



THÈSE

pour obtenir le grade de
Docteur de l'École des Mines de Paris
Spécialité "Géostatistique"



Contributions géostatistiques à la biologie halieutique

Mathieu WOILLEZ

14 Décembre 2007

Jury:

M. Laurent CARRARO

Rapporteur

M. Francis LALOE

Rapporteur

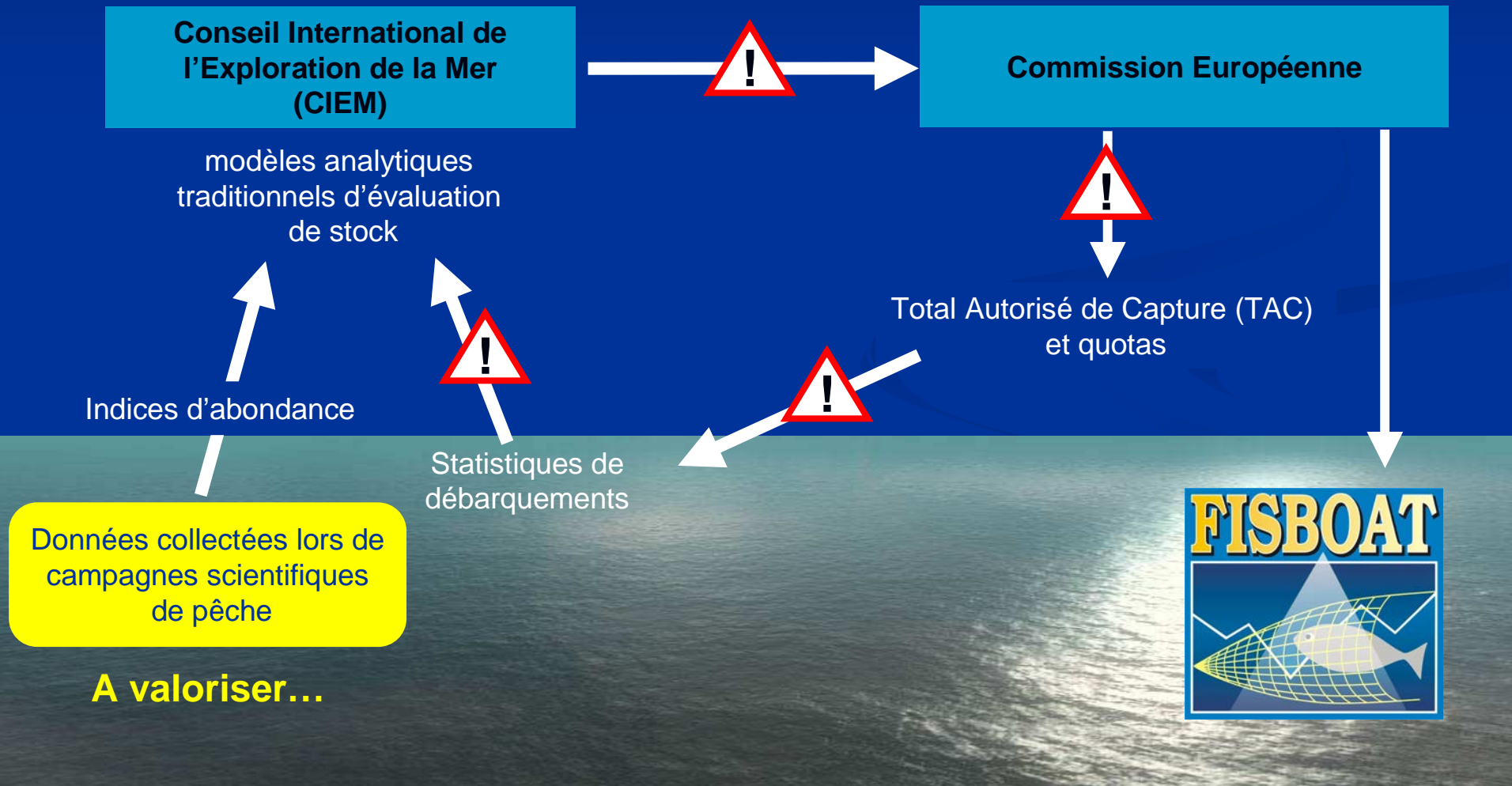
M. Pierre PETITGAS

Examineur

M. Jacques RIVOIRARD

Directeur de thèse

La gestion des pêches dans l'Atlantique Nord-Est



Plan

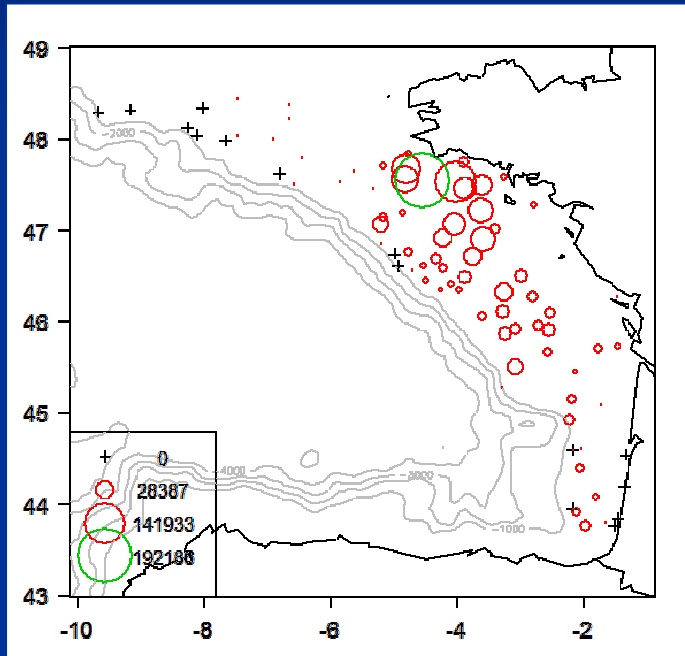
- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Plan

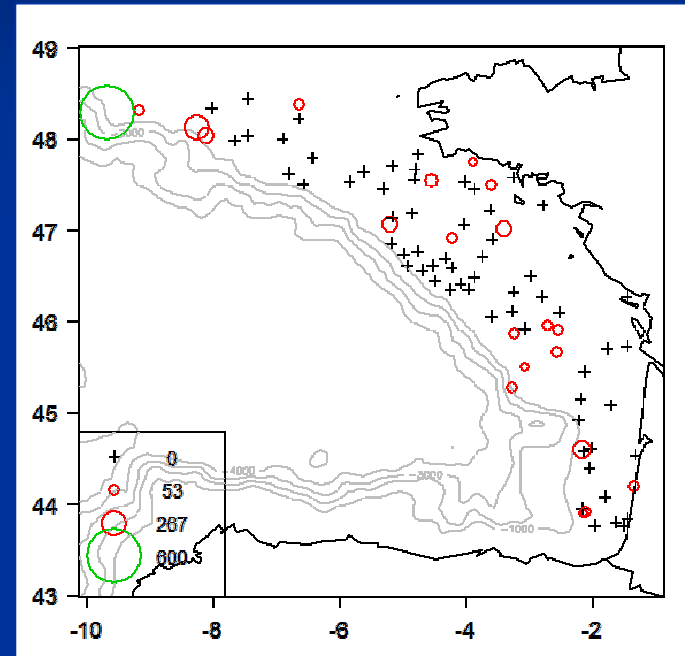
- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Contexte de suivi de population à limite diffuse

Campagne de chalut de fond, golfe de Gascogne, mesure de densités de merlu



Age 0 année 2004



Age 50+ année 2004

Couvrir la zone de présence

Incidence des valeurs nulles

Contexte de suivi de population à limite diffuse

- Avoir des indices spatiaux qui ne dépendent pas de la délimitation de la zone de présence de la population.

Contribution nulle des échantillons de valeur nulle :

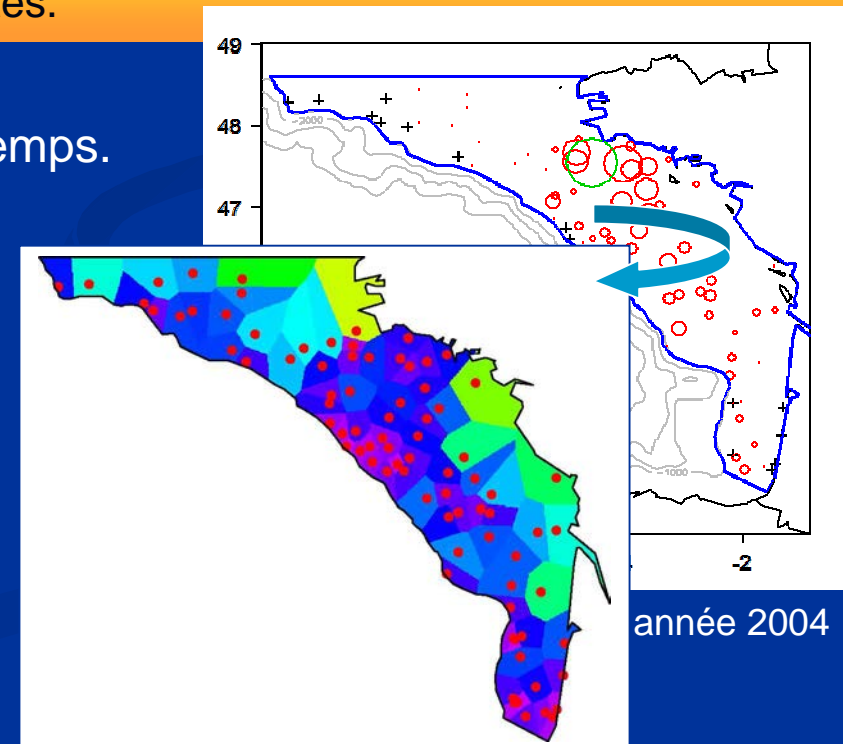
- Utilisation de moyennes pondérées par la probabilité d'un individu à être en une position x (i.e. $z(x)/Q$).
- Utilisation de sommes continues ou discrètes.

- Avoir une série comparable au cours du temps.

Pondération des échantillons par leur surface d'influence :

- Surfaces d'influence calculées dans le domaine maximal échantillonné.

⇒ Outils génériques capables de traiter différentes campagnes dans différentes régions.



Indices spatiaux (1)

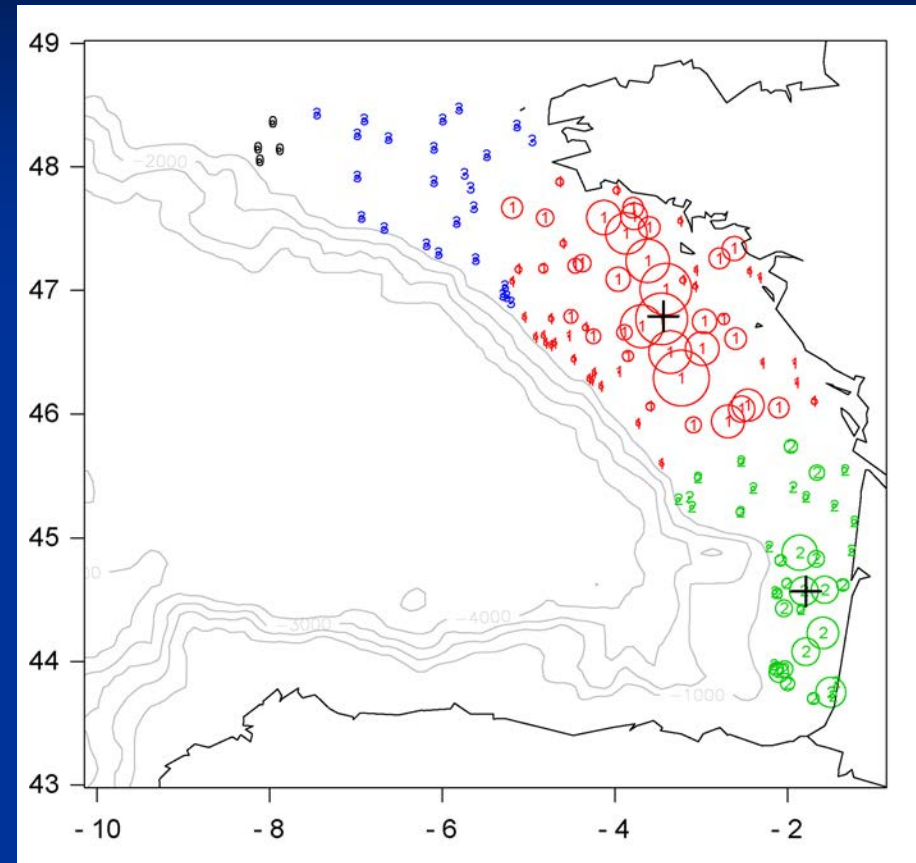
Centre de gravité : position moyenne de la population.

Inertie : dispersion de la population autour du centre de gravité

Isotropie : élongation de la distribution spatiale de la population.

Statistiques par individu (Bez, 1997)

Nombre de patch : Agrégation de poissons supérieure à l'échelle du banc.

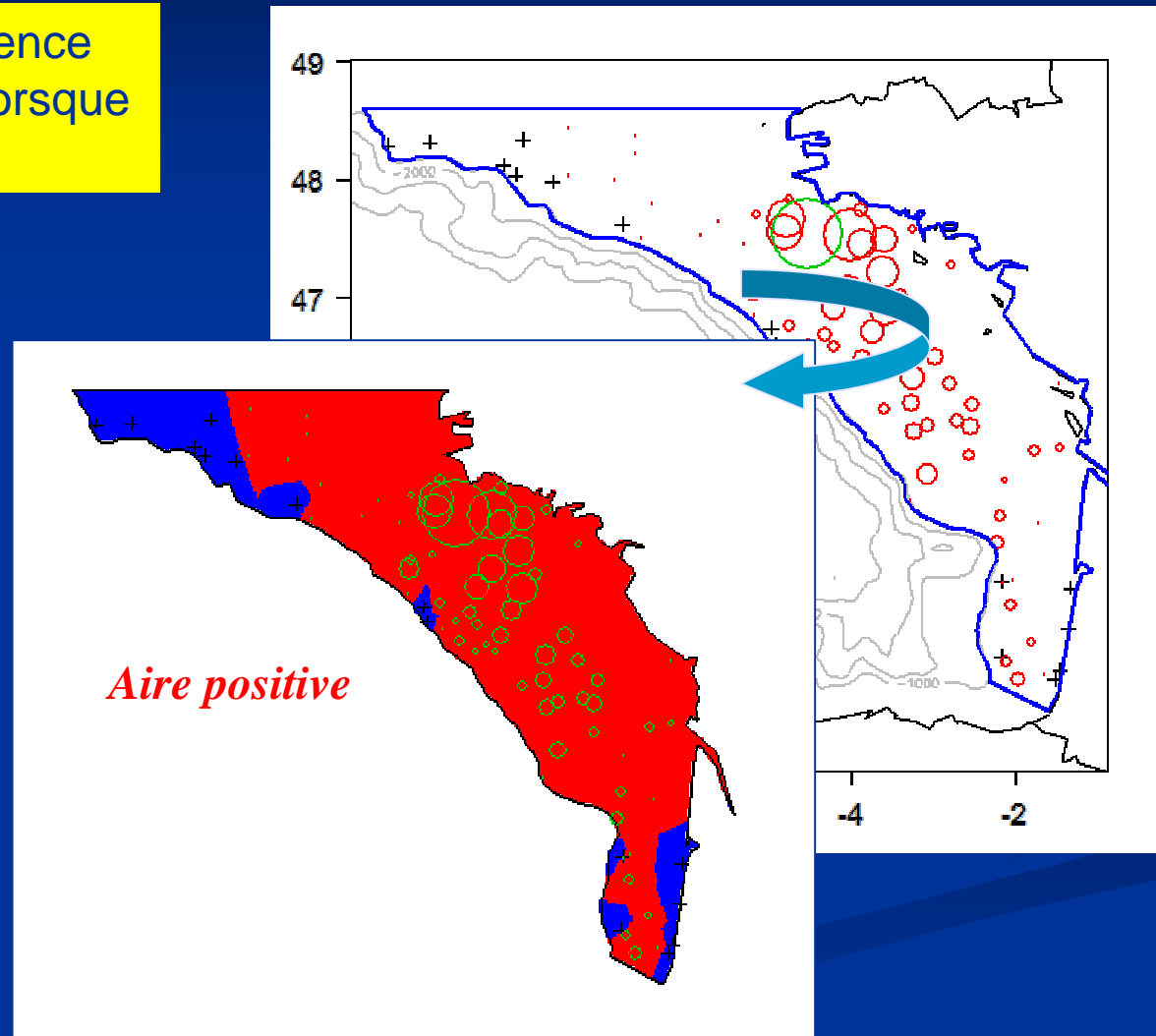


Age 0 année 1988

Indices spatiaux (3)

Aire positive : aire de présence occupée par le stock, même lorsque la densité est faible.

$$\text{Aire positive} = \sum_i s_i 1_{z_i > 0}$$



Indices spatiaux (4)

Aire équivalente : aire qui serait couverte par la population, si tous les individus avaient la même densité, égale à la densité moyenne par individu.

$$\text{Aire équivalente} = \frac{Q}{\int z(x) \frac{z(x)}{Q} dx}$$

1	2
3	0



Aire positive = 3

1	2
3	0

1	2
3	0



$$Q = \int z(x) dx \quad Q = 6$$

$$m_i = 1 \times \frac{1}{6} + 2 \times \frac{2}{6} + 3 \times \frac{3}{6}$$

$$m_i = \frac{14}{6} = 2.33$$

$$\text{Aire équivalente} = \frac{6}{2.33} = 2.57$$

2.33	
	0

(Bez et Rivoirard, 2001)

Indices spatiaux (6)

Aire d'étalement : elle mesure comment une population se distribue dans l'espace en prenant en compte les variations de densités de poisson.

T : aire cumulée occupée par les valeurs de densité, rangées dans un ordre décroissant

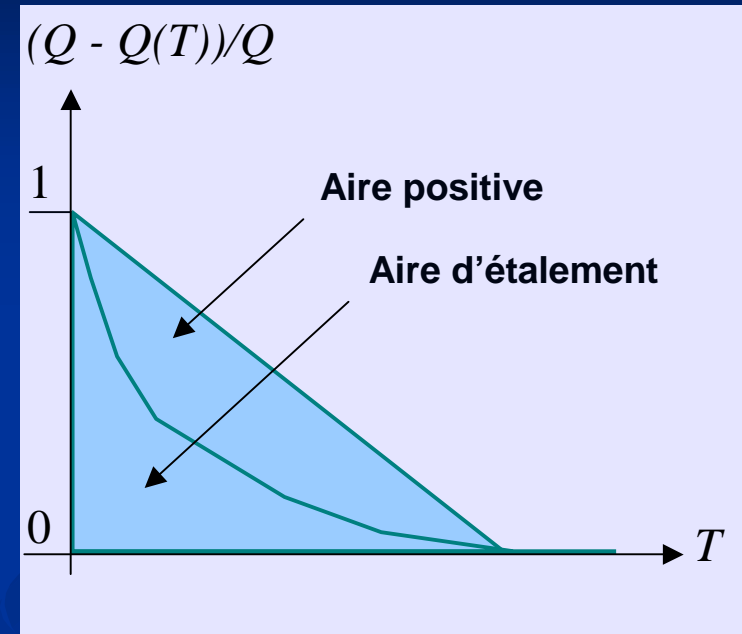
$Q(T)$: l'abondance cumulée correspondante

Q : l'abondance totale

$$\text{Aire d'étalement} = 2 \int \frac{Q - Q(T)}{Q} dT$$

Si les densités sont peu variables,
l'aire d'étalement est proche de l'aire positive.

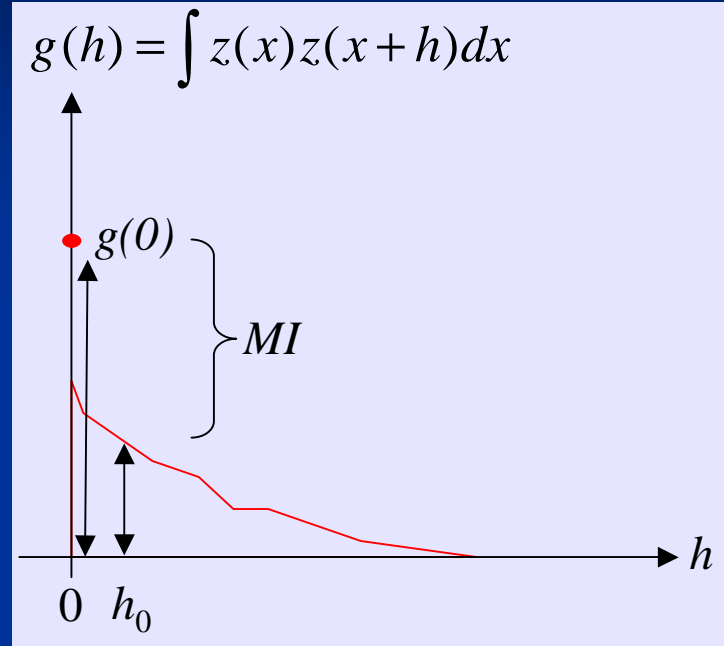
Si les densités ont une distribution dissymétrique,
l'aire d'étalement est faible par rapport à l'aire positive.



Indices spatiaux (5)

Indice de microstructure : il mesure la variabilité à petite échelle de la surface de densité de poissons.

$$MI = \frac{g(0) - g(h_0)}{g(0)}$$



Si la surface de densité est régulière,
l'indice de microstructure est proche de 0.

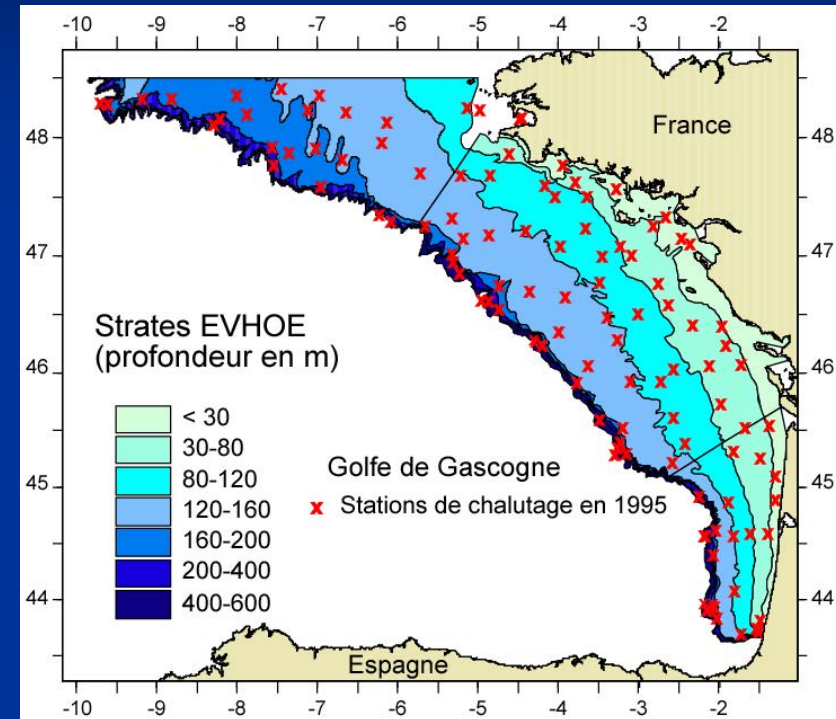
Si la surface de densité est irrégulière,
l'indice de microstructure est proche de 1.

Plan

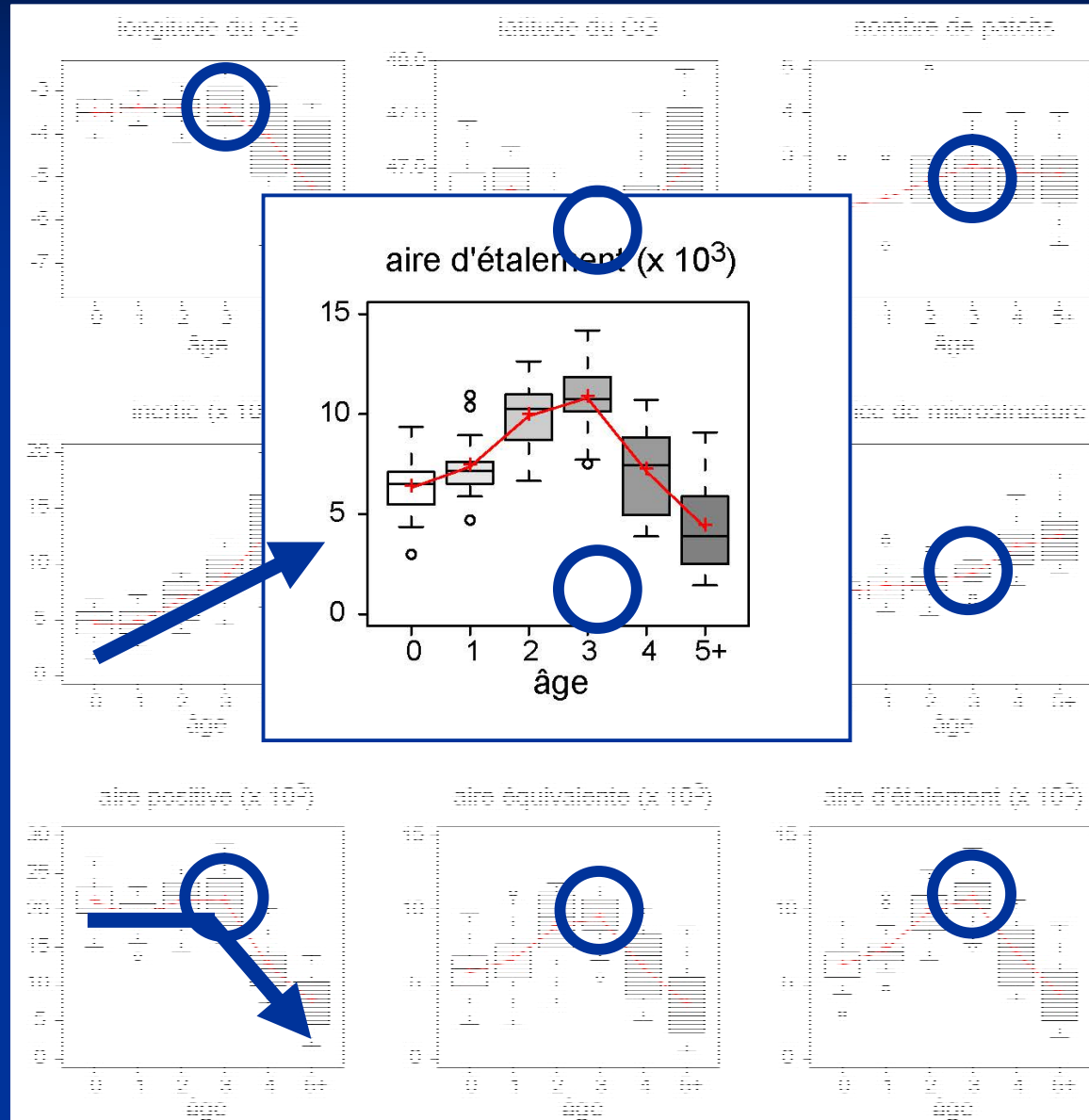
- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Les données de merlus

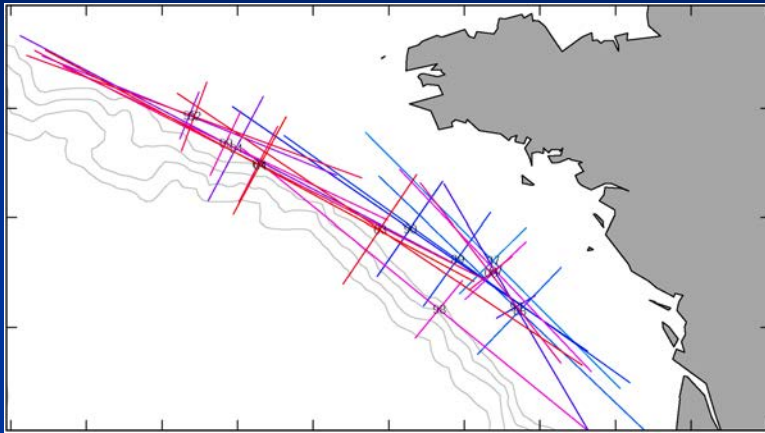
- Campagne de chalut de fond dans le golfe de Gascogne (série EVHOE).
- 15 années disponibles : 1987-2004 (années manquantes 1991, 1993 et 1996).
- Plan d'échantillonnage : aléatoire stratifié selon la latitude et la profondeur.
- Mesure de densités de merlu selon 6 classes d'âge (de 0 à 5+).



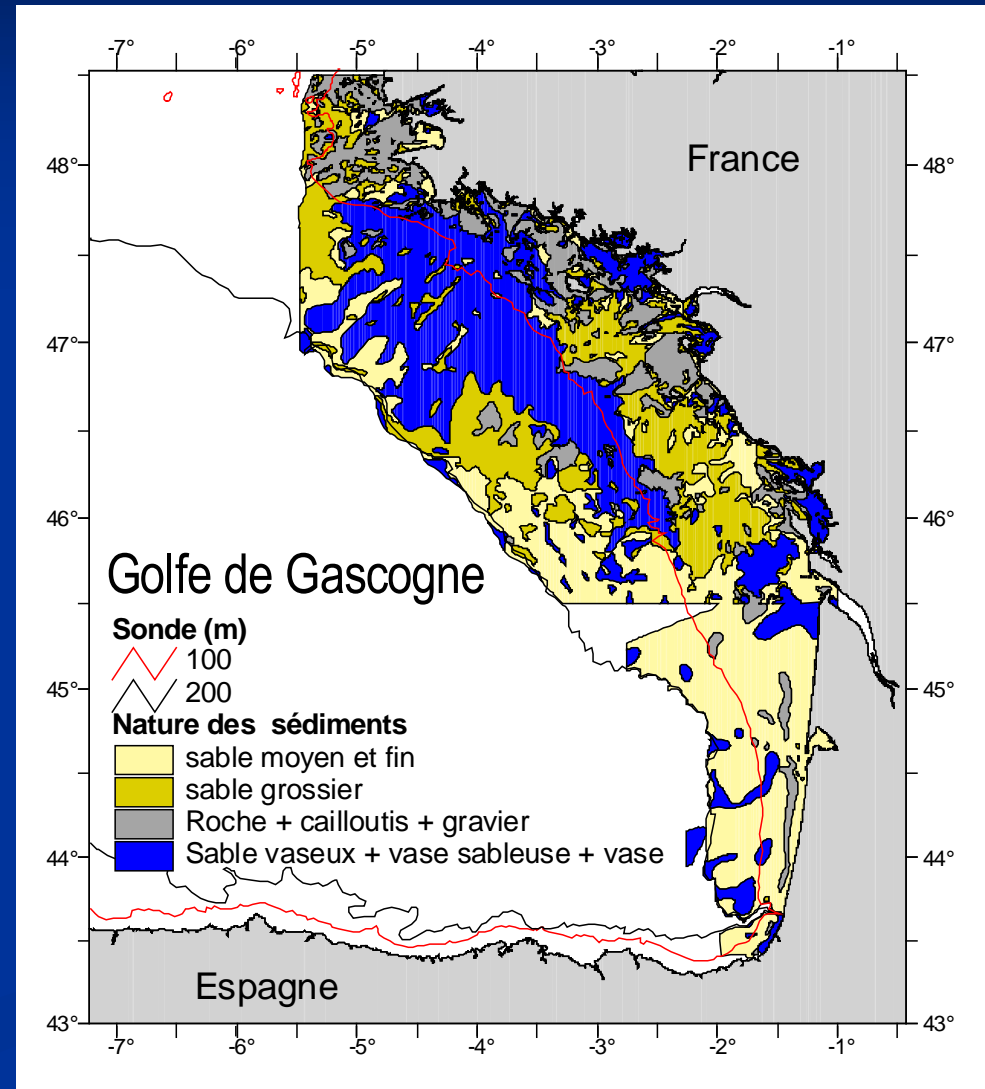
Analyse du comportement moyen (1)



Analyse du comportement moyen (2)

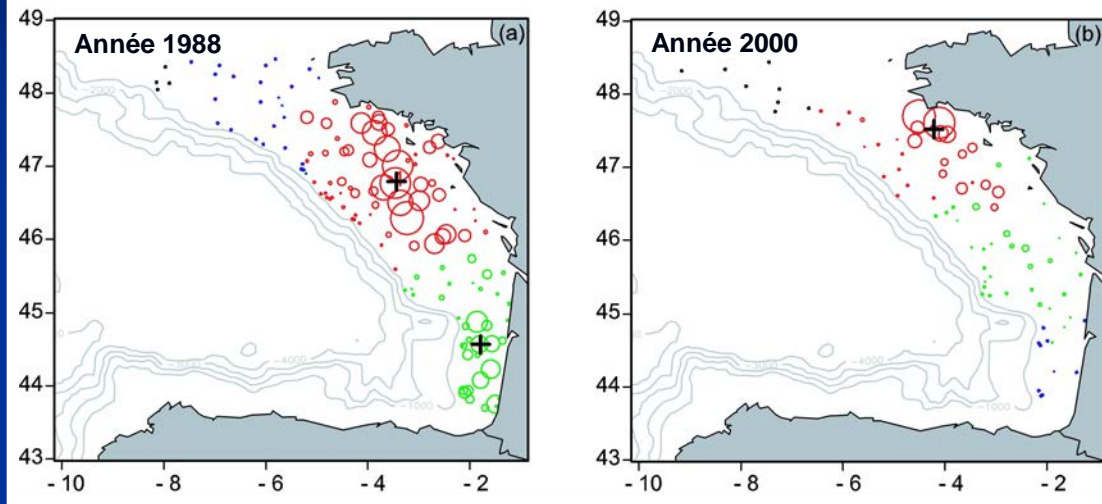


- Alignement des CGs d'âge 0 par rapport à la « grande vaseière ».
- Déplacement vers l'Ouest avec l'âge.
- Isotropie minimale pour les âges 0, 4 et 5+.
- Schéma ontogénique de distribution lié aux sédiments (âge 0), puis à la profondeur.



Détections des années hors normes, des années charnières et des tendances

Nombre de patches pour l'âge 0

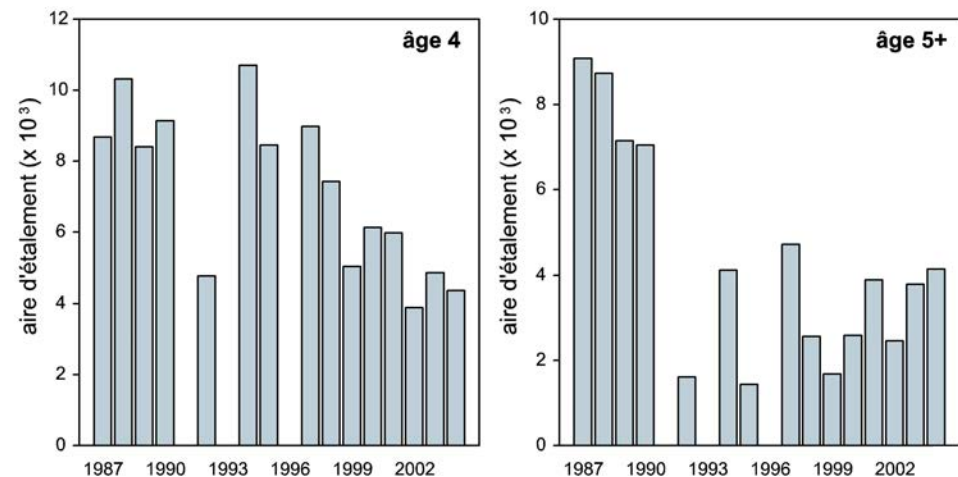


Deux patches en moyenne sur la série temporelle.

Un seul patch pour l'année 2000.

Diminution de l'aire équivalente pour l'âge 4 et 5+ respectivement à partir des années 1997 et 1990.

Séries temporelles de l'aire d'étalement

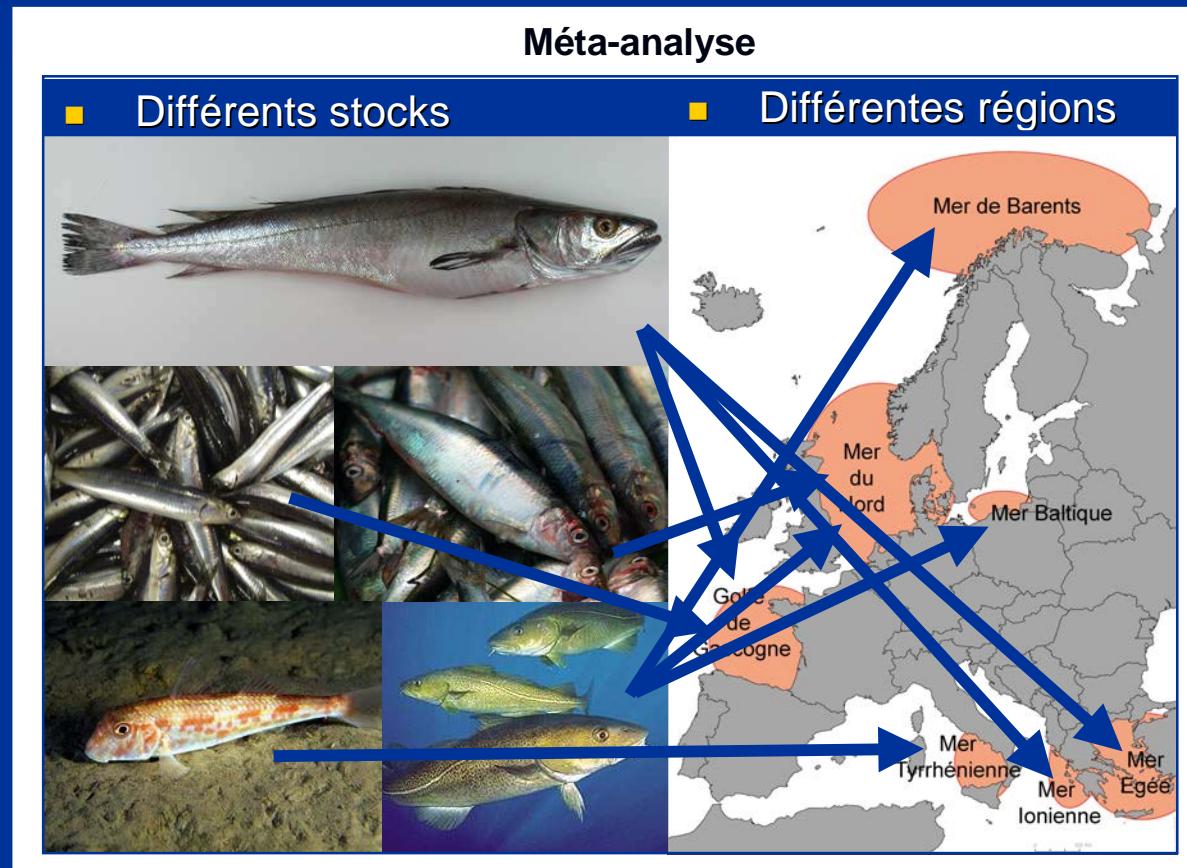


Plan

- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Objectifs

- Chercher des corrélations sur une variété de stocks entre indices spatiaux et la dynamique de la population au travers de l'abondance, du recrutement et de la mortalité.
- Fournir des indicateurs additionnels de l'état d'une population.



Résultats (1)

- Corrélations considérées significatives pour une valeur $P \leq 0.05$
- Corrélations entre indices spatiaux des matures et abondance des recrues suivantes :
 - On attend en moyenne :
 0.05×9 (spatial indices) $\times 10$ (stocks) = 4.5 corrélations significatives dans le cas où les vraies corrélations étaient nulles.
 - Nous avons trouvés 10 corrélations significatives.

⇒ L'inertie.

⇒ La morue de la mer Baltique.

Résultats (2)

- Corrélations entre indices spatiaux et abondance de l'année :

- 720 corrélations testées, soit 36 attendues en moyenne.
- 96 corrélations significatives trouvées.

⇒ Les indices d'aire.

⇒ La morue de la mer du Nord.

- Corrélations entre indices spatiaux et mortalité l'année suivante :

- 540 corrélations testées, soit 27 attendues en moyenne.
- 34 corrélations significatives trouvées.

Plan

- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Introduction

Des indices en nombre conséquent représentant un stock cible au cours de sa période d'étude :

- Indices spatiaux pour chaque âge :
 - La localisation (longitude et latitude du centre de gravité, inertie, anisotropie)
 - L'extension spatiale (aire positive, aire équivalente, aire d'étalement)
 - La microstructure (indice de microstructure)

Mais aussi...

- Indices d'abondance (en log) :
 - Abondance totale, abondance des recrues, abondance à l'âge.
- Indices biologiques :
 - Quantile de la longueur (Lbar, L25, L75)
 - Longueur de maturité (L.M50)
 - Mortalité (Z)

Comment diagnostiquer l'état d'un stock ?

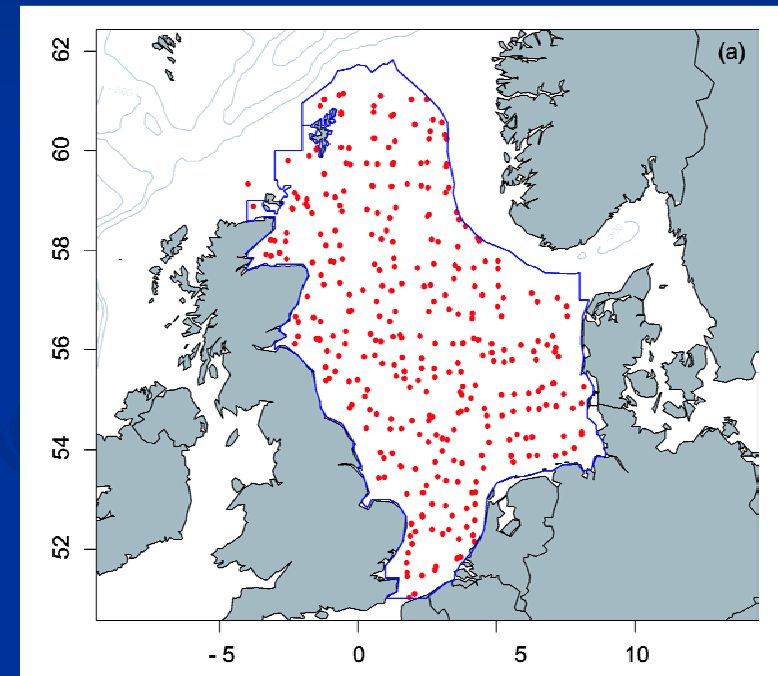
- Besoin de réduire le nombre d'indices et de garder ceux qui conduisent à un diagnostic fiable.
- La corrélation dans le temps utilisée comme critère pour :
 - Sélectionner les indices les plus continus.
 - Combiner ces indices en indicateurs continus de la population grâce à la méthode des « Min/max Autocorrelation Factors » (MAF).
- Les indicateurs combinés (les MAFs) nous permettront de :
 - Décrire l'état de la population.
 - Faire une projection dans le temps par modélisation de leur variogramme.

Les cas d'étude

La morue de la mer du Nord et l'anchois du golfe de Gascogne.



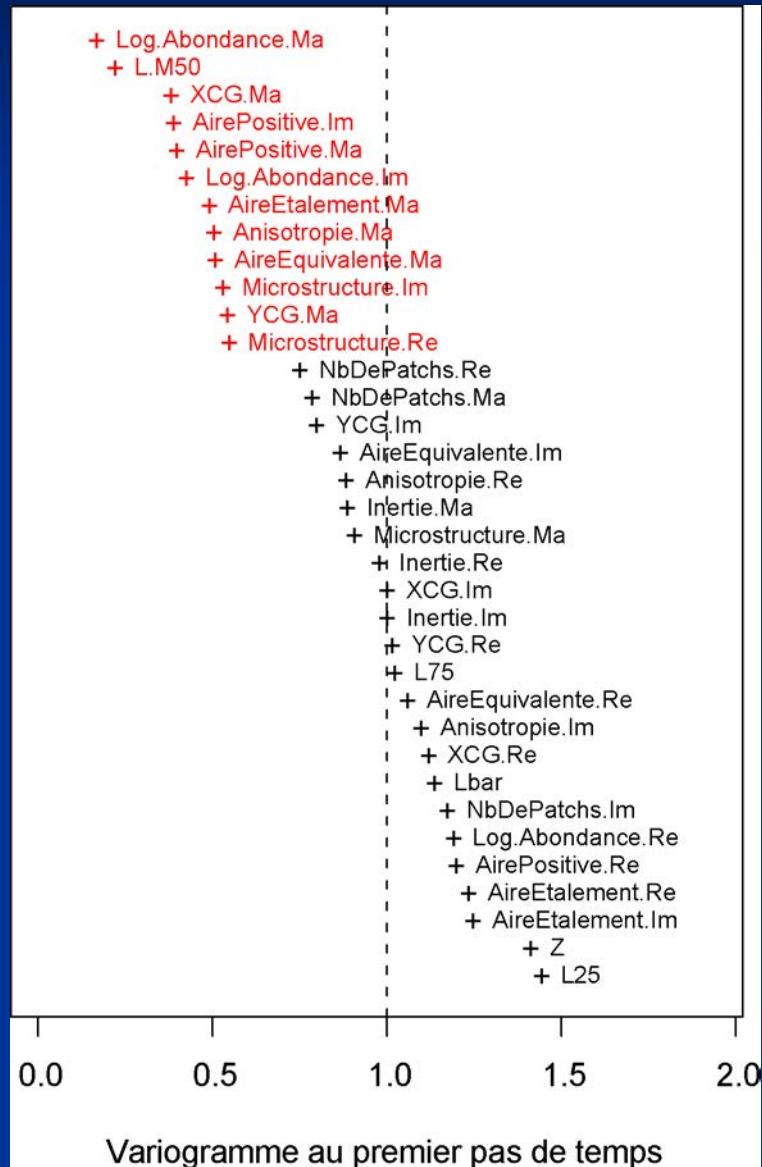
- Collectées à partir des campagnes internationales de chalut de fond (IBTS).
- 21 années disponibles : 1985-2005.
- Echantillonnage aléatoire stratifié avec 2 coups de chalut par carré CIEM.
- Mesure de densités de morues disponibles par âge (de 1 à 6 ans).



Indices spatiaux calculés par groupe représentatif :

- ⇒ Matures et immatures déterminés à partir d'une ogive de maturité.
- ⇒ Age 2 considéré comme représentatif des recrues.

L'étape de sélection



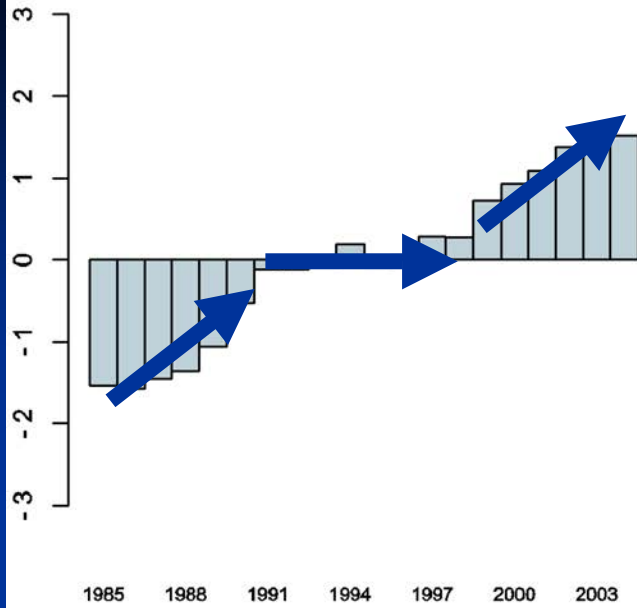
- 35 séries temporelles d'indices (normés à une variance de 1)
- 21 sur 35 avec $\gamma(1) \leq 1$
- 12 retenus ($0.167 \leq \gamma(1) \leq 0.547$) :
 - 1 trait d'histoire de vie
 - 2 indices d'abondance
 - 9 indices spatiaux

Min/max Autocorrelation Factors

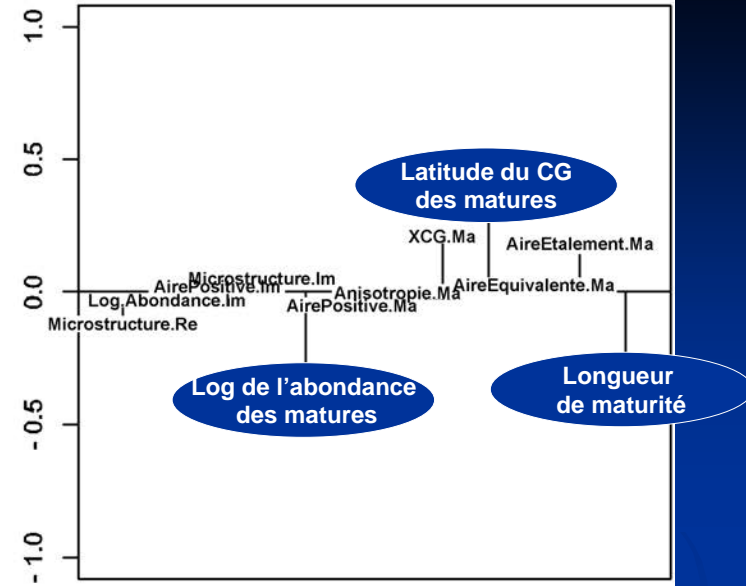
- Une méthode statistique multivariée, proche des classiques composantes principales, permettant à un jeu de variables initiales d'être décomposé en facteurs.
- Les facteurs (les MAFs) maximisent l'autocorrelation.
- Cas des séries temporelles courtes :
 - Surparamétrisation.

⇒ Robustification nécessaire

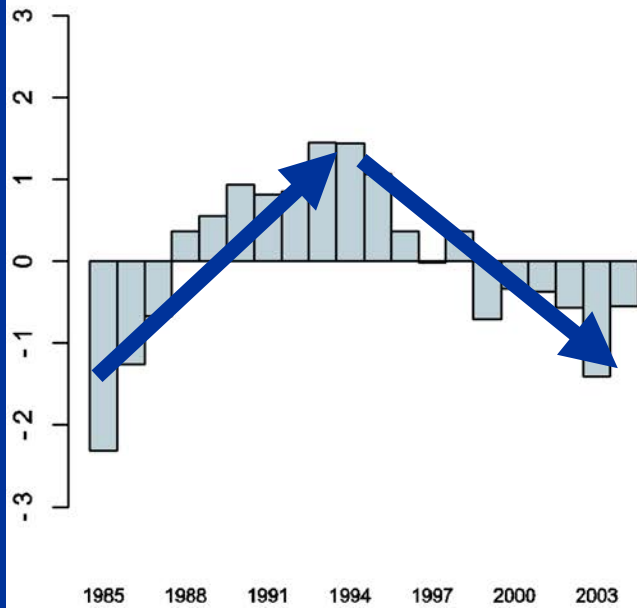
MAF 1



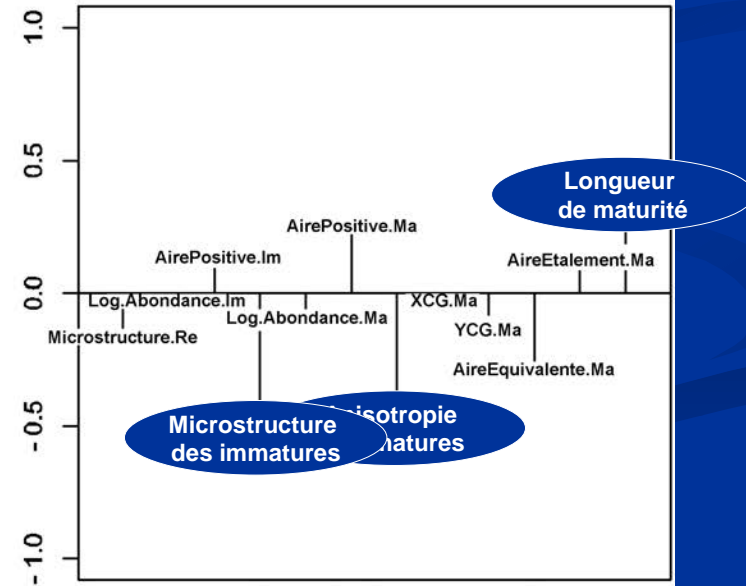
Coefficients



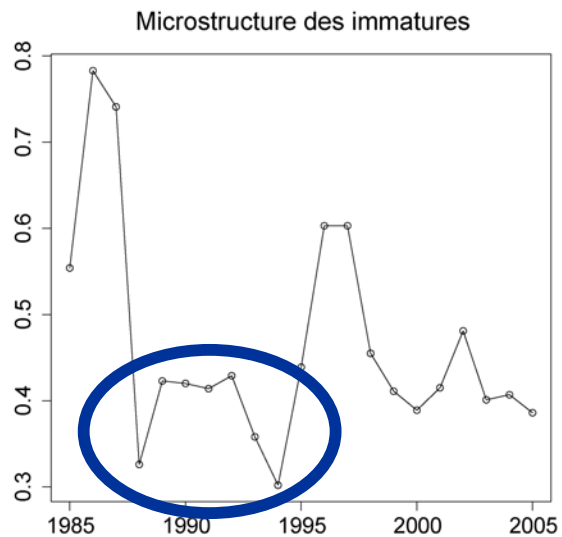
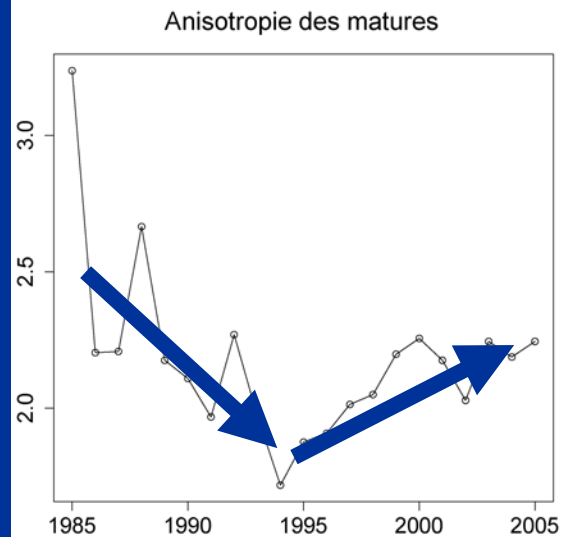
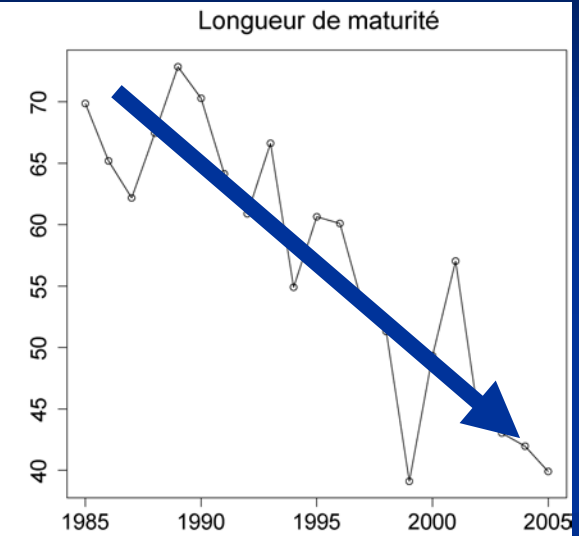
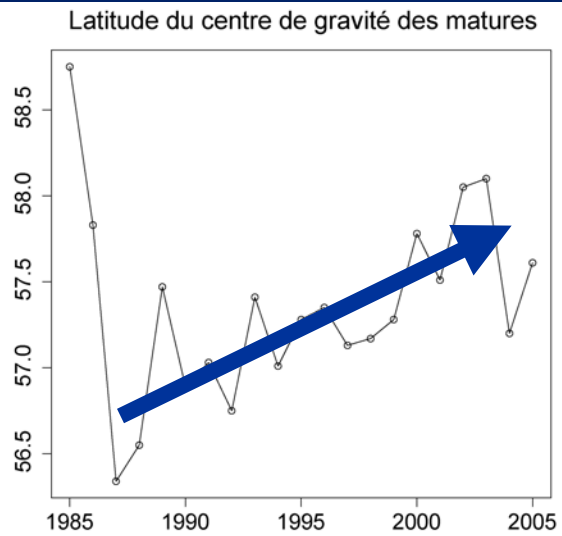
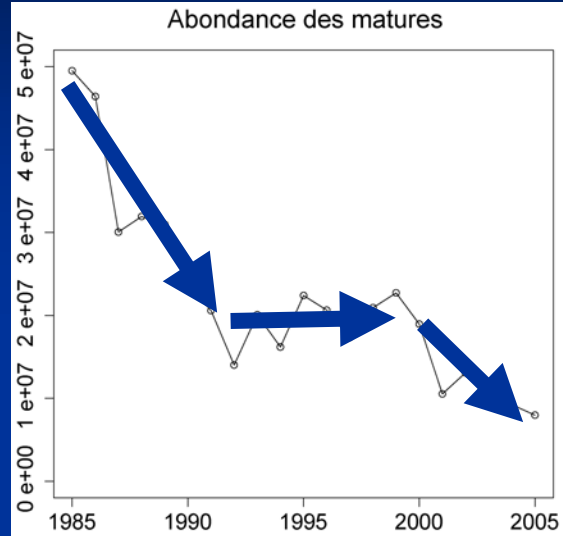
MAF 2



Coefficients



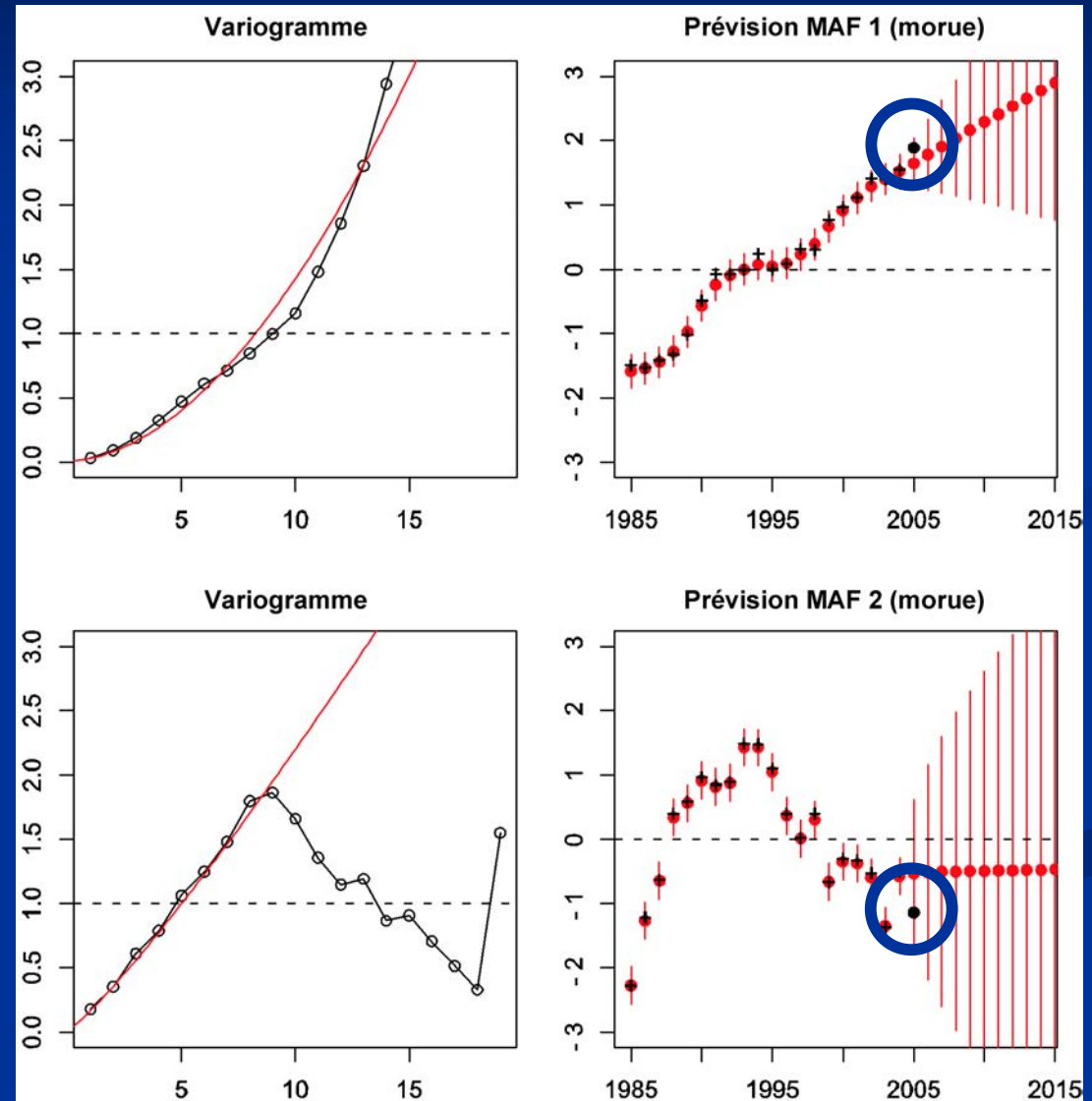
Retour aux séries temporelles des indices



MAF2

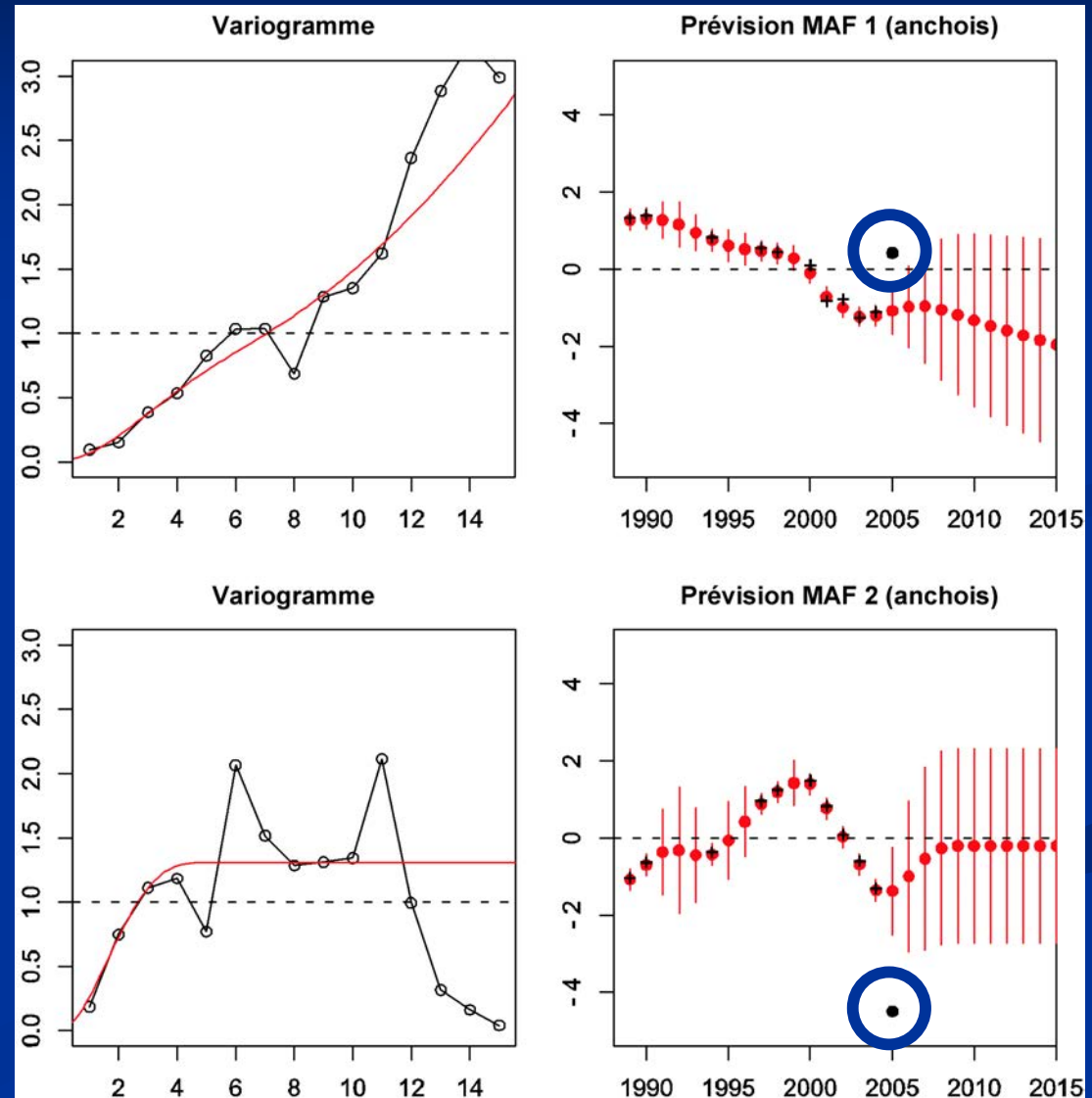
Prévisions de l'état d'un stock (1)

- Morue mer du Nord :
 - L'année 2005 (l'année prévue) en continuité
 - Pas de changement observé comparé à la prévision concernant l'évolution de l'état du stock (toujours dégradé).



Prévisions de l'état d'un stock (2)

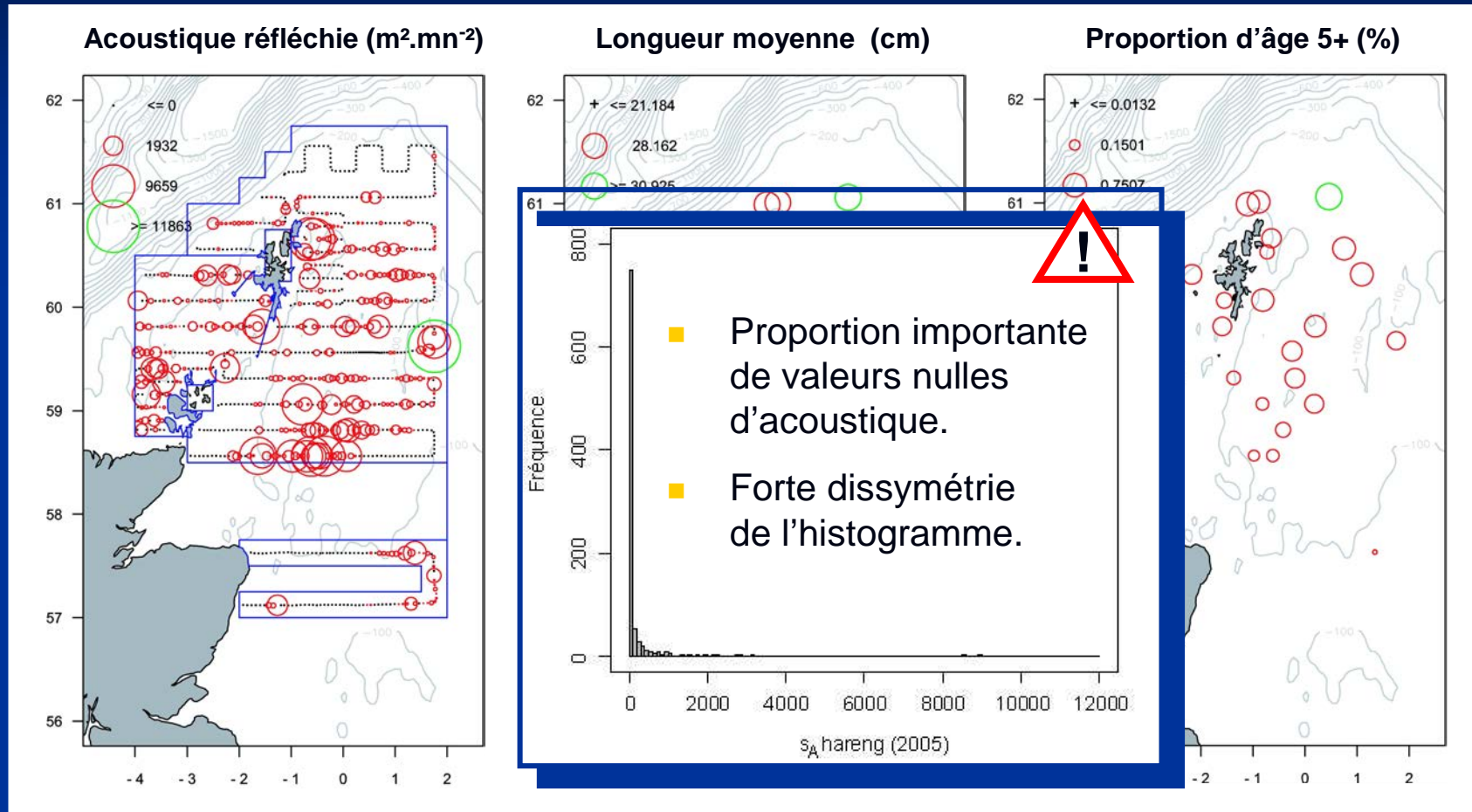
- Anchois du golfe de Gascogne :
 - Situation décrite par les MAFs : changement soudain dans les années récentes (2000-2001) vers un déclin.
 - L'année 2005 (l'année prévue) pas en continuité
 - Changement observé comparé à la prévision concernant l'évolution de l'état du stock. Dans quel sens ? Vers une aggravation au vu des indices de 2005.



Plan

- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Données collectées



Année 2005

Combinaison nécessaire pour estimer des abondances

Modèle géostatistique multivarié spécifique

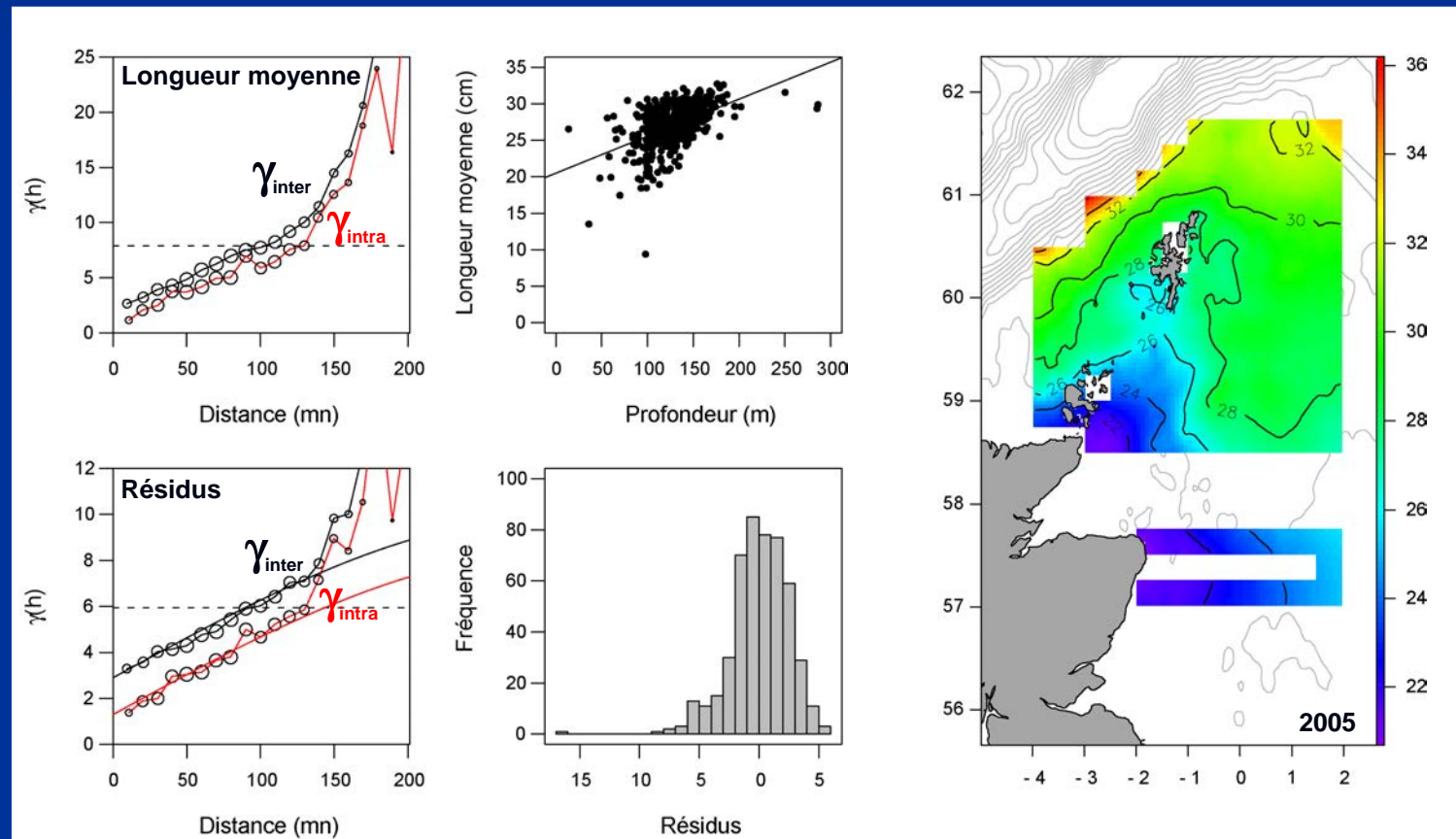
- Modèle identifié dès les années 90 sur les campagnes 1989-1994 (P. Fernandes, Ph. Guiblin, J. Rivoirard, J. Simmonds).
- Étendu sur la période 1989-2005.

Objectifs : Décrire les relations structurales des différentes variables.

- Modélisation de la longueur moyenne.
- Modélisation des proportions à l'âge.

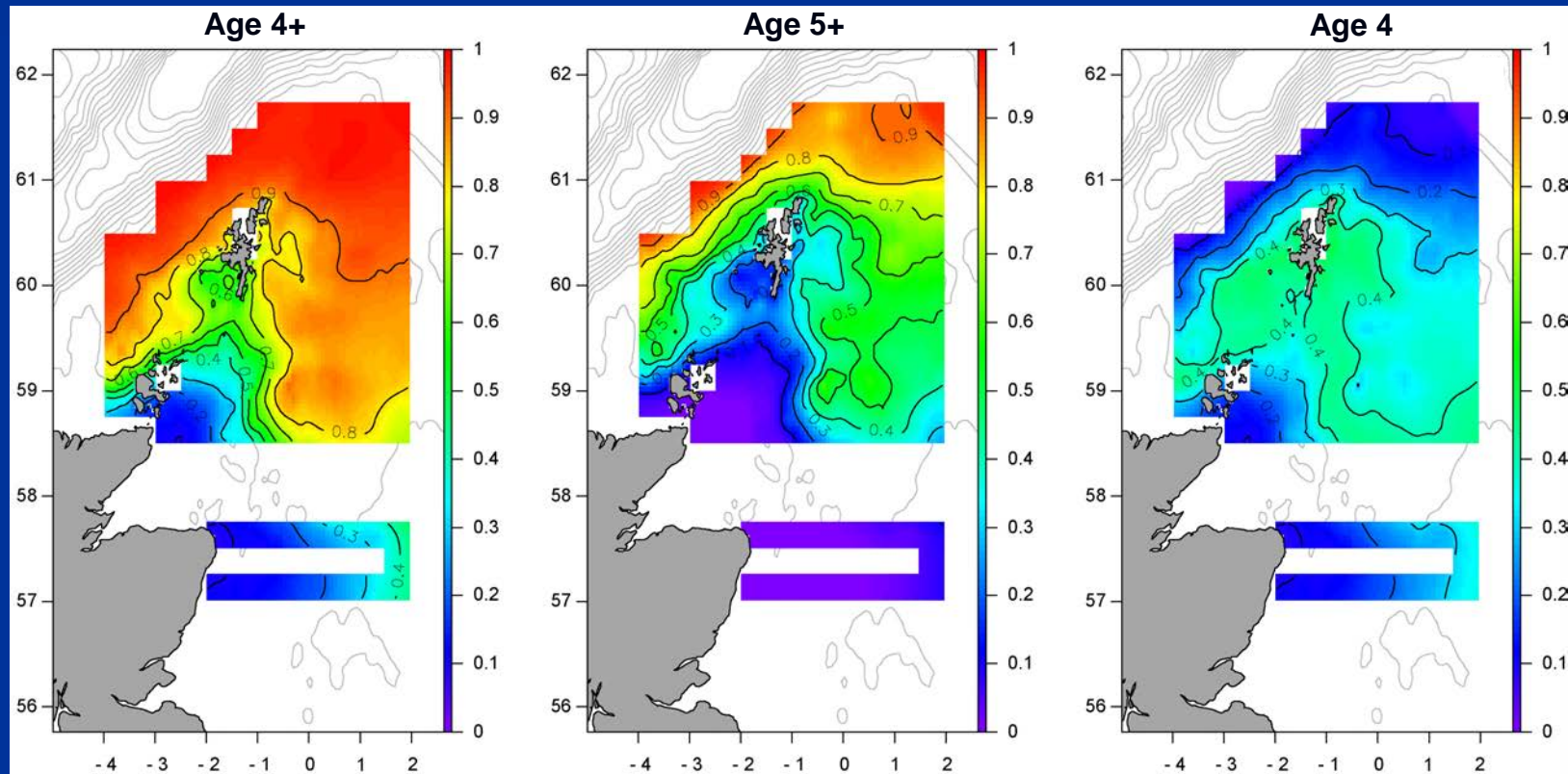
Modélisation de la longueur moyenne

- En utilisant les données de longueur des autres années disponibles dans un modèle adéquat à 3 dimensions (spatiale et temporelle).
- Par krigeage avec la profondeur comme dérive externe.



Modélisation des proportions d'âge

- Par krigeage avec la régression logistique sur la longueur moyenne comme dérive externe et une structure résiduelle sans discontinuité spatiale.
- Cartes de proportion d'âge obtenues en faisant la différence entre 2 cartes consécutives de proportions cumulées d'âge.



Année 2005

Simulation conditionnelle (2)

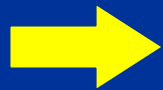
Soit $Z(x)$ la variable régionalisée à simuler.

valeur vraie = estimateur de krigeage + erreur de krigeage

$$Z(x) = Z^*(x) + \underbrace{[Z(x) - Z^*(x)]}$$

INCONNU

Cependant, dans le cas Gaussien, cette erreur est spatialement indépendante du processus de krigeage.



Simulation non conditionnelle

$$Z_{SC}(x) = Z^*(x) + [Z_{SNC}(x) - Z_{SNC}^*(x)]$$

Applicable sans transformation à la longueur moyenne
et aux proportions cumulées d'âge.



Mais pas suffisant pour la variable acoustique.

Simulation de la variable acoustique

- La variable acoustique $Z(x)$ fait référence à un modèle gaussien anamorphosé :

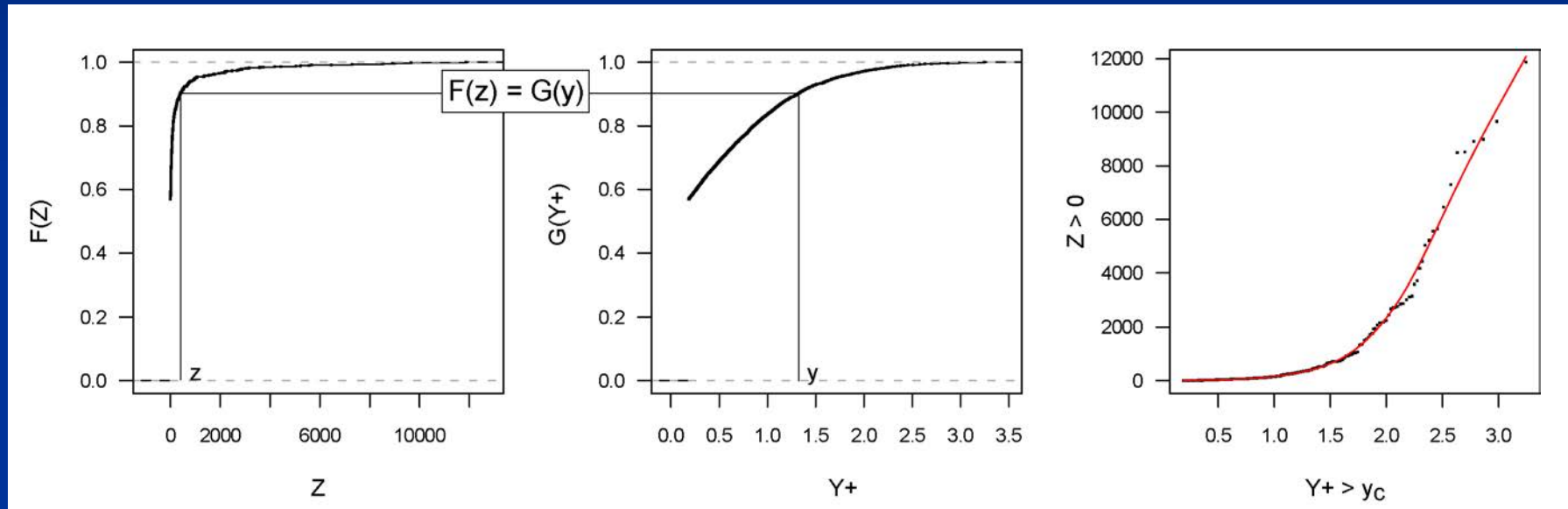
$$Z(x) = \varphi(Y(x))$$

- Le problème des zéros : N'importe quelle valeur de Y inférieur à y_c (déterminé par la proportion de zéros p_0) peut être associée à une valeur nulle de Z :

$$p_0 = P(Z = 0) = P(Y < y_c) = G(y_c)$$

- Adaptation nécessaire :
 - L'échantillonneur de Gibbs \Rightarrow simuler itérativement les valeurs de Y aux points de données où Z est zéro, conditionnellement au fait qu'elles soient inférieures à y_c et aux autres points de données.
 - Le variogramme de Y sera déterminé indirectement à partir de la gaussienne tronquée à gauche Y^+ , vu que le variogramme de n'importe quelle transformation de la variable gaussienne est une fonction du variogramme de Y .

Anamorphose gaussienne



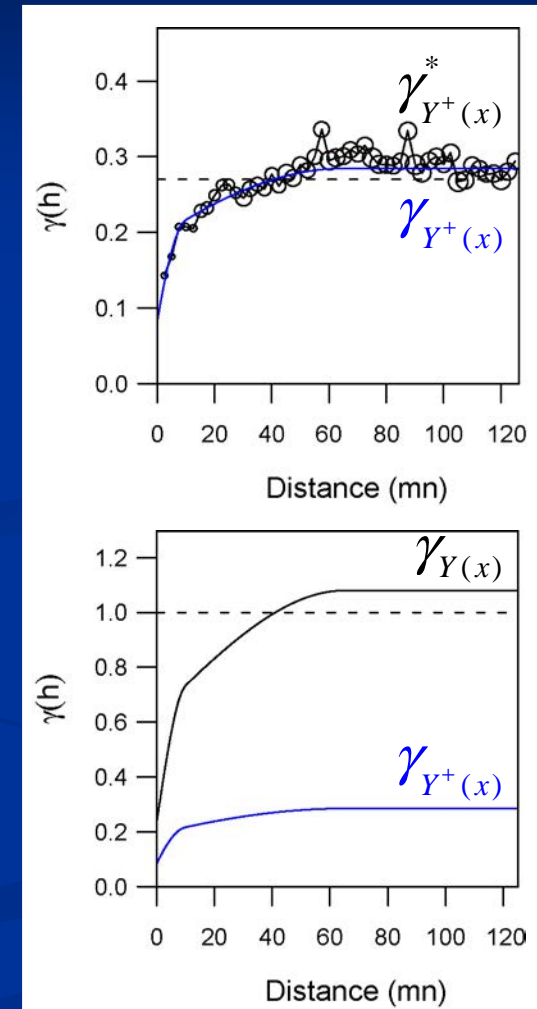
Elle associe à chaque valeur de z un y tel que les deux courbes de fréquences cumulées ont la même valeur.

Modélisation de la gaussienne

- Gaussienne tronquée à gauche Y^+ égale à :
 - y_c si $Z(x) = 0$ tel que $y_c = G^{-1}(p_0)$
 - $Y(x)$ si $Z(x) > 0$

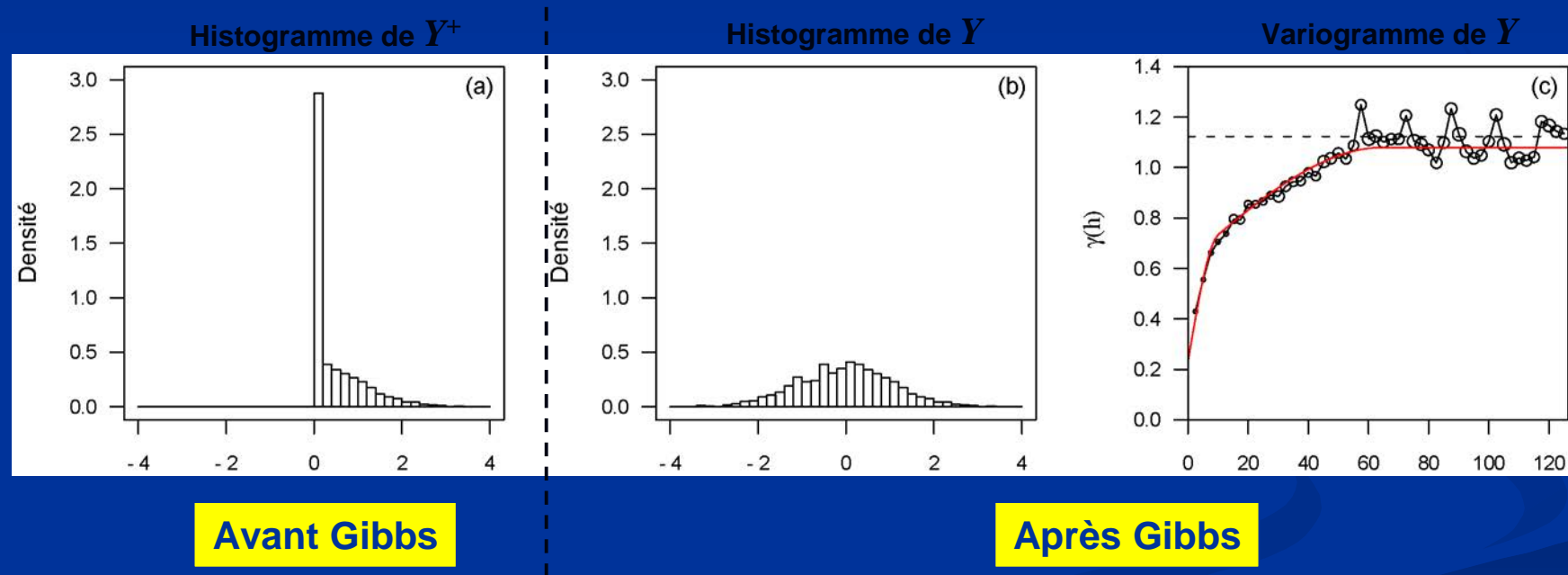
- N'importe quelle fonction $f(Y)$ d'une variable gaussienne peut être développée en polynôme d'Hermite, où sa covariance spatiale est :
 - $$\text{Cov}[f(Y(x+h)), f(Y(x))] = \sum_1 (f_n)^2 [\rho(h)]^n$$

- Modèle de la gaussienne inféré pour 2005 :
 - $$\gamma(h) = 0,23.\text{pep}(0) + 0,40.\text{sph}(10) + 0,45.\text{sph}(65)$$



Echantillonneur de Gibbs

- Simuler itérativement les valeurs gaussiennes correspondant aux valeurs zéros selon le modèle de Y .



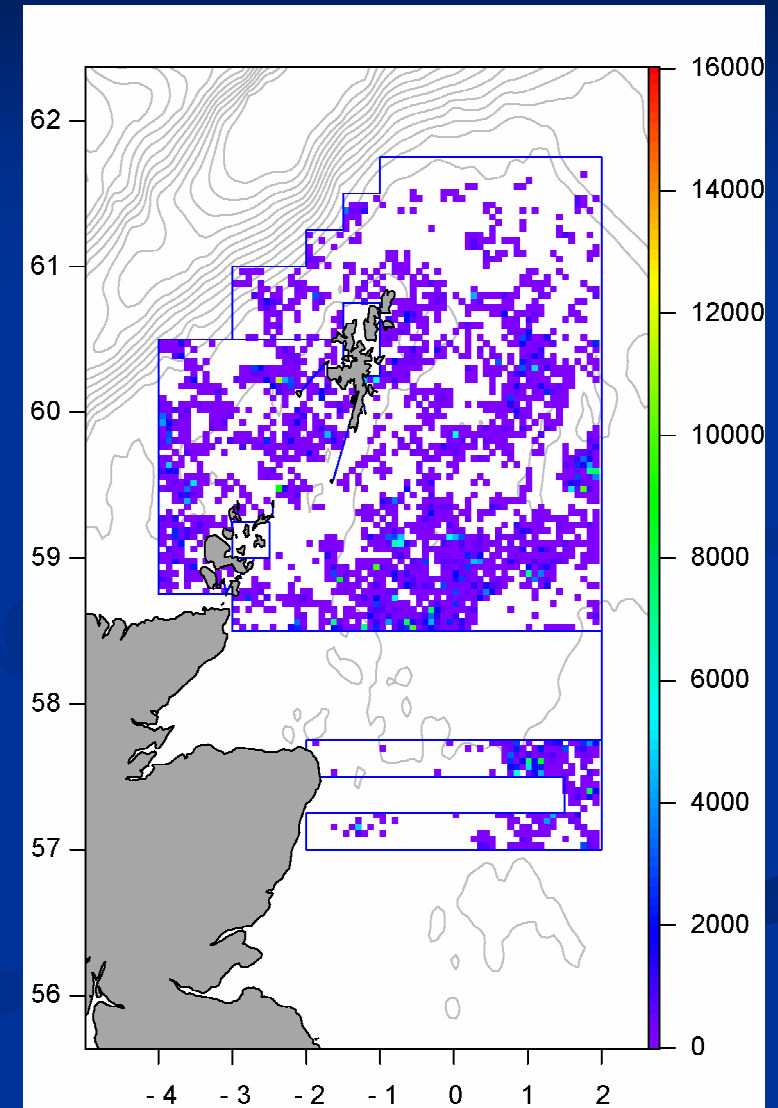
Simulation de l'acoustique

- Réalisation de la variable gaussienne produite par simulation conditionnelle :

$$Y_{sc} = Y^* + (Y_{snc} - Y_{snc}^*)$$

- Réalisation de la variable gaussienne transformée en réalisation brute d'acoustique :

$$Z_{sc} = \Phi(Y_{sc})$$

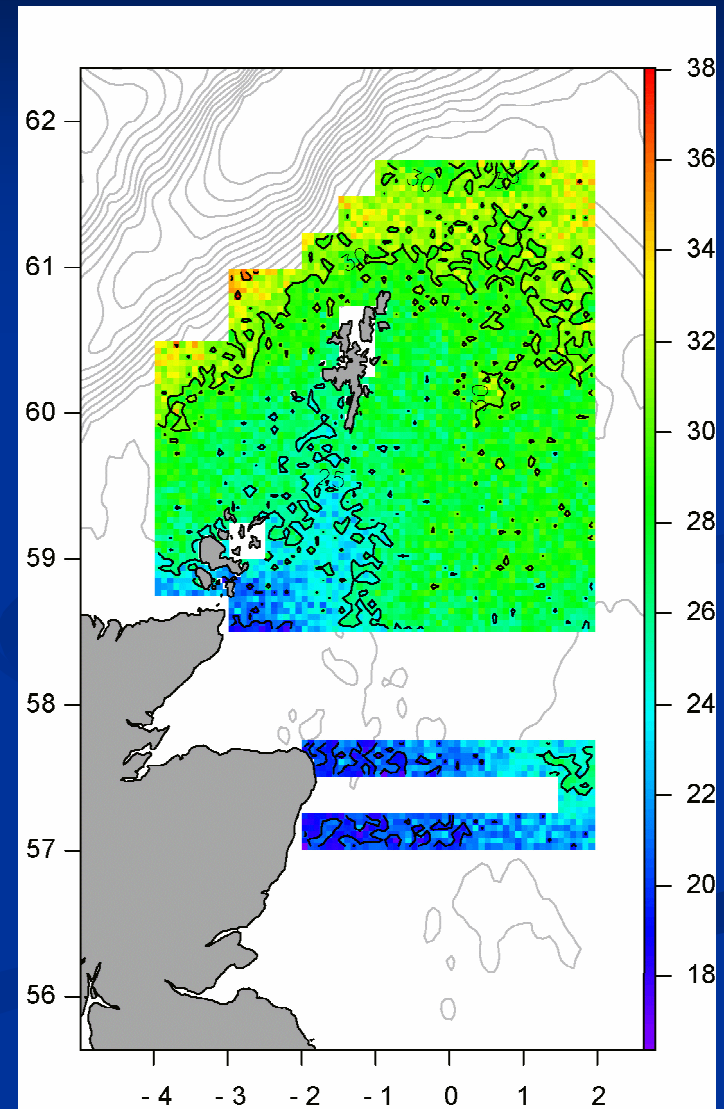


Année 2005

Simulation de la longueur moyenne

- Réalisation de la longueur moyenne produite par simulation conditionnelle et utilisant le modèle géostatistique qui a été déterminé précédemment:

$$L_{sc} = L^* + (R_{snc} - R_{snc}^*)$$



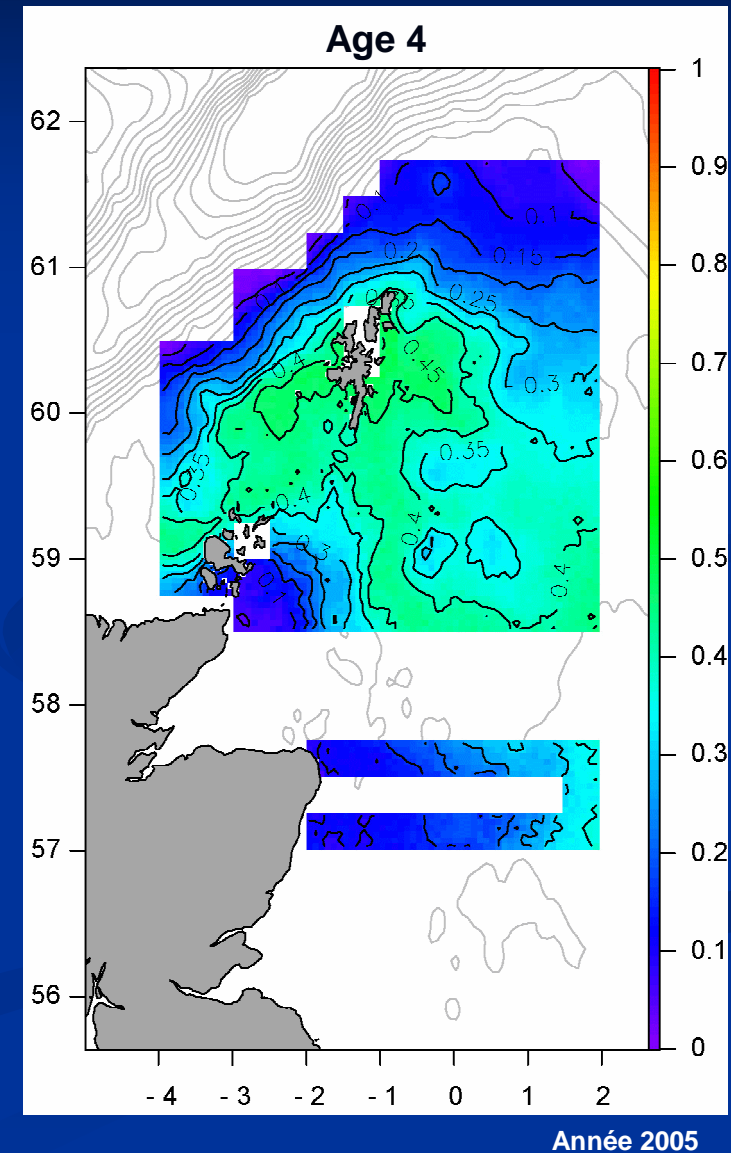
Année 2005

Simulation des proportions d'âge

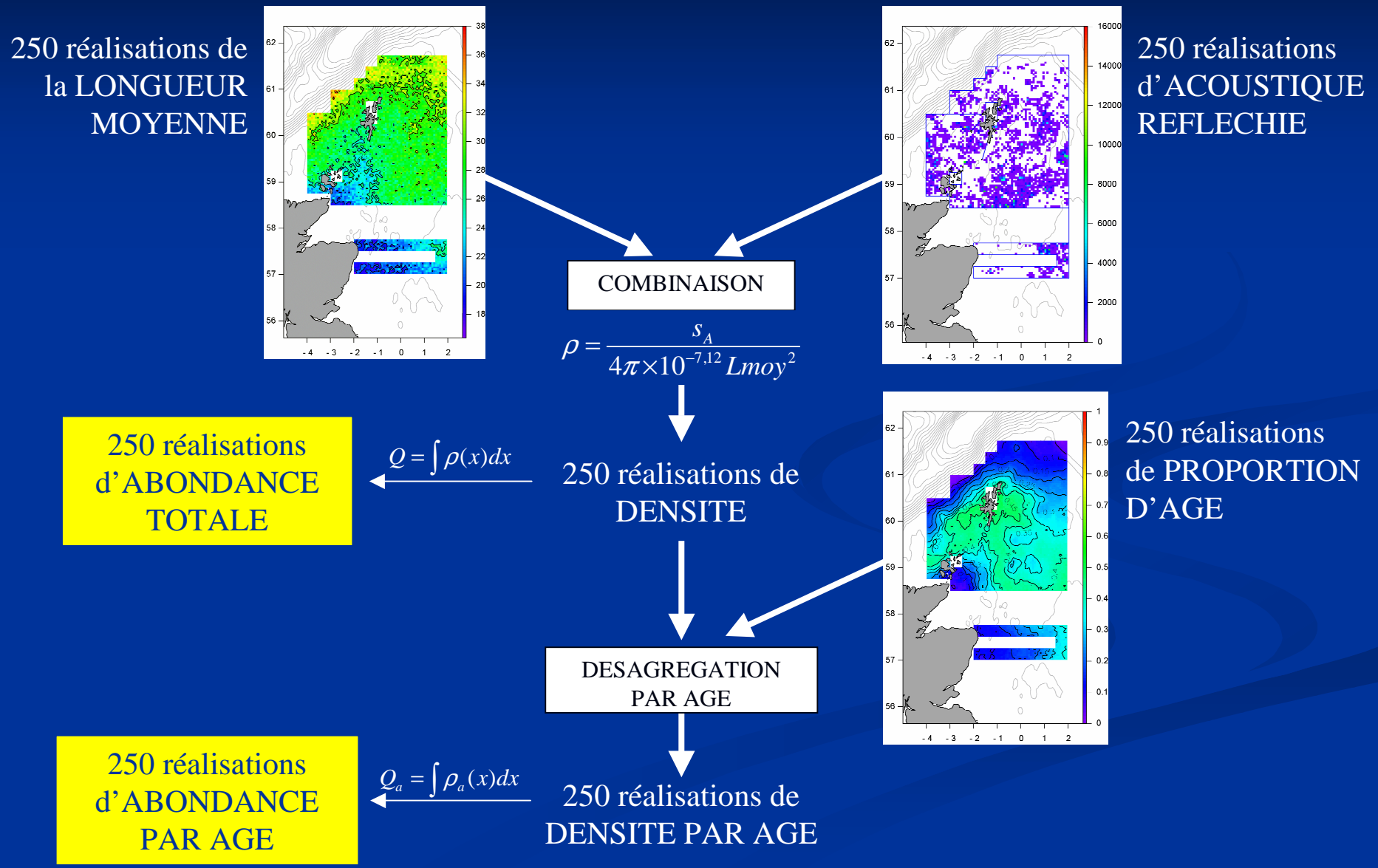
- Réalisation d'une proportion cumulée d'âge produite par simulation conditionnelle et utilisant le modèle géostatistique qui a été déterminé précédemment:

$$Ajp_{sc} = Ajp^* + sj(R_{snc} - R_{snc}^*)$$

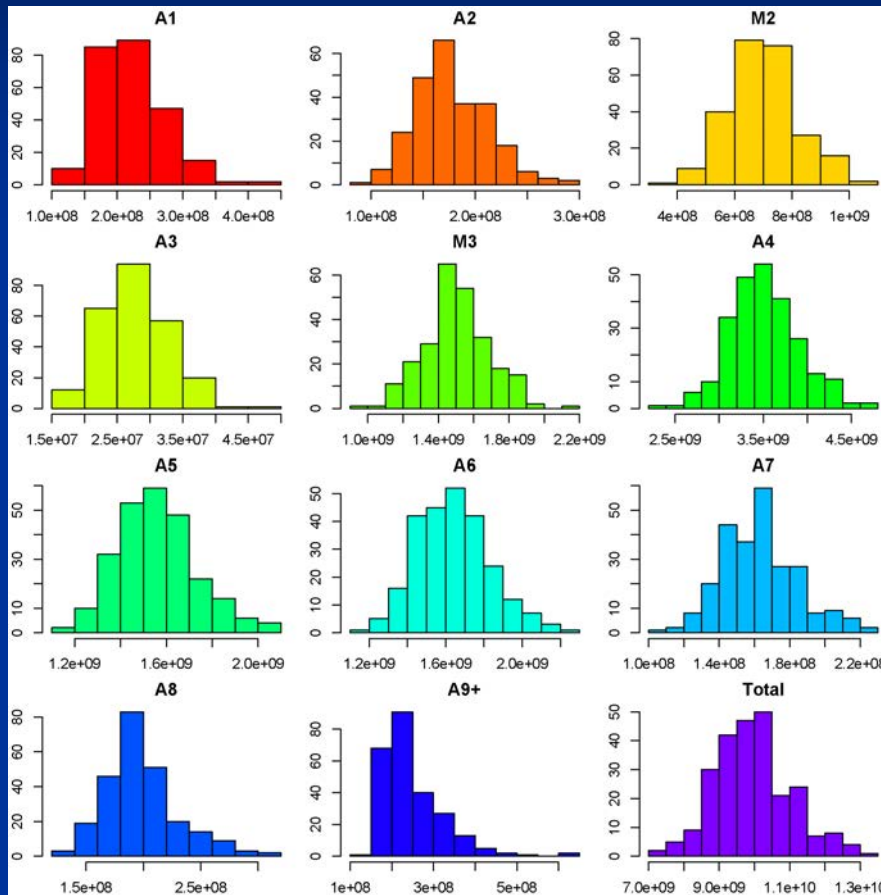
- Réalisation d'une proportion d'âge obtenue par différence entre deux réalisations de proportions cumulées d'âge consécutive.



Combinaison des incertitudes et désagrégation par âge



Estimation globale, variance d'estimation et structure de distribution des réalisations



Ages du hareng	Campagne acoustique 2005		
	Moyenne ¹	Ecart-type ¹	CV (%)
A1	224	52	23
A2	178	38	20
M2	608	117	17
A3	26	6	25
M3	1504	184	12
A4	3516	303	10
A5	1560	175	11
A6	1642	109	12
A7	160	22	13
A8	198	32	16
A9+	248	75	30
Total	9957	1073	11

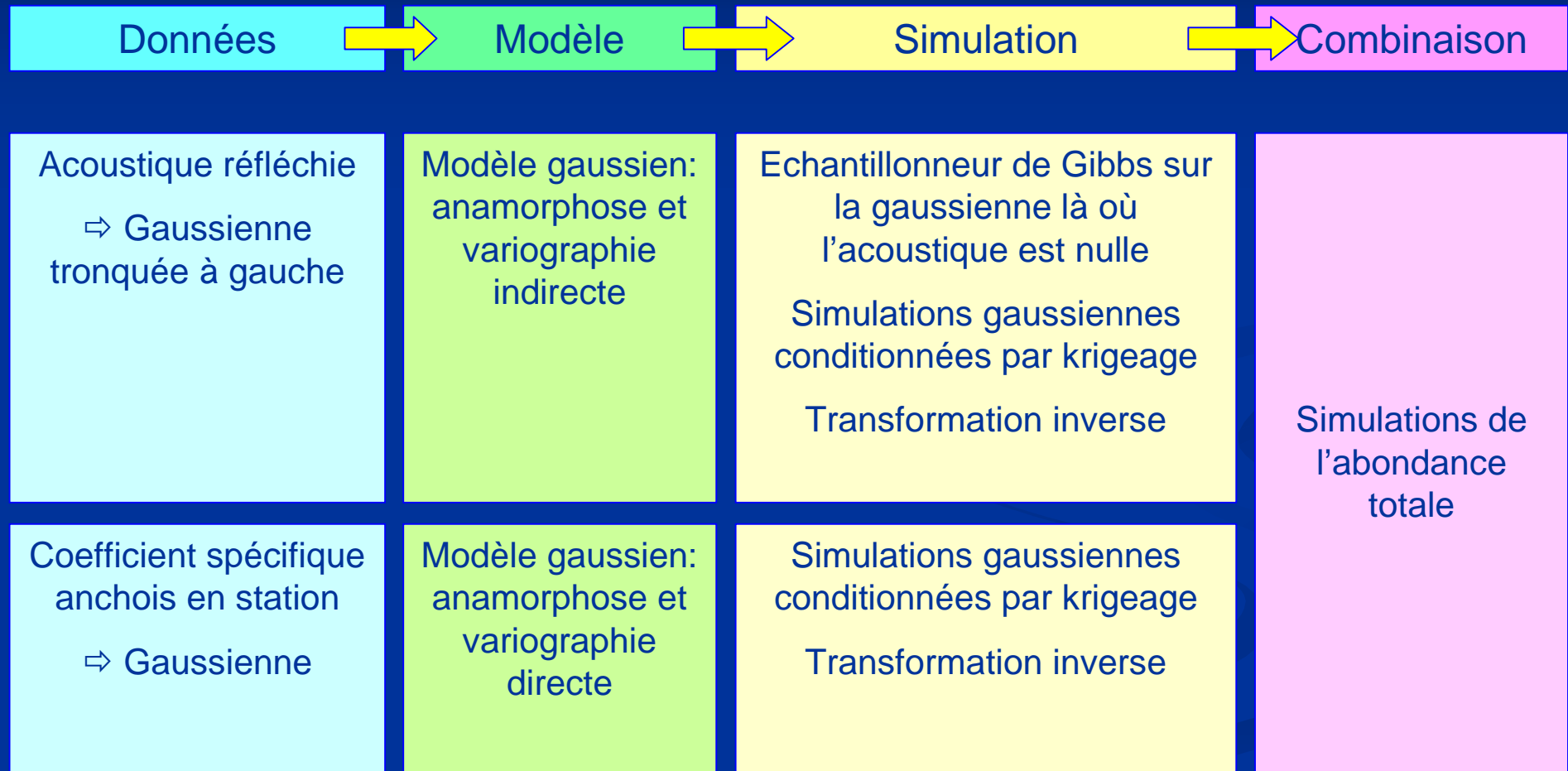
¹ en million d'individus.

Propagation des incertitudes envisageable

Plan

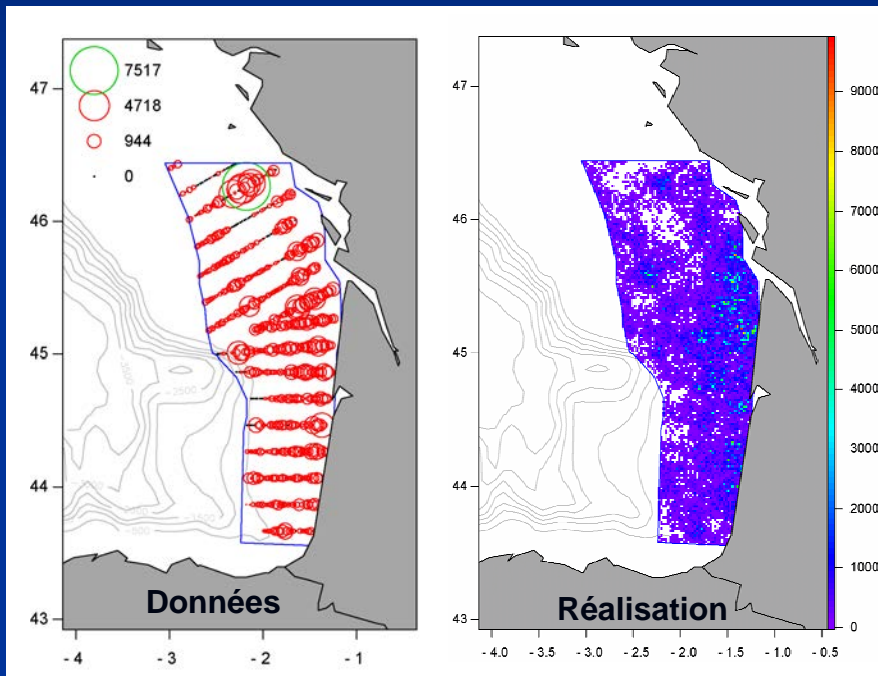
- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Méthodologie transférable



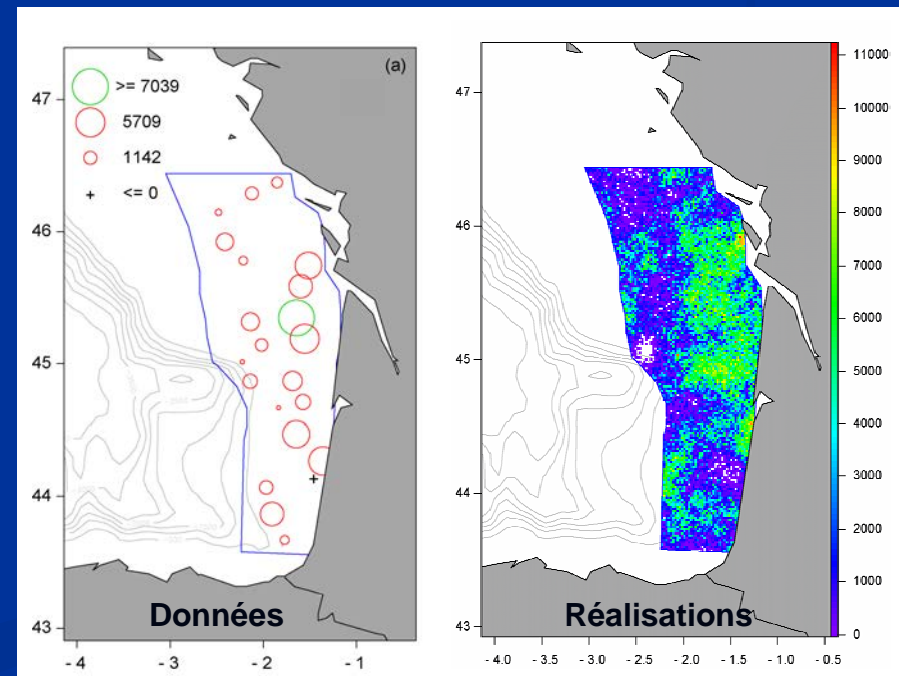
Simulations conditionnelles

Acoustique réfléchi



Année 2000

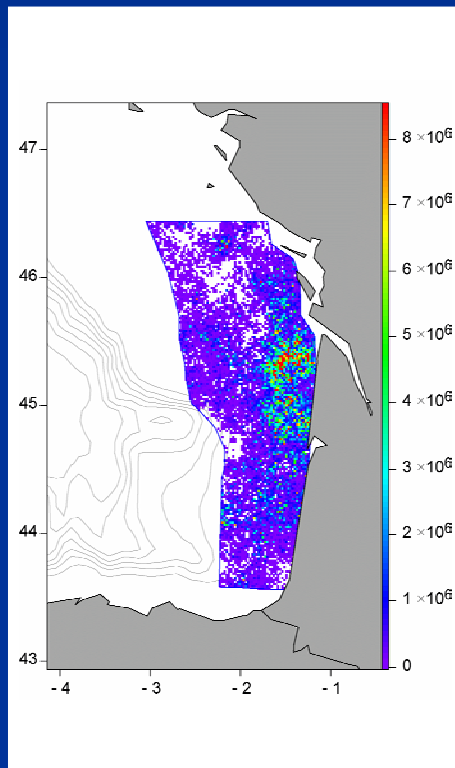
Coefficient spécifique anchois



Année 2000

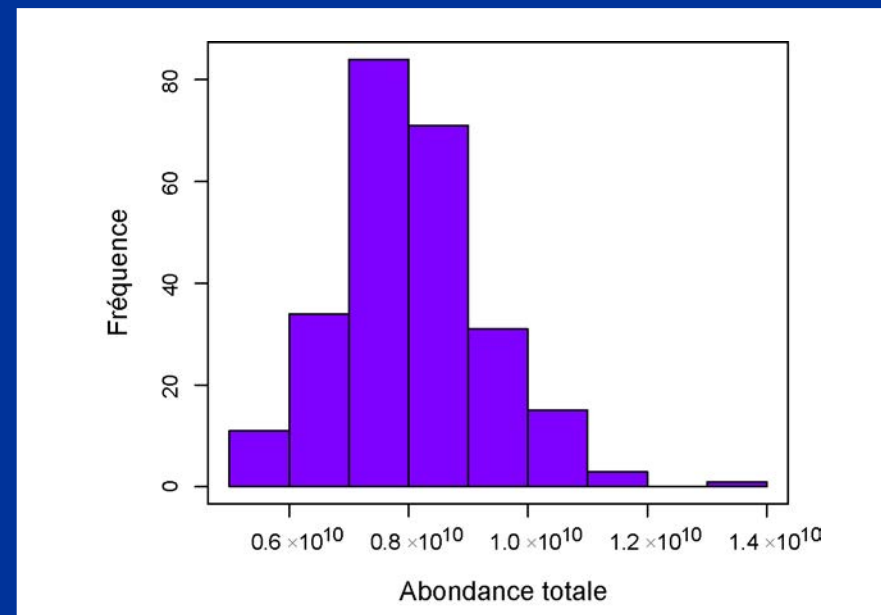
Incertitude sur l'estimation globale

Combinaison



Anchois	Campagne acoustique 2000		
	Moyenne ¹	Ecart-type ¹	CV (%)
Total	8 045	1 278	16

¹ en million d'individus.



Plan

- Introduction
- De l'étude de la dynamique spatiale de populations halieutiques à leur diagnostic
 - Caractériser la distribution spatiale
 - Application au stock de merlu du golfe de Gascogne
 - Relation démographie et dynamique spatiale
 - Des indicateurs au diagnostic
- Evaluation de l'incertitude associée aux estimations d'abondance des campagnes acoustiques
 - Le hareng autour des îles Shetland
 - L'anchois du golfe de Gascogne
- Conclusion

Le travail réalisé (1)

- Partie I :
 - Des outils simples insensibles aux valeurs nulles de densité de poissons, ne nécessitant pas la délimitation arbitraire d'un champ et permettant de quantifier les distributions spatiales de populations halieutiques.
 - Appliqués à la population de merlu du golfe de Gascogne, ces indices ont permis de décrire la dynamique spatiale des différents groupes d'âge.
 - Ce sont des indicateurs de la dynamique de la population.
 - La méthode des MAFs permet de combiner différents indices de campagnes en extrayant les composantes les plus continues dans le temps afin de diagnostiquer l'état d'un stock.

Le travail réalisé (2)

- Partie II :
 - L'évaluation de la part d'incertitude provenant de l'échantillonnage spatial dans l'estimation d'abondances à partir de campagnes acoustiques est désormais possible.
 - Nécessité de combiner les différentes sources d'incertitudes.
 - Recours à des simulations géostatistiques conditionnelles de modèles multivariés spécifiques.
 - La méthode a été appliquée au hareng autour des îles Shetland et à l'anchois dans le golfe de Gascogne.

Perspectives

- La précision des indices spatiaux pourrait être évaluée par le biais des simulations conditionnelles.
- Etendre l'approche par indicateurs à une palette plus large d'indices dans le but d'assurer un diagnostic écosystémique.
- Extension à d'autres campagnes acoustiques.
- Evaluer de nouvelles incertitudes ?
- Apport de la géostatistique dans l'analyse statistique de séries temporelles ?

A large school of silver fish, possibly sardines, swimming in deep blue water. The fish are densely packed and moving in a coordinated pattern. A dark blue horizontal bar is overlaid on the center of the image, containing the text "Merci de votre attention" in yellow.

Merci de votre attention

