'avère être donc la métrique le plus pertinente à tester en tant qu'indicateur potentiel.

E. Thème Socio-économie

Tableau n°IV-7: Tableau des variables utilisé dans la typologie de leurs modalités dans le référentiel PAMPA, des modalités recodées pour les besoins de la typologie et du nombre d'individus de chacune de ces modalités.

Variables	Modalités PAMPA	Modalités typologie	Nombre d'individus
E3_CSP	Actif	actif	31
	Agriculteur		
	Cadre		
	Artisan		
	Employé		
	Intermediaire		
	liberal		
	ouvrier		
	Étudiant	Pas actif	19
	retraité		
	Pas actif		
E2_an_nais	Quantitatif	≤1950	15
		[1951-1965]	20
		>1965	15
E4_commune	Code postal des Communes de résidence	97416	16
		97427	16
		97434	17

Pour chacune des variables utilisés, une recodage des données est effectué de manière à avoir un nombre de modalité par variable compris entre 2 et 4. Ce recodage doit permettre de conserver une cohérence de l'information afin d'être pertinent en temps qu'indicateur de socio-économie.

Typologie

Pour l'ACM du thème coscio-économie, Il n'y a pas de décroissance des valeurs propres des axes de moins de 10%, la décroissance minimale étant entre les 2e et 3e axes

(14%) . Néanmoins au delà du 3e axe, les axes factoriels ne présentent pas de corrélations de plus de 50% avec une variable de départ. De ce fait on conserve les trois premiers axes pour construire l'ACM de socio-économie.

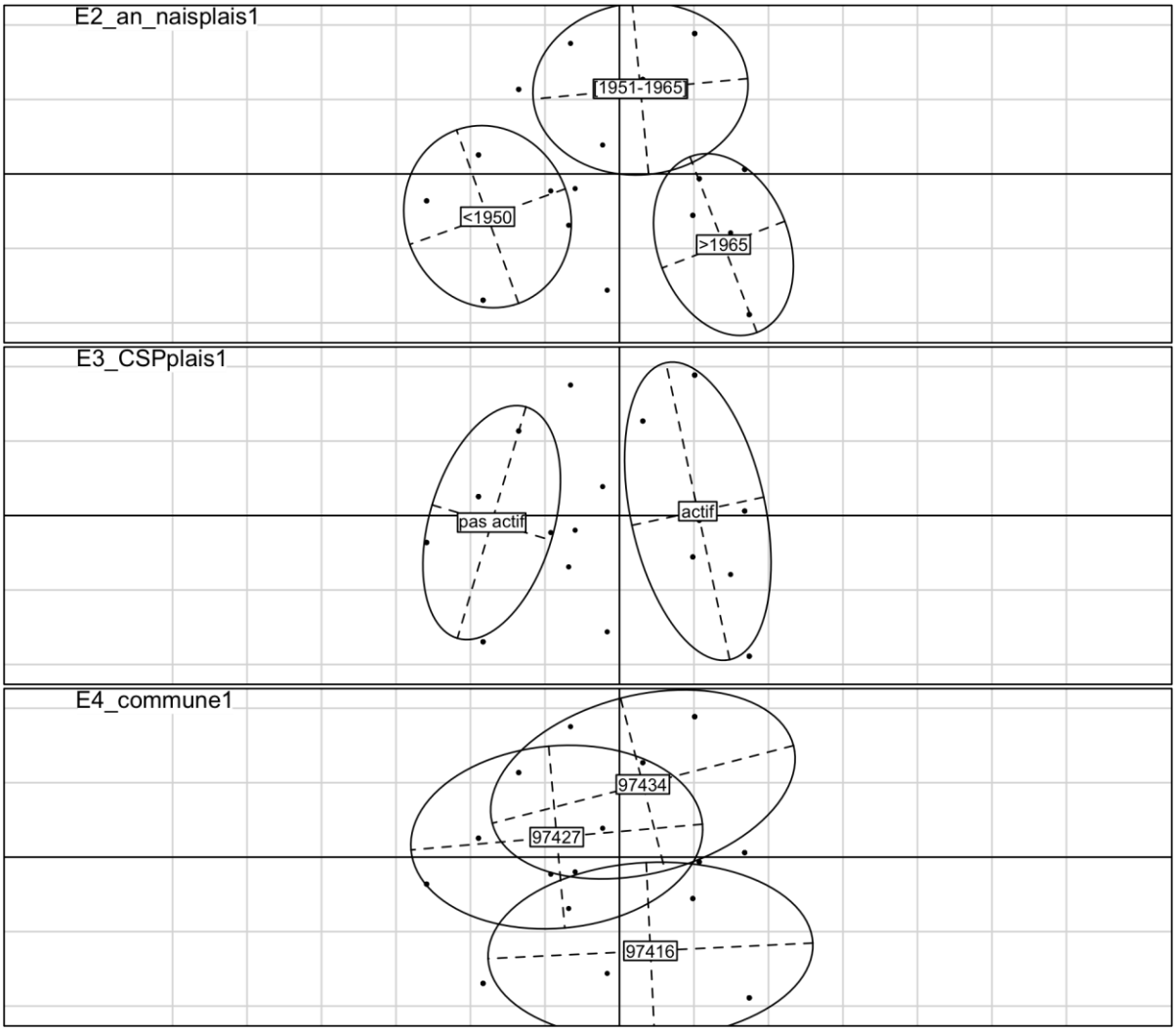


Figure n°IV-16: Projection des individus et des modalités pour chaque variable de socio-économie sur le plan factoriel de l'ACM

Sur la figure IV-16 les individus sont représentés par des points, tandis que les modalités des variables sont représentées par un cadre contenant le libellé de la modalité dont la position représente sa moyenne multivariée (centroïde) et des lignes pointillées dont la plus longue illustre la variance des coordonnées des individus sur le premier axe factoriel et la plus courte sur le second axe. L'inclinaison du cercle représente la corrélation du premier axe avec la ligne pointillée la plus longue.

Partitionnement

La valeur du critère de Calinski-Harabasz est la plus élevée pour 5 groupes (39,88). On choisit donc de former 5 groupes de pratique parmi l'échantillon de 50 individus.

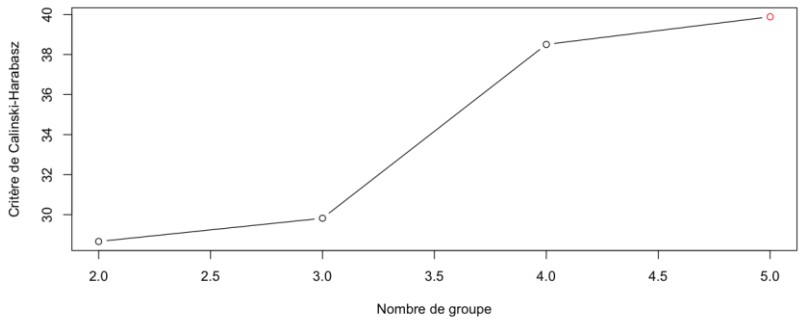


Figure n°IV-17: Valeurs du critère de Calinski-Harabasz en fonction du nombre de groupe

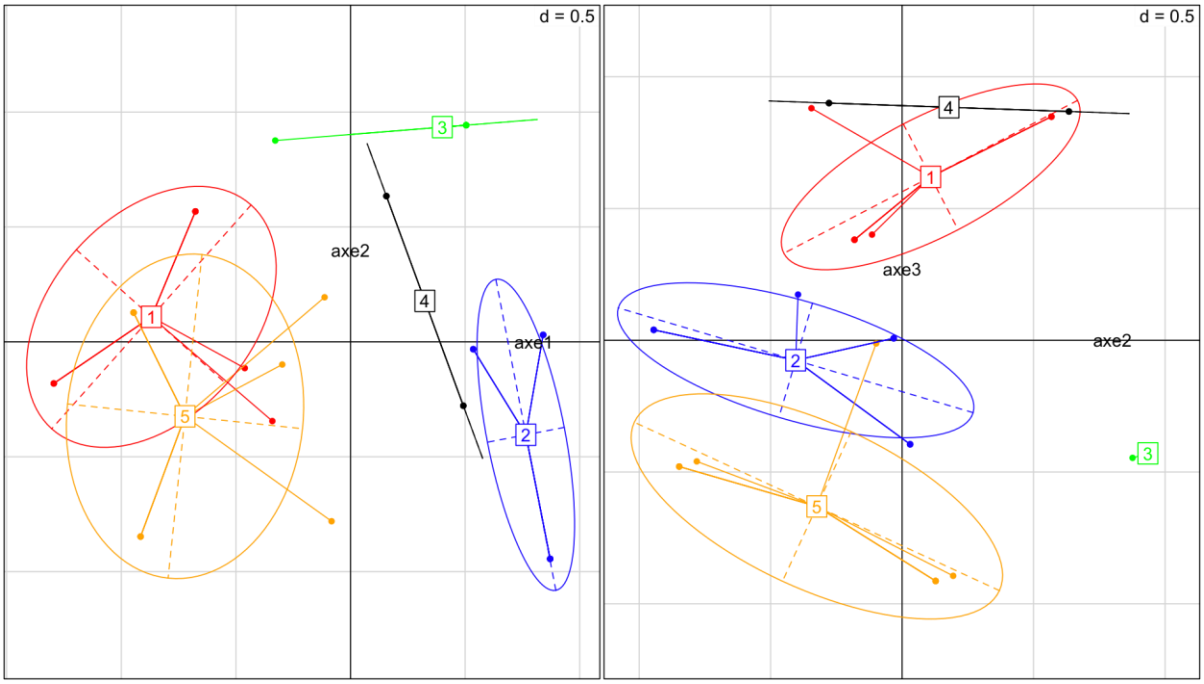


Figure n°IV-18: Projection des individus par groupe de socio-économie (points) sur les plans factoriels de l'ACM

Le groupe 1 rassemble les individus habitant dans a commune de l'Étang-Salé non-actifs(tableau IV-8 et figure 1V-18). Les individus du groupe 2 sont jeunes (≤45 ans) et actif. La totalité des individus du groupe 3 sont d'âge intermédiaire et habitent la commune de la Saline. Les individus du groupe 4 habitent la commune de l'Étang-Salé et sont actifs. Enfin les individus du groupe 5 ne sont pas actifs et ont plus de 60 ans.

Tableau n°IV-8: Tableau récapitulatif des statistiques et des résultats des tests entre les groupes de socio-économie

	Description de l'échantillon et des groupes			Groupes				
Groupes	***Nb d'ind.***	***Variables***	***Proportion***	***1***	***2***	***3***	***4***	***5***
Échantillon	***50***	***Année de naissance***	[1951-1965]: 40%					
		CSP	actif: 62%					
		Commune	97434; 97416; 97427: 33%					
Groupe 1: *"Habitant de la commune de l'Étang-Salé et non actif"*	***10***	***Année de naissance***	<1950: 50%					
		CSP	pas actif: 90%					
		Commune	97427: 100%					
Groupe 2: *"Jeune et Actif"*	***15***	***Année de naissance***	>1965: 73%					
		CSP	actif: 100%					
		Commune	97416: 67%					
Groupe 3: *"Habitant de la commune de La Saline"*	***8***	***Année de naissance***	[1951-1965]: 100%					
		CSP	actif: 88%					
		Commune	97434: 100%					
Groupe 4: *"Habitant de la commune de l'Étang-Salé et actif"*	***6***	***Année de naissance***	>1965; [1951-1965]: 50%					
		CSP	actif: 100%					
		Commune	97427: 100%					
Groupe 5: *"Âgé de plus de 60 ans"*	***11***	***Année de naissance***	<1950: 91%					
		CSP	pas actif: 82%					
		Commune	97416: 54%					

Sélection des métriques

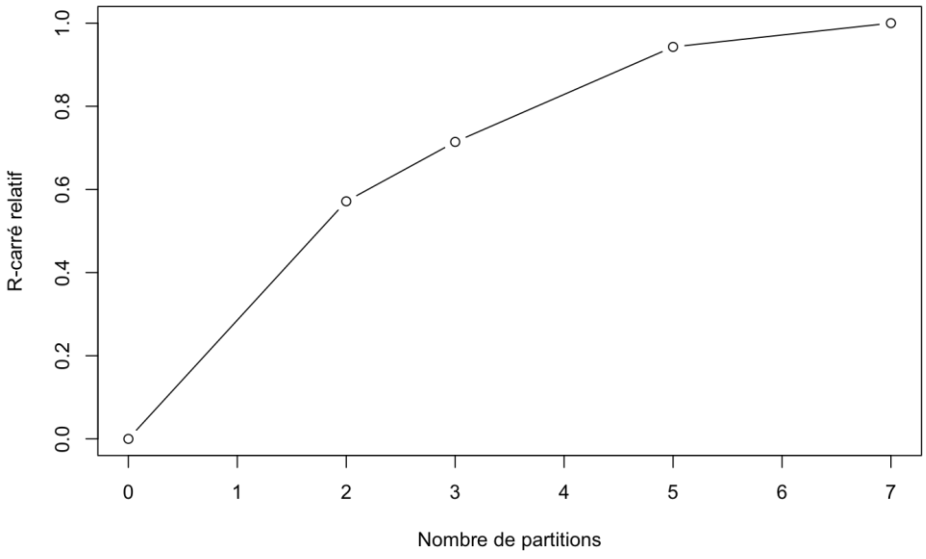


Figure n°IV-19: Courbe représentant le R-carrée relatif de l'arbre de segmentation en fonction du nombre de partitions

Le R-carré relatif de l'arbre de segmentation ne cesse d'augmenter avec le nombre de partition. Il est donc nécessaire de conserver la totalité des noeuds terminaux.

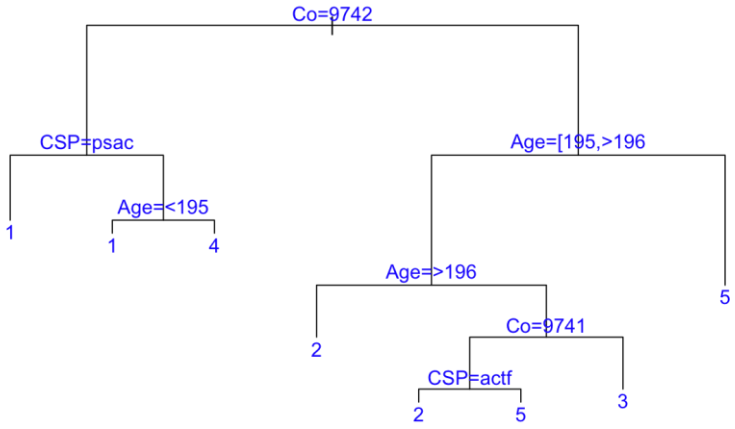


Figure n°26: Arbre de reclassification pour les groupes de socio- économie en fonction des variables du même thème

La commune d'habitation apparaît par la méthode de segmentation comme la métrique prioritaire à analyser comme indicateur potentiel. La deuxième métrique prioritaire a analyser est une combinaison des variable "année de naissance", "Commune" et "CSP" en prenant comme borne des classes les valeurs seuils indiquées par les 3 premières partitions de l'arbre.

Conclusion

On a mis en évidence 5 groupes de socio économie parmi l'échantillon en fonction de l'année de naissance, de la catégorie socio professionnelle, et de la commune d'habitation du pêcheur. La commune d'habitation du pêcheur est la variable qui reflète le mieux les groupes de socio économie. Elle s'avère être donc la métrique le plus pertinente à tester en tant qu'indicateur potentiel.

Bibliographie

Amand M. Et al. (2004) A step toward the definition of ecological indicators of the impact of fishing on the fish assemblage of the Abore reef reserve (New Caledonia). Aquatic living ressource, vol. 17 : 139-149

Calinski T. et Harabasz J. (1974) A dendrite method for cluster analysis. Communications in statistics, vol.3 : 1-27

Champely S. (2006) Tests statistiques paramétriques : Puissance, taille d'effetet taille d'échantillon (sous R), Université Lyon 1, France

Claudet J. et al. (2006) Assessing the effects of marine protected area (MPA) on a reef fish assemblage in a northwestern Mediterranean marine reserve: Identifying community-based indicators. Biological conservation, vol.130 (issues 3) : 349-369

Clua E. et al. (2005) Towards multidisciplinary indicator dashboards for coral reef fisheries management. - Aquatic living resources - vol.18 : 199-213

Cury PM. et Christensen V. (2005) Quantitative ecosystem indicators for fisheries management. ICES Journal of marine science, vol.62 : 307-310

Ferraris J. et al. (2005) Assessing the impact of removing reserve status on the Abore Reef fish assemblage in New Caledonia. Marine ecology process series, vol. 292 : 271-286

Garcia SM. et Staples DJ. (2000) Sustainability reference systems and indicators for responsible marine capture fisheries: a review of concept and elements for a set of guidelines. - Marine freshwater research, vol.51 : 385-426

Himes H. (2007) Performance indicators in MPA management: Using questionnaires to analyze stakeholder preferences. - Ocean & coastal management, vol.50 : 329-351

Link JS. et al. (2001) Ecosystem status in the Nrtheast United states continental shelf ecosystem: integration, synthesis, trends and meaning of ecosystem metrics. Or getting the brass tacks of ecosystem based fishery management. ICES document, CM 2001/T : 10. 41 pp.

Lloret J. et al. (2008) Evolution of a mediterranean coastal zone: human impacts on the environment of Cape Creus. Environmental management, vol.42 : 977:988

Milligan GW. et Cooper MC. (1985) an examination of procedures for determining the number of clusters in a data set. Psychometrika, vol.50 (number 2) 159-179

Pelletier D. Et al. (2005) Designing indicators for assessing the effects of marine protected areas on coral reef ecosystems: A multidisciplinary standpoint. Aquatic living ressource,

vol.18 : 15-33

Pomeroy RS. et al. (2005) How is your MPA doing? A methodology for evaluating the management effectiveness of marine protected areas. - Ocean & coastal management - vol.48 (issues 7-8) : 485-502.

Rice JC. et Rochet MJ. (2005) A framework for selecting a suite of indicators for fisheries management, vol.62 : 516-527

Russ GR. (1999) Management history of Sumilon and Apo marine reserves, Philippines, and their influence on national marine resource policy. Coral reefs, vol.18 : 307-319

Sainsbury KJ. et al. (2000) Design of opertational management strategies for achieving fishery ecosystem objectives. - ICES Journal of marine Science, vol.57 : 731-741.

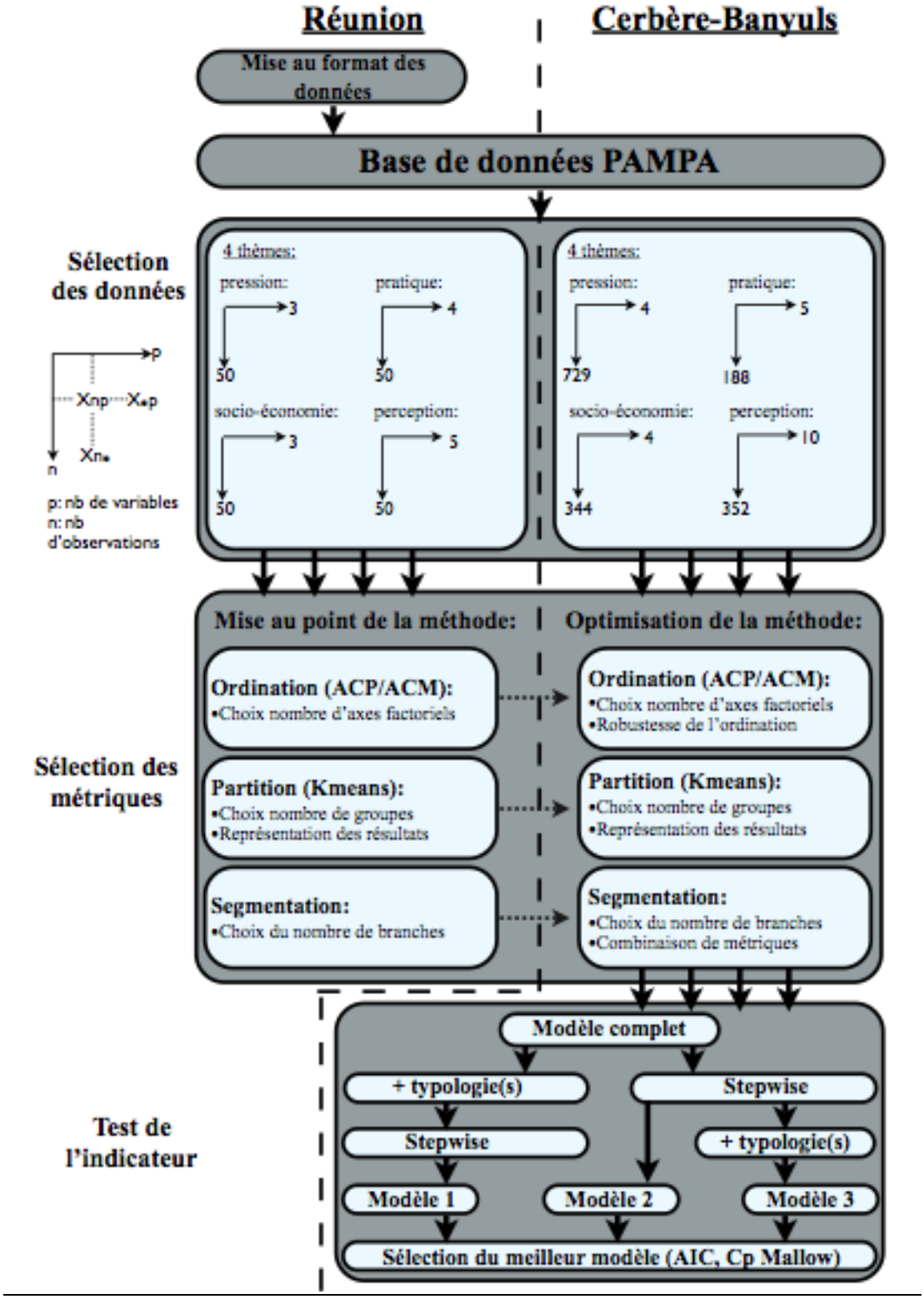
Stelzenmuller V. et al. (2008) Spatial assessment of fishing effort around European marine reserves: Implications for successful fisheries management. Marine pollution bulletin, vol.56 : 2018-2026

Suman D. et al. (1999) Perception and attitudes regarding marine reserves: a comparison of stakeholder groups in the Florida Keys National Marine Sanctuary. - Ocean & coastal management, vol.42 : 1019-1040

Thomassin A. et al., Social acceptability of a marine protected area: The case of Reunion Island, Ocean and coastal management (2010), doi:10:1016/j.ocecoaman.2010.01.008

ANNEXES

Annexe n°1: Schéma récapitulatif de la méthode utilisée pour les différents sites étudiés



P.S. : le modèle statistique utilisé dans la méthodologie est un Modèle linéaire généralisé

Annexe n°2: tableau récapitulatif des données utilisées pour la réalisation des typologies de chaque thème de chaque site (La Réunion, Banyuls et la Nouvelle Calédonie)[4]

Thèmes	Variables	Nature de la variable	RUN	BA	NC
Pression	B3_nb_pecheur	quantitatif	X	X	X
	B6_duree_sortie	quantitatif		X	X
	B6_duree_peche	quantitatif		X	
	B6_duree_peche/B6_duree_sortie	quantitatif			X
	Nb_engins=B5_nb_engin1+B5_nb_engin2+B5_nb_engin3	quantitatif	X	X	
	Nb_engins/B3_nb_pecheur	quantitatif			X
	C4_dureePec_moy	quantitatif	X		
Pratique	B6_pratique	qualitatif		X	
	C2_plannif_an	qualitatif	X	X	X
	C2_plannif_sem	qualitatif	X	X	X
	C2_plannif_jour	qualitatif	X	X	
	C4_nb_sortie_tot	qualitatif	X	X	
	Plannif_env=C2_plannif_env1+C2_plannif_env2	qualitatif			X
	C3_act_hab1	qualitatif			X
	Nb_act=C3_act_hab1+C3_act_hab2	qualitatif			X
Perception	D1_existence_amp	qualitatif		X	
	D1_amp_choix_lieu	qualitatif		X	
	D1_suffi_info	qualitatif		X	
	D1_adapt_reg	qualitatif		X	
	D1_respect_reg	qualitatif	X	X	
	D1_ass_process_dec	qualitatif		X	
	D1_connais_reg	qualitatif	X		
	D2_effet_ecosyst	qualitatif	X	X	
	D2_effet_peche	qualitatif	X	X	
	D2_effet_econo	qualitatif		X	
	Nb_conflits=Somme(D3_relation_XX="conflictuelle")	qualitatif	X	X	
Socio-économie	E2_an_nais	qualitatif	X	X	X
	E3_CSP	qualitatif	X	X	X
	E4_commune	qualitatif	X	X	X
	C8_budget_an	qualitatif		X	
Impact	C7_categ_capt	qualitatif			X
	B7_capture	qualitatif			X
	C7_capt_an/C4_nb_sortie_tot/B6_duree_peche	qualitatif			X
	Nb_esp=C7_esp_an1+C7_esp_an2+C7_esp_an3 +C7_esp_an4+C7_esp_an5	qualitatif			X

[4] *Gamp E. & Pelletier, D. Format d'entrée des données pour le calcul des métriques usages. Document interne PAMPA/WP3/Meth/5. Version du 08 août 2010. 27p.*