

**RAPPORT DU SOUS-GROUPE SUR LES STATISTIQUES**

(La Jolla, États-Unis, 14 - 18 juillet 1997)

**RAPPORT DU SOUS-GROUPE SUR LES STATISTIQUES**  
(La Jolla, États-Unis, 14 - 18 juillet 1997)

INTRODUCTION

1.1 La réunion du sous-groupe sur les statistiques, présidée par Georges Watters (États-Unis), s'est tenue du 14 au 18 juillet 1997 au Southwest Fisheries Science Center à La Jolla, aux États-Unis.

1.2 L'ordre du jour provisoire est présenté et discuté. Il est convenu d'y ajouter la question "Conception des campagnes d'évaluation synoptique". L'ordre du jour (supplément A) est adopté sans autre modification.

1.3 La liste des participants figure en supplément B et la liste des documents soumis à la réunion en supplément C.

1.4 Le rapport est rédigé par Ian Boyd et John Croxall (Royaume-Uni), Brian Manly (Nouvelle-Zélande), William de la Mare (Australie), A. Murray (Royaume-Uni), David Ramm (secrétariat) et George Watters (États-Unis).

EXAMEN DES SÉRIES CHRONOLOGIQUES MISES À JOUR DES INDICES DU CEMP

2.1 D. Ramm présente WG-EMM-97/25 dans lequel figurent des tableaux récapitulatifs de toutes les données soumises au CEMP (section 2), une sélection de chiffres illustrant ces données (section 3) et des informations liées à la découverte d'anomalies par les méthodes proposées par le sous-groupe l'année dernière (section 1).

2.2 Des remerciements sont adressés à D. Ramm et au secrétariat qui ont produit une série de documents fort complète.

2.3 En examinant la compilation des indices, le sous-groupe remarque quelques erreurs; celles-ci sont corrigées dans WG-EMM-97/25 rév.1.

2.4 Par ailleurs, le sous-groupe note que :

- i) sur l'illustration des données collectées selon la méthode A1B (section 3, A1B, figures 1 à 5), les années devraient être différenciées plus clairement; et
- ii) il existe maintenant des données qui permettraient d'évaluer si les régimes d'échantillonnage et la taille des échantillons recommandés dans plusieurs méthodes standard sont valables. Les membres qui possèdent de telles données devraient procéder aux évaluations et transmettre les résultats au WG-EMM.

## IDENTIFICATION DES ANOMALIES DANS LES INDICES DU CEMP

2.5 Le sous-groupe met en évidence deux questions liées à l'identification des anomalies :

- i) l'identification d'anomalies dans les données des distributions non-normales; et
- ii) le fait que certaines observations qui sont des "anomalies" sur le plan biologique ne sont pas forcément importantes sur le plan statistique.

2.6 Examen du document de Brian Manly et Darryl Mackenzie (WG-EMM-Stats-97/6). Les auteurs discutent des caractéristiques d'une méthode de détection des années anormales dans les indices du CEMP et appliquent le principe aux situations où les données présentent une tendance linéaire et sont autocorrélées, et où les données sont tirées d'une distribution constante autre qu'une distribution normale. Dans le cas des données d'une distribution non-normale, une transformation de Box-Cox est appliquée avant l'analyse. Cette méthode, qui demande encore à être approfondie, semble en général mieux convenir pour détecter les valeurs extrêmes que, par exemple, un changement permanent de la moyenne d'une série de données.

2.7 Examen du document de W. de la Mare (WG-EMM-Stats-97/7). L'auteur propose, en combinant les variables du CEMP, de produire un nombre moins élevé d'indices récapitulatifs. Il note également que la procédure actuelle de détection des anomalies n'est pas assez puissante lorsqu'il y a plusieurs valeurs extrêmes, et qu'un changement permanent de la moyenne et/ou de l'écart-type d'une série est plus aisément détecté en calculant les valeurs résiduelles standardisées au moyen de la moyenne et de l'écart type d'une base donnée dérivée de la série. De ce point de vue la détection des anomalies suivra les étapes suivantes :

- i) définir les classes de comportement d'une série à détecter (un changement de la moyenne, de la variance, une tendance, etc.);
- ii) sélectionner une transformation normalisante si nécessaire;
- iii) sélectionner une base dérivée de la série;
- iv) examiner les caractéristiques statistiques de la procédure en tenant compte des possibilités de corrélation en série, de valeurs manquantes, etc.; et
- v) examiner la capacité de la procédure à détecter les phénomènes dignes d'intérêt.

2.8 La discussion porte sur la nécessité de tenir compte des diverses utilisations des indices. Il est noté qu'ils sont essentiellement censés mesurer divers aspects de la nourriture disponible pour les prédateurs, avec une intégration à diverses échelles spatio-temporelles (tableau 1). Ceci souligne la nécessité de comprendre la relation entre les indices par le biais d'analyses à variables multiples, notamment s'ils sont combinés pour produire divers indices récapitulatifs.

2.9 Le terme "anomalies" peut prêter à confusion, car souvent, les valeurs extrêmes qui doivent être détectées font partie de la variation naturelle du système. Dans une certaine mesure, il est possible que ces valeurs soient le résultat de réponses largement non linéaires des prédateurs aux conditions de l'environnement. Il est recommandé d'utiliser un autre terme

tel que VOGON (valeur n'entrant pas dans les normes généralement observées). Dans ce cas, le terme "normes" s'applique aux conditions satisfaisantes pour les populations de prédateurs.

2.10 Des calculs sont effectués à titre d'exemple afin de démontrer la valeur potentielle de l'analyse à variables multiples. On se sert alors des données du tableau 2 sur l'île Bird. L'analyse en composantes principales sur la matrice de corrélation des indices des années 1990 à 1997 produit les résultats figurant au supplément D. Il apparaît que la première composante représente 53,0% de la variation des données, alors que les deuxième et troisième représentent respectivement 19,9% et 12,3%. Les deux premières composantes représentent donc ensemble 72,9% de la variation et les trois premières, 85,3%. L'application de l'analyse aux données transformées donne des résultats similaires.

2.11 La première composante est essentiellement la moyenne de la durée des sorties alimentaires des femelles d'otaries (s'il y a un signe négatif, les valeurs les plus proches de zéro représentent des conditions favorables), du succès reproductif des manchots papous, du poids des jeunes gorfous macaroni à la première mue, de la proportion de krill dans le régime alimentaire des gorfous macaroni et des manchots papous, de la moyenne du poids de la dernière pesée des jeunes femelles d'otaries et de celle du poids de la dernière pesée des jeunes mâles d'otaries. Cette composante est considérée comme l'*état biologique général*. La deuxième composante reflète principalement les taux de croissance des jeunes otaries, mâles ou femelles, qui peuvent être biaisés par une mortalité élevée les années médiocres. Pour cette raison, les valeurs élevées ne sont pas toujours associées à des conditions favorables. Cette composante peut être interprétée comme étant la *croissance des jeunes otaries*. La troisième composante concerne principalement le *succès reproductif des gorfous macaroni*. Or les années médiocres, ces manchots sont capables d'ajuster leur régime alimentaire, ce qui laisse entendre que cette composante ne représente pas non plus une mesure fiable des conditions biologiques générales.

2.12 Le sous-groupe estime que les résultats de cette analyse en composantes principales permettent de clarifier la relation entre chacun des divers indices et les conditions annuelles. Il recommande de réaliser ce type d'analyse sur d'autres sites et avec d'autres variables.

2.13 Une première exploration de l'indice d'une combinaison simple suggéré dans WG-EMM-Stats-97/7 est préparée à partir de données sur les espèces dépendantes du CEMP, à savoir celles sur les otaries, les gorfous macaronis et les manchots papous de Géorgie du Sud. Il est possible de combiner les paramètres sélectionnés pour cet exemple, car ils se réfèrent à des échelles spatio-temporelles semblables. Les paramètres en question figurent au tableau 2.

2.14 L'indice simple est calculé en transformant et standardisant les divers paramètres suivant la méthode adoptée par le WG-EMM en 1996. Chaque paramètre est transformé de manière à suivre approximativement une distribution normale standard. On ajoute ensuite les valeurs paramétriques, puis on les standardise de nouveau au moyen de l'écart-type estimé à partir de la somme effectuée en utilisant la matrice de la covariance (corrélation). Les valeurs sont également standardisées en vertu du signe, les valeurs positives par exemple indiquent des conditions supérieures à la moyenne pour les prédateurs. C'est pour cette raison que l'on a inversé le signe de la valeur transformée de la durée de la sortie alimentaire des otaries. Il est possible de calculer l'indice simple de toutes les années pour lesquelles il existe des données.

2.15 Les valeurs moyennes et la matrice de la covariance nécessaires pour la standardisation des séries de données sont calculées à partir des données de la période de 1989 à 1997, années pour lesquelles on dispose des données sur tous les paramètres. Avant d'être standardisées, les données suivent les transformations acceptées à l'heure actuelle pour chaque paramètre. Cette période est utilisée pour fournir une moyenne et une matrice de la covariance de base pour calculer l'indice depuis le début des séries de données en 1977. Le sous-groupe n'a pas cherché à déterminer si cette période était une base adéquate; les résultats présentés ici ne le sont qu'à titre d'exemple. La matrice de la corrélation obtenue est donnée au tableau 3.

2.16 La figure 1 illustre l'indice simple utilisant toutes les données disponibles. Les deux années médiocres connues, 1977 et 1984, y apparaissent clairement. L'indice suggère également que les années 1987, 1988 et 1994 sont médiocres, mais pas tant pour cette dernière que ne le laisse entendre l'évaluation effectuée par le WG-CEMP cette année-là. En raison du fait que les paramètres de croissance des jeunes otaries ne comptaient pas pour beaucoup dans la première composante principale de l'analyse en composantes principales (paragraphe 2.11), on a recalculé l'indice sans utiliser ces données. L'exclusion de ces données de l'indice (ligne en tirets) entraîne tout au plus un abaissement plus prononcé du point de 1994, mais est sans grande importance. Étant donné que 1994 était une année extrêmement mauvaise pour les otaries, le fait que l'indice soit insensible à la croissance des jeunes otaries laisse entendre que ce paramètre ne donne pas un indice représentatif du succès reproductif des otaries. Il est suggéré d'envisager un ajustement de ces paramètres, en remplaçant les taux de croissance de chaque jeune par le taux de croissance de la biomasse totale des jeunes.

2.17 La figure 2 est une comparaison de l'indice simple calculé sans les taux de croissance des jeunes otaries (ligne en tirets) et de celui fondé uniquement sur le succès reproductif des deux espèces de manchots (seuls paramètres représentés toutes les années). Elle indique que, dans ce cas notamment, l'indice n'est pas particulièrement sensible à l'absence de certains des paramètres.

2.18 Le sous-groupe estime que les résultats sont encourageants et recommande de poursuivre l'étude en vue de mettre au point un certain type d'indices simples combinés à l'échelle régionale et temporelle appropriée. Le sous-groupe note également que l'indice simple peut s'avérer plus robuste que les indices paramétriques séparés lorsqu'il s'agit d'identifier les VOGON, car la distribution d'une somme de variables aléatoires se rapproche d'une distribution normale, même lorsque les variables aléatoires elles-mêmes ne sont pas de distribution normale.

2.19 Le sous-groupe note que par le passé, on s'est inquiété du fait que la méthode de détection des VOGON ne les identifiait pas toujours, alors que leur importance biologique était reconnue (SC-CAMLR-XV, annexe 4, paragraphe 4.72). Il convient que lorsque la distribution d'un indice (ou de sa transformation) n'est pas proche de la normale, le niveau  $\alpha$  de 0,05 risque de ne pas être assez souple pour détecter les VOGON qui sont importants sur le plan biologique. Il est également suggéré qu'il pourrait s'avérer utile de mettre au point une procédure d'identification d'un VOGON lorsqu'une grande proportion des indices sont proches de leur niveau critique dans une même année, sans toutefois le dépasser.

2.20 En vue de fournir deux exemples de la rigidité du niveau  $\alpha$  de 0,05, le sous-groupe a estimé quel niveau  $\alpha$  serait nécessaire pour détecter tous les VOGON importants sur le plan

biologique dans la série chronologique du succès reproductif des manchots papous (indice A6a) et des albatros à sourcils noirs (indice B1) de l'île Bird. J. Croxall a identifié dans chaque série les VOGON importants sur le plan biologique.

2.21 Le calcul pour chaque exemple est effectué en quatre étapes :

- transformation de l'indice par une probabilité logarithmique;
- identification du VOGON qui, important sur le plan biologique, est le moins extrême;
- calcul de la valeur critique ( $Z_c$ ) pour détecter le VOGON le moins extrême à partir de :

$$Z_c = \frac{\bar{x} - LEV}{s}$$

où  $\bar{x}$  et  $s$  sont la moyenne et l'écart-type de l'indice transformé, et  $LEV$  la valeur du VOGON le moins extrême; et

- identification du niveau  $\alpha$  correspondant à  $Z_c$  par la simulation de 1 000 séries chronologiques sur 20 ans d'écart normal standardisés; on compte le nombre de cas où la valeur absolue de l'écart simulé est  $\geq Z_c$ , et on divise ce chiffre par 20 000.

2.22 Les résultats des exemples de calculs apparaissent au tableau 4. Pour détecter tous les VOGON importants sur le plan biologique dans la série chronologique des manchots papous,  $\alpha$  doit être égal à 0,22 et dans celle des albatros à 0,69. Dans les deux cas,  $\alpha = 0,05$  serait trop rigoureux.

2.23 Vu les résultats de ces calculs, le sous-groupe convient que la sélection du niveau  $\alpha$  pour identifier les VOGON doit se faire en fonction de l'indice, après avoir déterminé avec précaution si l'indice (ou sa transformation) est distribué normalement. S'il n'est pas normal, le niveau  $\alpha$  devrait probablement être situé entre 0,2 et 0,3.

#### ÉVALUATION CRITIQUE DES HYPOTHÈSES ET DES VALEURS PARAMÉTRIQUES DU MODÈLE D'AGNEW ET PHEGAN (1995) DU CHEVAUCHEMENT RÉALISÉ

3.1 L'année dernière, le WG-EMM a chargé le sous-groupe sur les statistiques d'évaluer les hypothèses et valeurs paramétriques du modèle à échelle précise du chevauchement de la demande alimentaire des manchots et de la pêcherie de krill des îles Shetland du Sud et de la péninsule Antarctique (Agnew et Phegan, 1995) (SC-CAMLR-XV, annexe 4, paragraphe 6.80). Ce modèle calcule la demande alimentaire des manchots dans le dessein de calculer un indice de chevauchement de l'approvisionnement et de la pêcherie pendant la période critique de décembre à mars. Les données de la sous-zone 48.1 sur les caractéristiques de l'approvisionnement des manchots, leur demande énergétique et l'effectif de leur population, ainsi que sur les captures mensuelles de krill par case à échelle précise servent de données d'entrée du modèle.

3.2 En vue de faciliter ce processus, le secrétariat avait demandé (SC CIRC 97/2) des données et analyses qui fourniraient des estimations :

- i) de la composition mensuelle du régime alimentaire (des manchots et des otaries);
- ii) de la distance maximale et moyenne/modale de l'approvisionnement;
- iii) de la direction moyenne de l'approvisionnement ; et
- iv) des données à échelle précise sur la distribution de l'approvisionnement.

3.3 Le document WG-EMM-Stats-97/5 contient de telles données sur les manchots papous, les gorfous macaroni et les otaries de Kerguelen de l'île Bird, en Géorgie du Sud (sous-zone 48.3). Des données sur les manchots à jugulaire de l'île Seal ont été soumises au secrétariat à l'intention du WG-EMM, mais elles n'étaient pas disponibles à la réunion du sous-groupe. Il est décevant que des données de ce type n'aient pas été fournies pour d'autres sites, notamment ceux de la sous-zone 48.1 où ont été effectuées plusieurs études approfondies sur le régime alimentaire et l'approvisionnement.

3.4 L'examen du modèle amène à considérer les principaux points suivants :

- i) la distance de l'approvisionnement;
- ii) la direction de l'approvisionnement
- iii) les taux de consommation des prédateurs;
- iv) les recensements de population; et
- v) la structure du modèle.

3.5 Le modèle présume que les manchots s'approvisionnent à des distances calculées selon une distribution normale à partir d'une distance moyenne des colonies. Les valeurs utilisées dans le modèle sont : distance moyenne d'approvisionnement du manchot à jugulaire : 20 km, avec un écart-type de 8 km  $\sim N(20,8)$ ; du manchot Adélie,  $\sim N(38,15)$ ; du manchot papou,  $\sim N(10,4)$ ; et du gorfou macaroni,  $\sim N(28,11)$ . La distance moyenne d'approvisionnement est fixée à des écarts-types correspondant à la moyenne + 2.

3.6 Le modèle présume que les manchots s'approvisionnent dans une direction calculée selon une distribution uniforme proche d'une ligne perpendiculaire à la côte sur laquelle est située la colonie. Les données sur la direction de l'approvisionnement des colonies de la sous-zone 48.1 ne concernent que l'île Seal. Les valeurs utilisées dans le modèle s'étendent en général jusqu'à 40° de part et d'autre de la ligne perpendiculaire à la côte.

3.7 Les données utilisées dans le modèle relativement à la distance d'approvisionnement et à la direction sont certainement appropriées pour le secteur de l'île Seal. Le groupe fait remarquer que les données disponibles sont trop rares pour permettre d'élargir le modèle de manière à ce qu'il comprenne d'autres régions de la sous-zone 48.1. Il recommande de ne procéder qu'avec précaution à l'extrapolation aux régions sur lesquelles on ne dispose pas de données.

3.8 Il est peu probable que la distribution des distances d'approvisionnement soit normale. A priori, on peut s'attendre à une certaine distribution exponentielle; d'après les observations en mer, la distribution serait asymétrique. En ce qui concerne la direction de l'approvisionnement, il n'y a apparemment aucune raison, ou preuve liée à l'observation, de

supposer que la distribution n'est pas uniforme. La distribution des deux paramètres devra être revue à la lumière de nouvelles données et de la littérature sur le déplacement des animaux.

3.9 Le modèle utilise les valeurs moyennes du taux de consommation des prédateurs qui étaient les meilleures estimations disponibles jusqu'en 1984. Les estimations utilisées dans le modèle pourraient être améliorées vu les nombreuses données complètes dont on dispose de nos jours sur le taux métabolique et les besoins énergétiques en mer des manchots (voir par ex., WG-EMM-96/19 et SC-CAMLR-XV, annexe 4, paragraphe 6.41).

3.10 Les recensements des populations de manchots utilisés dans le modèle, qui sont dérivés d'un jeu de données à long terme sur les dénombrements de manchots, étaient les meilleurs qui soient disponibles en 1992. Un jeu de données mis à jour est maintenant disponible (SC-CAMLR-XV/BG/29).

3.11 Le sous-groupe examine les quatre étapes suivies par le modèle :

- i) estimation du nombre total de manchots de toutes les colonies qui s'approvisionnent dans le secteur;
- ii) calcul du nombre de manchots supposés s'approvisionner dans une aire de 10 x 10 milles;
- iii) calcul de la consommation totale de krill des manchots; et
- iv) calcul de l'indice de chevauchement approvisionnement-pêche (FFO).

Le sous-groupe convient de l'à-propos de l'approche spatiale de base de la modélisation. Cependant, il lui semble d'une part, que les aspects temporels de l'approvisionnement des manchots ne sont pas capturés convenablement dans le modèle et qu'il conviendrait d'en poursuivre l'étude et d'autre part, que l'indice FFO n'est pas une mesure directe du chevauchement, mais concerne plutôt la quantité totale de krill prélevée du secteur d'alimentation durant la période critique. L'indice FFO est le produit : [consommation totale de krill des manchots] x [capture totale de krill par la pêche] avec des unités de (masse)<sup>2</sup>.

3.12 Le sous-groupe propose de développer un nouvel indice standardisé, fondé sur la théorie du chevauchement des niches (SC-CAMLR-XV, annexe 4, appendice H), tel l'indice de Schroeder

$$I_t = 1 - 0.5 \sum |p_{i,t} - q_{i,t}|$$

où  $p_{i,t}$  est la proportion de krill consommé par un ou plusieurs prédateurs dans la case  $i$  du quadrillage pendant la période  $t$ , et  $q_{i,t}$  est la proportion de krill prélevée par la pêche dans la même case et durant la même période  $t$ . Ce type d'indice varierait de  $I_t = 0$ , lorsqu'il n'y a aucun chevauchement spatial entre la consommation des prédateurs et les prélèvements de la pêche durant la période  $t$ , à  $I_t = 1$  lorsque la consommation des prédateurs et les prélèvements de la pêche se chevauchent complètement durant la période  $t$ . À l'heure actuelle, il est possible de calculer  $p_{i,t}$  d'après la structure proposée dans Agnew et Phegan (1995).

3.13 Il est recommandé d'appliquer ce nouvel indice en premier lieu à la sous-zone 48.1, en utilisant les données existantes sur l'île Seal. Le secrétariat devrait procéder à cet exercice dont les résultats seraient présentés à la réunion du Comité scientifique en octobre prochain.

3.14 Le sous-groupe recommande d'inclure dans les nouvelles tâches liées à l'étude du chevauchement réalisé :

- i) l'examen de la sensibilité de l'indice *I* aux diverses hypothèses relatives à l'effort d'approvisionnement et à la consommation de proies des manchots;
- ii) la saisie des données pertinentes sur l'effort et la répartition de l'approvisionnement à partir des sites de la sous-zone 48.1 en plus de l'île Seal. Ces données doivent être soumises au plus tôt sur les formulaires du secrétariat (SC CIRC 97/2), et le cas échéant, en fournissant les données et analyses de manière analogue à celles présentées dans WG-EMM-Stats-97/5; et
- iii) l'application du modèle à la sous-zone 48.3. Il est noté que, la pêcherie s'y déroulant en hiver à l'heure actuelle, l'interaction avec les prédateurs dépendants de krill pendant la période critique de décembre à mars est minime. Il pourrait toutefois être utile d'effectuer des analyses portant sur les données d'années où la pêche de krill se déroulait en été.

3.15 Il conviendrait à l'avenir d'examiner le chevauchement de la demande alimentaire des manchots et de la pêcherie de krill pendant d'autres périodes qui pourraient être aussi critiques, comme la période faisant suite à la mue, par exemple, lorsque de nombreux jeunes commencent à s'approvisionner indépendamment et que les adultes se nourrissent davantage en vue de la mue annuelle. Il semblerait, d'après de récentes études, que l'hiver peut également connaître des périodes critiques. Il existe peu de données empiriques, sinon aucune, sur la plupart de ces périodes. Les études à réaliser durant l'hiver simultanément sur la répartition de l'approvisionnement des prédateurs et la pêcherie de krill devront principalement porter sur l'otarie, le gorfou macaroni et le manchot à jugulaire.

#### DÉVELOPPEMENT D'INDICES DE COMPORTEMENT EN MER ET MÉTHODES DE DÉRIVATION PAR L'ANALYSE D'ÉCHANTILLONS DE JEUX DE DONNÉES

4.1 Par le passé, le WG-EMM avait convenu qu'il était nécessaire de coordonner l'approche de l'analyse des données sur le comportement en mer des prédateurs plongeurs tels que les manchots et les otaries. En effet, ceci devait principalement permettre de contrôler le comportement des prédateurs plongeurs à des échelles spatio-temporelles plus précises que ne le permettaient les indices du CEMP. Cette approche pourrait par ailleurs fournir des données d'entrée dans l'indice de chevauchement réalisé (paragraphe 3.12). Plusieurs jeux de données existants seraient également utilisés. Des méthodes de mesure du comportement en mer et de déploiement des instruments permettant d'effectuer ces mesures ont déjà été adoptées (WG-EMM-96).

4.2 Le sous-groupe avait été chargé :

- i) de revoir les échelles spatio-temporelles du développement des indices de comportement en mer (SC-CAMLR-XV, annexe 4, paragraphes 3.61 à 3.65 et 7.58);
- ii) d'examiner les échantillons de jeux de données et d'analyses (SC-CAMLR-XV, annexe 4, paragraphes 4.44 et 7.58);
- iii) de développer des indices et des méthodes de calcul à partir de l'analyse d'échantillons de jeux de données (SC-CAMLR-XV, paragraphe 5.38 i));
- iv) de fournir des avis sur les indices qu'il conviendrait le mieux d'inclure dans la base de données du CEMP (SC-CAMLR-XV, annexe 4, paragraphes 4.44 et 7.58).

4.3 Le sous-groupe examine plusieurs échantillons de jeux de données sur les otaries de Kerguelen. D'un jeu de données à deux variables tenant compte du temps et de la profondeur (relevés à des intervalles de 5 à 15 secondes), il est possible de dériver plusieurs paramètres supplémentaires tels que la profondeur et la durée de la plongée, ainsi que le temps passé en surface entre les plongées. Ces paramètres donnent alors des informations sur la fréquence des plongées, la proportion des plongées effectuées aux différents moments de la journée et les séquences de plongées. Les études ont montré par le passé que ces derniers paramètres pouvaient fournir des informations sur la variabilité du comportement en mer d'une année à une autre, et refléter ainsi la variation de la quantité de nourriture disponible.

4.4 La littérature n'est pas unanime sur la manière de comparer le comportement en mer des divers individus et d'une année à une autre. En principe, le sous-groupe recommande de baser les comparaisons sur les procédures qui tiennent correctement compte de la variabilité des données. L'attention est notamment attirée sur l'utilité potentielle de l'analyse spectrale. Elle a pour avantage d'incorporer toutes les données dans une seule méthode analytique tout en réduisant le nombre d'hypothèses sur la définition des unités de comportement des individus, telles que les plongées ou séquences de plongées.

4.5 Une seconde approche qui met également fin à bien des hypothèses sur la définition des plongées et des séquences de plongées, consiste à examiner le temps cumulé passé sous l'eau au cours d'une sortie alimentaire par rapport au temps cumulé passé en mer. La courbe de ce rapport fournirait un paramètre unique qui refléterait presque entièrement la variabilité du comportement en mer en un indice unique.

4.6 La comparaison du comportement en mer d'une année à une autre est compliquée par un degré potentiellement élevé de variabilité entre les individus et parce que de nombreux paramètres qui sont fréquemment utilisés pour mesurer le comportement en mer ont des distributions largement biaisées. Quelques-uns dénotent peut-être également un certain degré de bimodalité.

4.7 Le sous-groupe recommande d'examiner la possibilité d'étudier la variabilité interannuelle des indices au moyen d'une épreuve du caractère aléatoire. B. Manly suggère que cet examen pourrait se faire par la procédure suivante :

- i) présumer que les données consistent en relevés de sorties alimentaires individuelles provenant de divers animaux;

- ii) grouper les sorties alimentaires deux par deux et mesurer la différence de l'une à l'autre (par exemple, par une mesure Kolmogorov-Smirnov de la différence entre les distributions de l'indice). On obtient alors une matrice des différences pour les prédateurs, dans laquelle l'élément  $a(i, j)$ , (ligne  $i$ , colonne  $j$ , est la différence pour les prédateurs  $i$  et  $j$ ;
- iii) créer une deuxième matrice dans laquelle les éléments sont les similarités entre les échantillons comme cela est souvent recommandé pour la procédure de permutation à réponses multiples (Mielke *et al.*, 1976). Ainsi, l'élément  $b(i, j)$  (ligne  $i$ , colonne  $j$ ) contient 0 pour deux cas en des années différentes et  $1/(n-1)$  dans deux cas en une année avec un échantillon de taille  $n$ ;
- iv) vérifier si la corrélation entre  $a(i, j)$  et  $b(i, j)$  est grandement négative en la comparant à la distribution obtenue en permutant au hasard les marques des échantillons pour l'une des matrices, c'est-à-dire qu'il faut faire un test de permutation de la matrice de Mantel (1967) comme cela est décrit dans Manly (1997); et
- v) le test peut être effectué avec n'importe quelle statistique mesurant la différence entre le comportement de deux prédateurs.

4.8 En raison de la taille des jeux de données et de la nécessité d'examiner soigneusement comment ces techniques analytiques peuvent être appliquées aux mesures du comportement en mer, il serait peu pratique d'examiner ces méthodes pendant la réunion. I. Boyd et A. Murray conviennent de procéder à une analyse qui servirait d'exemple d'évaluation de cette méthode en utilisant les données de plusieurs années sur les otaries de Kerguelen et de faire un compte rendu des résultats à une prochaine réunion du WG-EMM.

4.9 C'est au moyen d'une analyse spectrale que l'on peut déterminer avec le plus de succès les échelles de variabilité du comportement en mer. I. Boyd, par exemple, montre, au moyen d'une telle analyse, que plusieurs pics du spectre correspondent aux différentes échelles de comportement, à savoir, la plongée, les séquences de plongée et la variabilité circadienne. A. Murray suggère qu'en remplaçant les hypothèses de formes d'ondes de sinus associées aux transformations de Fourier, on peut obtenir un autre spectre qui fournirait de nouvelles informations. I. Boyd et A. Murray conviennent également d'étudier cette question pendant la période d'intersession.

4.10 Le sous-groupe examine également l'intérêt d'inclure les données de position fournies par les marques satellite comme variables décrivant le comportement en mer. La précision des données de position est suffisante pour que celles-ci soient entrées dans l'indice de chevauchement réalisé prédateurs-pêcheries (paragraphe 3.12). Toutefois, à ce stade, la précision des positions fournies par satellite ne permet pas d'évaluer la variabilité de l'emplacement des secteurs d'alimentation aux échelles spatiales les plus précises des données de temps-profondeur.

4.11 En conclusion, le sous-groupe estime qu'il est prématuré de faire des recommandations précises sur les indices de comportement en mer qui devraient être entrés dans la banque de données du CEMP. Une fois que les diverses méthodes discutées par le sous-groupe auront été testées, cette question devra de nouveau être examinée.

## MÉTHODES PROPRES À FAIRE FACE AUX DONNÉES MANQUANTES DANS LES JEUX DE DONNÉES MULTIPLES

5.1 A. Murray présente son document WG-EMM-Stats-97/8. Ce document fait état de trois stades d'analyse des jeux de données incomplets :

- i) comprendre les mécanismes à la source des valeurs manquantes (sont-ils l'effet du hasard?);
- ii) décider de l'analyse des données qui conviendrait pour conforter les déductions requises (l'estimation des tendances, ou l'identification des valeurs anormales, etc.)
- iii) choisir et appliquer une méthode appropriée pour l'interpolation des données manquantes et procéder ensuite à l'analyse des données.

Les classes des mécanismes des valeurs manquantes et les catégories générales des méthodes d'interpolation sont examinées. Pour qu'une valeur soit considérée comme "manquante au hasard", la probabilité qu'elle soit manquante doit être indépendante des valeurs observées et manquantes. L'analyse d'un exemple de jeu de données des recensements d'une colonie de manchots à jugulaire de l'île Signy est présentée pour illustrer quatre méthodes d'interpolation.

5.2 L'une des méthodes d'évaluation de l'effet de l'interpolation des valeurs manquantes sur l'analyse consisterait à prendre un jeu de données complet et à essayer divers schémas (aléatoires ou non) et divers degrés de suppression de données. Les valeurs imputées pourraient ensuite être comparées aux valeurs originales et les analyses des jeux de données complétés pourraient être comparées à l'analyse du jeu de données entier. Il serait ainsi possible d'évaluer le succès des procédures d'interpolation. De nombreuses études de ce type sont rapportées dans la littérature et dans certains cas au moins, bien que les valeurs individuelles ne correspondent pas vraiment aux données originales, les statistiques telles que les moyennes peuvent être proches des valeurs originales. À titre d'information, un exercice de ce type pourrait être utile pour un exemple de jeu de données du CEMP.

5.3 WG-EMM-Stats-97/8 souligne combien il est important de comprendre les mécanismes qui mènent à l'absence de données et invite à une discussion de cette question dans le contexte des séries du CEMP. Les données des indices du CEMP peuvent manquer pour plusieurs raisons.

- i) Les données n'ont pas été collectées soit parce que cela n'était pas prévu, soit pour des problèmes logistiques tels que des moyens d'accès limités ou une panne d'équipement. Il peut être considéré que l'absence de ces données est tout à fait aléatoire.
- ii) Les données n'ont pas été collectées en raison des conditions adverses de l'environnement. Les glaces de mer pourraient avoir empêché l'accès à un site ou une journée de travail sur le terrain pourrait avoir dû être annulée par cause de mauvais temps. Selon la nature de la variable en question, de telles raisons ne doivent peut-être pas être considérées comme aléatoires. Par exemple, pour

certaines paramètres biologiques tels que la date d'arrivée, la présence des glaces de mer peut avoir une grande influence et la raison pour laquelle les données manquent pourrait également affecter la valeur. L'absence de ces données ne peut être considérée comme aléatoire.

- iii) Les données n'ont pas été collectées en raison de circonstances biologiques. Les animaux en question sont peut-être morts pendant la saison, comme lorsque, certaines années, les poussins meurent avant l'émancipation. Il est peu probable que ceci se produise au hasard et cela constitue, en soi, un indicateur biologique important du statut de l'écosystème de cette année-là.
- iv) Les données ne sont pas relevées bien que l'on sache qu'elles dépassent un seuil donné, dans le cas par exemple de données qui dépassent la capacité d'enregistrement de l'instrument. Il s'agit ici de censure qui est une procédure courante à laquelle on a recours dans les cas d'observation de la durée d'un événement, comme le retour d'une sortie alimentaire, qui ne se ferait pas avant la fin de la période d'observation disponible. Les raisons peuvent en être soit biologiques dans le cas de sorties alimentaires prolongées ou incomplètes, lors de saisons médiocres, soit non biologiques dans le cas de panne d'équipement ou lorsque la capacité de stockage de données est dépassée. Alors que le premier cas ne peut être considéré comme aléatoire, le second pourrait l'être parfois. Il existe des méthodes statistiques standard pour estimer les paramètres de distribution (tels que les moyennes) lorsque les observations de certains éléments de l'échantillon sont censurées. Il est estimé que la méthode standard d'évaluation de la durée des sorties alimentaires des otaries (méthode C1) mérite d'être revue pour déterminer si l'adoption de cette méthode d'analyse permettrait de produire des jeux de données plus complets de cet indice.
- v) Les données n'ont pas été déclarées alors que les valeurs étaient en fait nulles, par exemple, lorsque certaines proies ne faisaient pas partie des contenus stomacaux. Ces valeurs devraient être retrouvées et remplacées par zéro dans la banque de données.

5.4 Le sous-groupe reconnaît qu'il est important d'évaluer les séries du CEMP afin de déterminer les raisons pour lesquelles les données manquent avant de procéder à une analyse formelle. Cette évaluation devrait être effectuée au plus tôt. Les fournisseurs des données devraient être encouragés à fournir les informations nécessaires, et à cet effet, il semble approprié de leur soumettre un questionnaire à choix multiples correspondant aux options présentées au paragraphe 5.3.

5.5 Les données peuvent manquer à deux niveaux dans les séries du CEMP : au niveau des échantillons qui permettent de calculer la valeur soumise, et au niveau du calcul des indices du CEMP.

5.6 Il est important de découvrir dans les échantillons de données, si des techniques ont été appliquées pour faire face aux données manquantes dans les valeurs qui ont été soumises à la CCAMLR. Dans certains cas, par exemple, il manque le recensement d'une colonie dans une série de recensements à un site, et il serait possible d'interpoler la valeur d'un site. Le sous-groupe recommande, lorsque de tels cas sont reconnus, de soumettre les données brutes pour que des techniques statistiques appropriées puissent être examinées et appliquées.

5.7 Dans les séries chronologiques entrées dans la banque de données du CEMP, les valeurs manquantes ne devraient être interpolées au cours d'une analyse qu'à des fins spécifiques. Les méthodes utilisées devraient tenir compte des raisons, procurées par les fournisseurs des données, pour lesquelles ces données manquent, et du but de l'analyse. Les données ainsi interpolées ne devraient pas être saisies dans la banque de données de la CCAMLR, ni être utilisées comme des données réelles. Elles ne servent en fait qu'à permettre d'analyser des données existantes et il est vrai que des valeurs peuvent être interpolées différemment dans le contexte d'analyses diverses. Il est important de veiller à ce que les méthodes d'interpolation utilisées permettent d'utiliser toutes les données provenant des observations sans leur ajouter d'effets artificiels. C'est-à-dire que les données interpolées devraient être aussi "neutres" que possible quant à leur effet sur l'estimation de moyennes, corrélations, tendances, etc.

5.8 L'interpolation devrait être aussi réaliste que possible et il convient de tenir compte des facteurs biologiques et spatio-temporels appropriés pour décider quelles données utiliser dans les techniques d'interpolation à plusieurs variables. Par exemple, l'interpolation peut porter sur un échantillon représentatif fondé sur l'utilisation de valeurs pour la même variable ou des variables voisines à des colonies ou sites différents en une même année, ou bien être "longitudinale" et utiliser des valeurs d'années consécutives, ou encore combiner ces deux techniques.

#### CONCEPTION D'UNE CAMPAGNE D'ÉVALUATION SYNOPTIQUE

6.1 Le sous-groupe rappelle que l'objectif premier de la campagne d'évaluation synoptique est de fournir une estimation de la biomasse du krill et de sa variabilité aux fins du modèle de rendement du krill. Les autres objectifs (étude de la structure spatiale des concentrations de krill, par ex.) sont secondaires. Le sous-groupe note que la conception de cette campagne repose sur deux points clés : la stratification et le placement au hasard ou systématique des transects.

6.2 Le sous-groupe partage l'opinion qui avait été exprimée par le WG-EMM (SC-CAMLR-XV, annexe 4, paragraphe 3.75 v)) selon laquelle la campagne devrait être stratifiée en fonction des différences spatiales à grande échelle de la densité du krill. Il note qu'il existe de nombreux jeux de données anciens (FIBEX, AMLR, LTER, par ex.) qui permettraient d'estimer comment répartir l'effort d'échantillonnage entre les strates.

6.3 Pour entamer la discussion sur le placement des transects, le sous-groupe fait remarquer qu'une disposition aléatoire faciliterait les estimations de variance de la biomasse du krill fondées tant sur la conception (par les estimateurs de Jolly et Hampton, par ex.) que sur un modèle (géostatistique, par ex.). La disposition systématique des radiales repose sur une estimation de la variance effectuée à partir du modèle. Les paramètres d'estimation de la variance fondés sur un modèle peuvent être plus efficaces que ceux qui reposent sur la conception, mais ils sont fonction de l'efficacité du modèle. Il est nécessaire de procéder à une étude par simulation pour comparer l'efficacité relative des placements aléatoire et systématique des radiales dans une campagne d'évaluation synoptique du krill. Ce n'est que par une telle étude que l'on peut comparer les deux types de campagnes de manière quantitative.

6.4 Le sous-groupe reconnaît qu'il conviendrait d'effectuer au plus tôt une étude par simulation; le mieux serait de terminer les travaux d'ici un an environ. Un petit groupe de parties intéressées devrait être convoqué dès que possible pour définir les objectifs et les limites réalistes de l'étude par simulation. Selon le sous-groupe, la simulation devrait au minimum prendre en compte les points suivants :

- i) le coût (en heure-navire, par ex.) des diverses conceptions (et le coût des divers degrés de placement au hasard);
- ii) les biais inhérents aux migrations verticales circadiennes du krill; et
- iii) les effets de la cohérence spatiale des distributions du krill qui sont différents en différentes directions.

Il pourrait également être bon de considérer si à un certain point, l'utilité marginale de la réduction de la variance devient limitée. Ceci pourrait être examiné en considérant le point où les résultats du modèle de rendement du krill deviendront plus sensibles à la variabilité du recrutement du krill qu'à l'incertitude liée à la biomasse de krill.

6.5 B. Manly et A. Murray annoncent qu'ils seraient prêts à développer l'étude par simulation en collaboration avec un collègue néo-zélandais expert en géostatistique. Ils ajoutent qu'ils aimeraient recevoir l'assistance de parties intéressées, et plus particulièrement si elles possèdent des jeux de données d'anciennes campagnes d'évaluation du krill. W. de la Mare se charge, avec le secrétariat, d'examiner l'utilité marginale de la réduction de la variance dans les estimations de biomasse.

6.6 Faute d'étude par simulation, le sous-groupe convient que l'utilisation de transects parallèles placés au hasard représenterait une conception prudente car les paramètres d'estimation de la variance fondés sur la conception et ceux fondés sur les modèles pourraient servir à analyser les données.

AVIS AU WG-EMM

7.1 Le sous-groupe fait la synthèse de ses recommandations.

Question 2 de l'ordre du jour

7.2 Remplacer le terme anomalie par le terme VOGON (valeur n'entrant pas dans les normes généralement observées) (paragraphe 2.9).

7.3 Effectuer une analyse en composantes principales pour certains sites et indices (paragraphe 2.12).

7.4 Examiner et ajuster l'indice de croissance des jeunes otaries (C2b) dont l'efficacité est mise en doute quant à sa capacité à mesurer le succès reproductif (paragraphe 2.16).

7.5 Entreprendre de nouvelles études pour développer, aux échelles spatio-temporelles appropriées, des séries d'indices du CEMP qui soient plus robustes pour identifier les VOGON que ne le sont les indices individuels (paragraphe 2.18).

7.6 Envisager de développer une procédure pour reconnaître les situations dans lesquelles une proportion élevée d'indices donne des "quasi-VOGON" (paragraphe 2.19).

7.7 Calculer les niveaux voulus d' $\alpha$  permettant d'identifier les VOGON sur la base de chaque indice, en considérant d'utiliser des valeurs dépassant 0,05 pour les données non normales (paragraphe 2.23).

#### Question 3 de l'ordre du jour

7.8 Modifier le modèle d'Agnew et Phegan (1995) pour en améliorer les aspects temporels (paragraphe 3.11).

7.9 Appliquer un nouvel indice de chevauchement des niches, tel que l'indice de Schroeder, à la sous-zone 48.1 (paragraphe 3.12).

7.10 Poursuivre l'étude du chevauchement réalisé par des analyses de la sensibilité, la saisie des nouvelles données de la sous-zone 48.1 et l'application à la sous-zone 48.3 (paragraphe 3.14).

7.11 Tenir compte de l'interaction manchots-pêche lors de l'ajustement de l'indice de chevauchement réalisé, pendant d'autres périodes susceptibles d'être critiques (paragraphe 3.15).

7.12 Soumettre de nouvelles données qui permettent de faire progresser les travaux mentionnés ci-dessus (paragraphe 3.3).

#### Question 4 de l'ordre du jour

7.13 Mettre au point des méthodes de comparaison des indices de comportement en mer d'un site à l'autre et d'une année à l'autre par des épreuves du caractère aléatoire (paragraphe 4.7 et 4.8).

7.14 Développer des indices récapitulant le comportement en mer, en faisant usage des données fournies par satellite (paragraphe 4.10), et étudier les propriétés de ces indices (paragraphe 4.9).

7.15 Se pencher sur les questions mentionnées aux paragraphes 7.13 et 7.14 avant de retenir les indices à entrer dans la banque de données du CEMP.

#### Question 5 de l'ordre du jour

7.16 Explorer divers cas possibles de valeurs manquantes avec un jeu complet de données du CEMP (paragraphe 5.2).

7.17 Recueillir au plus tôt des informations sur les raisons pour lesquelles il manque des valeurs dans les données du CEMP, selon les directives suggérées au paragraphe 5.3 (paragraphe 5.4).

7.18 Identifier les séries dans lesquelles il manque des données, et les méthodes par lesquelles celles-ci peuvent être interpolées afin d'obtenir une valeur indispensable pour l'examen d'un paramètre dans une série du CEMP (paragraphe 5.6).

7.19 Explorer la méthode d'analyse des séries multivariées dans lesquelles il manque des valeurs, afin de permettre d'effectuer de telles analyses à l'avenir (paragraphe 5.7 et 5.8).

#### Question 6 de l'ordre du jour

7.20 Mener une étude par simulation pour comparer l'espacement aléatoire à l'espacement systématique des transects pour la campagne d'évaluation synoptique du krill, et convoquer un comité qui définirait les buts et limites réalistes de l'étude (paragraphe 6.4).

7.21 Se préparer à utiliser le modèle de rendement du krill pour examiner l'utilité marginale de la réduction de l'incertitude dans l'estimation de la biomasse de krill (paragraphe 6.5).

7.22 À moins qu'une étude par simulation ne soit effectuée, espacer les transects au hasard dans la campagne d'évaluation synoptique (paragraphe 6.6).

#### CLÔTURE DE LA RÉUNION

8.1 Le rapport est adopté. En clôturant la réunion, le responsable remercie le Southwest Fisheries Science Center et R. Holt de leur accueil, ainsi que tous les participants.

#### RÉFÉRENCES

Agnew, D.J. et G. Phegan. 1995. Development of a fine-scale model of land-based predator foraging demands in the Antarctic. *CCAMLR Science*, 2: 99–110.

Manly, B.F.J. 1997. *Randomisation, Bootstrap and Monte Carlo Methods in Biology*, 2nd Edition Chapman and Hall, London.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27: 209–220.

Mielke, P.W., K. J. Berry et E.S. Johnson. 1976. Multi-response permutation procedures for *a priori* classifications. *Communications in Statistics*, A5: 1409–1424.

Tableau 1 : Échelles temporelles d'intégration de variables contrôlées des prédateurs.

2 - 10 ans	1 an	0,5 – 2 ans	Environ 6 mois (hiver)	1 – 6 mois (été)
Survie des jeunes	Survie des adultes	Taille de la population	Poids des adultes à l'arrivée	Durée de la sortie alimentaire  Taux de croissance des jeunes  Poids au sevrage/à la première mue  Succès de la reproduction  Composition du régime alimentaire  Poids du repas

Tableau 2 : Données de l'île Bird utilisées pour fournir un exemple d'analyse à plusieurs variables et de production d'indices récapitulatifs. L'indice de la durée des sorties alimentaires des otaries est précédé du signe moins afin que les valeurs les moins négatives représentent les bonnes conditions.

Année	C1 Durée des sorties alimentaires d'otaries femelles * (-1)	C2b Croissance des jeunes otaries femelles	C2b Croissance des jeunes otaries mâles	A6a Succès de la reproduction des gorfous macaroni	A6a Succès de la reproduction des manchots papous	A7 Poids des gorfous macaroni à la 1 <sup>ère</sup> mue	A7 Poids des manchots papous à la 1 <sup>ère</sup> mue	A8 Proportion de krill dans le régime alimentaire des gorfous macaroni	A8 Proportion de krill dans le régime alimentaire des manchots papous	Poids des otaries femelles à la dernière pesée	Poids des otaries femelles à la dernière pesée
1977				0.476	0.598						
1978				0.250	0.006						
1979				0.473	0.294						
1980				0.602	0.577						
1981				0.527							
1982				0.509	0.048						
1983				0.491	0.506						
1984				0.092	0.285						
1985				0.477	0.428						
1986				0.504	0.418						
1987				0.361	0.427						
1988				0.364	0.468						
1989				0.608	0.457	3450	5464				
1990	-80	1.89	2.38	0.592	0.356	3237	5800	0.998	0.594	11.24	13.07
1991	-203	2.77	3.26	0.583	0.010	3112	5043	0.694	0.191	11.48	12.73
1992	-94	2.14	2.58	0.408	0.631	3507	5791	0.988	0.499	12.84	14.81
1993	-123	2.67	3.69	0.553	0.894	3318	5482	0.833	0.845	12.45	15.02
1994	-469	2.48	2.66	0.456	0.040	2913	5065	0.112	0.129	10.66	11.89
1995	-103	2.12	3.31	0.505	0.583	3025	5239	0.536	0.544	11.21	13.92
1996	-90	2.25	2.78	0.445	0.789	3179	5502	0.999	0.243	11.84	14.31
1997	-97	2.25	2.95	0.484	0.500	3300	5960	0.986	0.362	11.93	14.95



Tableau 4 : Détermination des seuils de  $\alpha$  requis pour détecter des VOGON ayant été identifiés sur le plan biologique.

	Manchot papou	Albatros
Années démontrant des VOGON significatifs sur le plan biologique	1978, 1982, 1991, 1994	1980, 1984, 1987, 1991, 1994
Années exclues de l'analyse - motif de l'exclusion	1981 – aucune donnée	1988, 1995 – conditions environnementales défavorables identifiées comme cause principale d'échec de reproduction
Longueur de la série chronologique après ajustement	20 ans	20 ans
Année possédant le VOGON le moins extrême	1982	1987
Moyenne de l'indice transformé	-0.7210	-1.4650
Écart-type de l'indice transformé	1.8508	2.1379
Valeur du VOGON le moins extrême	-2.9874	-2.3259
Valeur critique requise pour détecter le VOGON le moins extrême	1.2245	0.4027
Valeur de $\alpha$ pour valeur critique	0.22	0.69

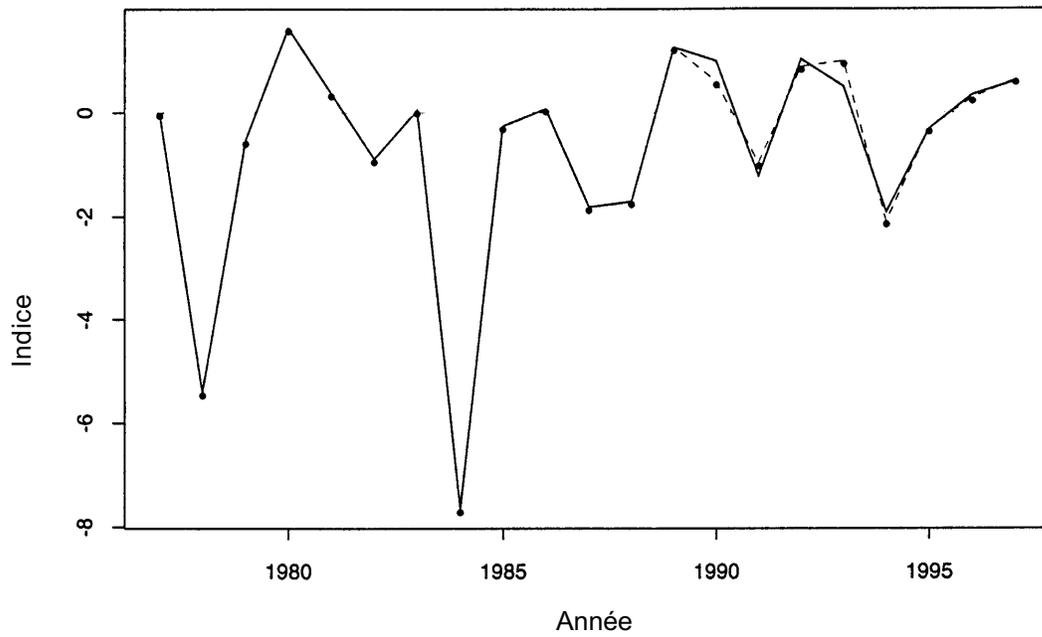


Figure 1 : Illustration de l'indice simple des espèces dépendantes en Géorgie du Sud, combinant les données sur les otaries et les manchots en rapport avec la saison de reproduction. La ligne continue représente l'indice fondé sur toutes les données; celle en tirets révèle l'effet de la suppression des données sur la croissance des jeunes otaries.

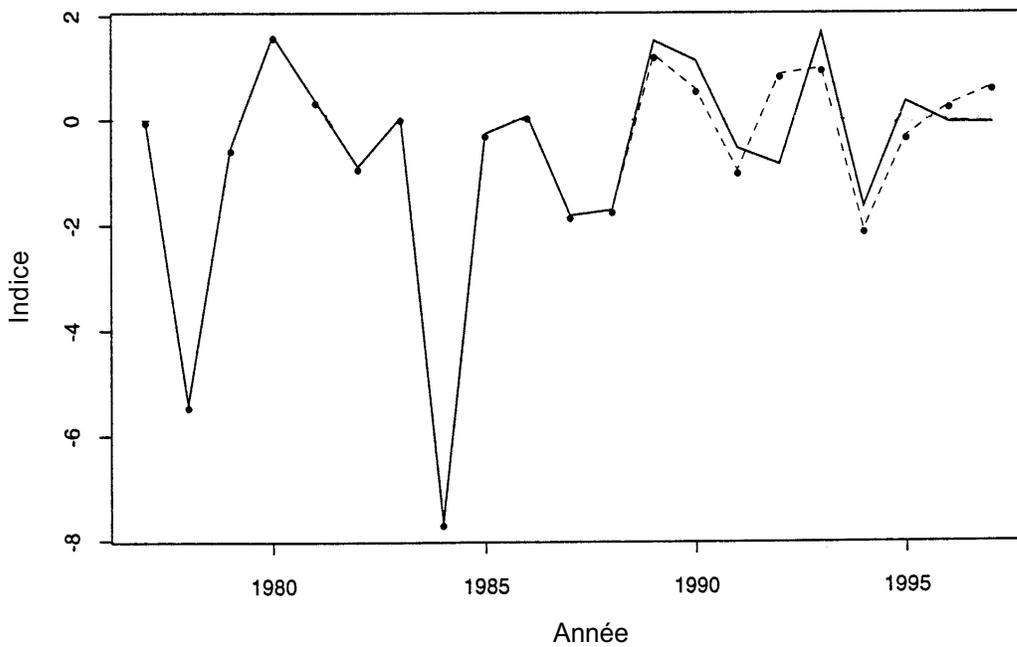


Figure 2 : Illustration de l'indice simple des espèces dépendantes en Géorgie du Sud, combinant les données sur les otaries et les manchots en rapport avec la saison de reproduction. La ligne continue représente l'indice, fondé uniquement sur les données sur le succès de la reproduction des manchots; celle en tirets révèle l'effet de l'inclusion de toutes les autres données, sauf celles sur la croissance des jeunes otaries.

**ORDRE DU JOUR**

Sous-groupe sur les statistiques  
(La Jolla, États-Unis, 14 - 18 juillet 1997)

1. Introduction
  - i) Ouverture de la réunion
  - ii) Organisation de la réunion et adoption de l'ordre du jour
2. Nouvel examen de l'identification des anomalies dans les indices du CEMP
  - i) Examen à jour des séries chronologiques d'indices du CEMP
  - ii) résumé des problèmes récents/suggestions pour l'identification des anomalies (divers problèmes et suggestions figurent dans SC-CAMLR-XV, annexe 4, paragraphes 4.58 à 4.61, 4.70, 4.72, 4.75 et 7.1)
  - iii) Discussion et mise en place de méthodes propres à traiter le problème/donner suite aux suggestions pour l'identification des anomalies (SC-CAMLR-XV, paragraphe 5.38 ii))
3. Évaluation critique des hypothèses et valeurs paramétriques du modèle de chevauchement réalisé d'Agnew et Phegan (1995)
  - i) Examen et synthèse des données et analyses soumises en réponse à SC CIRC 97/2 (sous-groupe sur les statistiques du WG-EMM - demande de données et d'analyses)
  - ii) Évaluation des hypothèses et valeurs paramétriques utilisées dans le modèle d'Agnew et Phegan (SC-CAMLR-XV, paragraphe 5.38 iv))
  - iii) Ajustement possible du modèle d'Agnew et Phegan grâce aux données présentées en réponse à SC CIRC 97/2 ou développement d'un nouvel indice de chevauchement réalisé
4. Développement d'indices de comportement en mer et méthodes permettant de les dériver par l'analyse des jeux de données des échantillons
  - i) Examen des échelles spatio-temporelles appropriées au développement d'indices utiles (des informations de support sur cette question figurent dans SC-CAMLR-XV, annexe 4, paragraphes 3.61 à 3.65 et 7.58)
  - ii) Étude des jeux de données et analyses des échantillons (SC-CAMLR-XV, annexe 4, paragraphes 4.44 et 7.58)

- iii) Développement d'indices et de méthodes pour leur calcul par l'analyse des jeux de données des échantillons (SC-CAMLR-XV, paragraphe 5.38 i))
  - iv) Formulation d'avis sur les meilleurs indices à inclure dans la banque de données du CEMP (SC-CAMLR-XV, annexe 4, paragraphes 4.44 et 7.58)
5. Méthodes propres à faire face aux données manquantes dans les jeux de données multiples
- i) Examen des méthodes d'interpolation des données manquantes dans les matrices des séries chronologiques des indices du CEMP collectées dans un groupe de colonies de prédateurs (SC-CAMLR-XV, paragraphe 5.38 iii) et annexe 4, paragraphe 4.63)
6. Conception des campagnes d'évaluation synoptiques
7. Avis au WG-EMM
8. Clôture de la réunion.

## LISTE DES PARTICIPANTS

Sous-groupe sur les statistiques  
(La Jolla, USA, du 14 au 18 juillet 1997)

BOYD, Ian (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom i.boyd@bas.ac.uk
CROXALL, John (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom
DE LA MARE, William (Dr)	Australian Antarctic Division Channel Highway Kingston Tasmania 7050 Australia bill_de@antdiv.gov.au
EVERSON, Inigo (Dr)	British Antarctic Survey High Cross, Madingley Road Cambridge CB3 0ET United Kingdom i.everson@bas.ac.uk
HOLT, Rennie (Dr)	US AMLR Program Southwest Fisheries Science Center PO Box 271 La Jolla, Ca. 92038 USA rholt@ucsd.edu
ICHI, Taro (Mr)	National Research Institute of Far Seas Fisheries Orido 5-7-1, Shimizu Shizuoka 424 Japan ichii@enyo.affrc.go.jp
MANLY, Bryan (Dr)	University of Otago PO Box 56 Dunedin New Zealand bmanly@maths.otago.ac.nz

MILLER, Denzil (Dr)

Sea Fisheries Research Institute  
Private Bag X2  
Roggebaai 8012  
South Africa  
dmiller@sfri.wcape.gov.za

MURRAY, Alastair (Dr)

British Antarctic Survey  
High Cross, Madingley Road  
Cambridge CB3 0ET  
United Kingdom  
a.murray@bas.ac.uk

WATTERS, George (Dr)

US AMLR Program  
Southwest Fisheries Science Center  
PO Box 271  
La Jolla, CA 92038  
USA  
gwatters@amlr.ucsd.edu

Present address:  
Inter-American Tropical Tuna Commission  
8604 La Jolla Shores Drive  
La Jolla CA 92038  
USA  
gwatters@iattc.ucsd.edu

SECRETARIAT:

RAMM, David (Dr) (Data Manager)  
MARAZAS, Rosalie (Secretary)

CCAMLR  
23 Old Wharf  
Hobart Tasmania 7000  
Australia  
ccamlr@ccamlr.org

**LISTE DES DOCUMENTS**

Sous-groupe sur les statistiques  
(La Jolla, USA, du 14 au 18 juillet 1997)

WG-EMM-Stats-97/1	PROVISIONAL AND ANNOTATED PROVISIONAL AGENDA FOR THE 1997 MEETING OF THE WG-EMM SUBGOUPE ON STATISTICS
WG-EMM-Stats-97/2	LIST OF PARTICIPANTS
WG-EMM-Stats-97/3	LIST OF DOCUMENTS
WG-EMM-Stats-97/4	DEVELOPMENT OF INDICES OF AT-SEA BEHAVIOUR I.L. Boyd (UK)
WG-EMM-Stats-97/5	DIET AND FORAGING RANGE OF PENGUINS AND FUR SEALS AT SOUTH GEORGIA J.P. Croxall, I.L. Boyd, K. Reid and P.N. Trathan (UK)
WG-EMM-Stats-97/6	TESTS FOR ANOMALOUS YEARS IN THE CCAMLR INDEX SERIES (DRAFT) B.F. Manly and D. MacKenzie (New Zealand)
WG-EMM-Stats-97/7	SOME CONSIDERATIONS FOR THE FURTHER DEVELOPMENT OF STATISTICAL SUMMARIES OF CEMP INDICES W.K. de la Mare (Australia)
WG-EMM-Stats-97/8	TREATMENT OF MISSING VALUES IN CEMP DATA SETS A. Murray (UK)
OTHER DOCUMENTS	
WG-EMM-97/25	CEMP INDICES 1997: SECTIONS 1 TO 3 Secretariat

**RÉSULTATS D'UNE ANALYSE EN COMPOSANTES PRINCIPALES  
DES DONNÉES SUR L'ÎLE BIRD DE 1990 À 1997.**

Les variables sont données dans le même ordre qu'au tableau 2; les abréviations correspondant aux noms anglais.

Données sur l'île Bird (non transformées)

Axe d'ACP	1	2	3	4	5	6	7
Eigenvalue	5.83	2.19	1.36	0.82	0.47	0.20	0.13
% du total	53.02	19.92	12.32	7.46	4.27	1.78	1.22
% cumulé	53.02	72.94	85.26	92.72	96.99	98.78	100.00

Eigenvectors (pondération des composantes)

SEALFD (C1)	0.36	0.02	0.27	-0.02	-0.49	0.26	-0.33
SEALPG-F (C2b)	-0.16	0.51	-0.28	0.45	0.03	-0.12	0.35
SEALPG-M (C2b)	0.02	0.65	-0.04	-0.13	-0.20	-0.25	-0.35
MACBS (A6a)	-0.06	0.29	0.73	0.26	0.04	-0.06	0.17
GENBS (A6a)	0.34	0.15	-0.16	-0.47	-0.13	0.13	0.65
MACFW (A7)	0.37	-0.05	-0.10	0.37	0.34	0.16	-0.17
GENFW (A7)	0.34	-0.29	0.10	0.10	0.17	-0.74	0.08
MACPK (A8)	0.36	-0.09	0.17	0.34	-0.34	0.09	0.33
GENPK (A8)	0.27	0.27	0.31	-0.36	0.61	0.13	-0.02
SEALWT-F	0.35	0.14	-0.31	0.28	0.19	0.31	-0.12
SEALW-M	0.38	0.14	-0.21	-0.12	-0.16	-0.38	-0.17

Valeurs des composantes principales

1990	0.22	-0.60	0.90	0.03	0.15	0.04	0.08
1991	-0.88	0.50	0.17	0.60	-0.19	0.10	-0.08
1992	0.99	-0.44	-0.50	0.16	0.24	0.18	-0.13
1993	0.71	1.07	-0.00	-0.09	0.26	-0.03	0.12
1994	-1.74	-0.29	-0.36	-0.14	0.26	-0.07	0.07
1995	-0.21	0.23	0.18	-0.61	-0.17	0.05	-0.19
1996	0.32	-0.25	-0.30	-0.10	-0.42	0.10	0.21
1997	0.59	-0.21	-0.08	0.16	-0.12	-0.37	-0.07

Diagramme des composantes principales de chaque année

