

ÉCOLE CENTRALE MARSEILLE
ED PHYSIQUE ET SCIENCES DE LA MATIÈRE 352
INSTITUT FRESNEL

THÈSE

présentée en première version en vue d'obtenir le grade de
Docteur, spécialité « Optique, photonique et traitement d'image »

par

Sylvain Jay

ESTIMATION ET DÉTECTION EN IMAGERIE HYPERSPECTRALE : APPLICATION AUX ENVIRONNEMENTS CÔTIERS

Thèse dirigée par M^{me} MIREILLE GUILLAUME et soutenue le
31 Octobre 2012 devant le jury composé de :

M. YANNICK DEVILLE	Université Paul Sabatier (Toulouse)	(Rapporteur)
M. GRÉGOIRE MERCIER	Institut Mines-Telecom / Telecom Bretagne (Brest)	(Rapporteur)
M. JACQUES BLANC-TALON	DGA/DS/QIS (Bagneux)	(Examinateur)
M. JOCELYN CHANUSSOT	Grenoble INP	(Examinateur)
M ^{me} AUDREY MINGHELLI-ROMAN	Université du Sud Toulon Var (La Valette du Var)	(Examinatrice)
M ^{me} MIREILLE GUILLAUME	Ecole Centrale Marseille	(Examinatrice)

A Jean-Claude, mon "grand-père adoptif"...

REMERCIEMENTS

Une thèse, c'est un peu comme un marathon : c'est long, difficile, et souvent, les derniers kilomètres sont les plus durs. Par conséquent, je suis redevable de chaque personne que j'ai cotoyée au cours de ces trois dernières années, et qui m'a permis de près ou de loin, de surmonter toutes les difficultés auxquelles j'ai eu à faire face, tant sur le plan professionnel que sur le plan privé. La liste est longue, mais je vais essayer d'être le plus exhaustif possible, afin de ne négliger personne.

En tout premier lieu évidemment, je tiens à remercier chaleureusement ma directrice de thèse, Mireille Guillaume. Je pense qu'il eût été difficile de trouver meilleure personne pour ce rôle : pour moi, l'aspect scientifique est indissociable de l'aspect humain, et de ces deux points de vue-là, j'ai pris beaucoup de plaisir à échanger lors des nombreuses discussions que nous avons eues au cours de ces trois ans. J'ai ainsi pu travailler dans des conditions très agréables et, à n'en pas douter, cela a beaucoup contribué à la réussite de cette thèse.

Dans un deuxième temps, je voudrais remercier la Direction Générale de l'Armement (DGA) qui nous a accordé sa confiance en acceptant de financer ce projet de thèse. En particulier, je remercie chaleureusement mes deux responsables successifs à la DGA, Jacques Blanc-Talon, chargé de mission Sécurité-Interopérabilité et Systèmes d'Information, et Véronique Serfaty, responsable du Domaine Scientifique à la Mission pour la Recherche et l'Innovation Scientifique (DGA/MRIS). Ils ont suivi régulièrement mes travaux tout au long de cette thèse, et y ont apporté un regard extérieur très intéressant et instructif.

Merci également aux Professeurs Yannick Deville (IRAP, Université Paul Sabatier, Toulouse) et Grégoire Mercier (Lab-STICC, Institut Mines-Telecom / Telecom Bretagne, Brest) pour avoir rapporté ma thèse, ainsi que Jacques Blanc-Talon (DGA/DS/QIS, Bagneux) et Audrey Minghelli-Roman (LSEET, ISITV, La Valette du Var), pour l'avoir examinée. Merci enfin au Professeur Jocelyn Chanussot (GIPSA-Lab, Grenoble INP, Grenoble) pour avoir également examiné ma thèse, et présidé mon jury.

Concernant le personnel de l'Institut Fresnel, je voudrais d'abord remercier Salah Bourennane et Amélie Litman, qui m'ont accueilli au sein de leur équipe, "Groupe Signaux Multidimensionnel" (GSM), et "Hyperfrequency, Instrumentation, Processing, Experimentation" (HIPE) respectivement, et qui m'ont permis d'évoluer dans un contexte scientifique enrichissant. Merci également à tout le personnel permanent travaillant dans ces deux équipes. Je remercie notamment Stéphane

Derrode, Caroline Fossati, Ali Khalighi, et Julien Marot pour GSM, ainsi que Pierre Sabouroux et Redha Abdeddaim pour HIPE. Un merci particulier à Mouloud Adel, qui a toujours été de bon conseil à chaque fois que j'ai eu besoin de lui. Je remercie également nos collègues de l'équipe φ -TI, et en particulier Nicolas Bertaux, Philippe Réfrégier, Muriel Roche, et Antoine Roueff qui m'ont transmis leur intérêt pour le traitement du signal et des images notamment à travers les cours dispensés lors de mon cursus à l'École Centrale Marseille. Merci enfin à Laurent Gallais pour nos sessions d'escalade du midi, qui auront fait tant de bien avant de retourner travailler !

J'ai aussi eu la chance de croiser de nombreux doctorants à l'Institut Fresnel tout au long de ces trois années. De ces rencontres sont nées des amitiés, ou tout simplement de bons moments passés autour d'un repas au "RU" ou sur un des paliers du laboratoire. Dans l'ordre chronologique, merci donc à Alexis Huck, Damien Letexier, Riad Khelifi, Xuefeng Liu, Alexis Cailly, Zhang Yi, Olivier Casadessus, Chadi Gabriel, Nabil Boughnim, Ivan Voznyuk, et Elodie Georget. Un très grand merci également à un ami sincère et fidèle, en la personne d'Aurélien Arnaubec, qui a toujours été d'un grand soutien, et avec lequel nous avons partagé et nous partagerons, je l'espère, beaucoup de bons moments.

Enfin, mes derniers remerciements concernant l'Institut Fresnel, vont à l'ensemble du personnel administratif. Si le laboratoire fonctionne aussi bien, c'est en grande partie grâce à eux, et c'est toujours avec un immense plaisir que je me serais rendu à chaque fois dans ce couloir. Je remercie donc Josiane Martin, Evelyne Santacroce, Magali Griess, Laetitia Le Guay, Guylène Deguero-Maynard, Frédéric Forestier, Jean Cayzac, Christophe Juillard, et Alexan Hatchikian. Un merci tout particulier à quatre d'entre eux, avec lesquels nous avons partagé beaucoup de rires et de bonne humeur, et qui sont devenus au fil du temps, plus que de simples collègues de travail, mais de vrais amis : Nelly Bardet, Claire Guéné, Nadège Guillem, et Marc Traversari.

D'une manière générale, je tiens à remercier toutes les personnes que j'ai croisées à l'Institut Fresnel, notamment à travers son directeur, Hugues Giovannini d'abord, puis Stefan Enoch. L'ambiance qui règne dans ce laboratoire a vraiment contribué à ce que mon séjour s'y déroule du mieux possible.

Par ailleurs, par l'intermédiaire du projet HypLitt financé par la DGA, j'ai eu l'occasion de travailler avec la société ActiMar, ce qui a ainsi permis de former une passerelle intéressante avec le monde de l'entreprise. Merci donc à Marc Lennon, Guillaume Sicot, et Nathalie Vigneau pour cette collaboration.

J'aimerais également remercier Dominique Bigliazzi, du Département de la Recherche à l'École Centrale Marseille, pour son aide et son extrême gentillesse.

Nombreuses ont donc été les personnes qui m'ont soutenu dans le cadre professionnel. Mais je ne serais rien aujourd'hui si mes proches n'avaient pas été présents depuis le tout début. Famille et amis comptent beaucoup pour moi, et ces quelques pages sont l'occasion de les remercier (de manière un peu moins conventionnelle...) pour tout ce qu'ils m'apportent chaque jour.

Un énorme merci donc aux deux mariés néo-ch'tis Doumé et Doumette Bonato, à mon Hollandaise "avec un grand H" préférée Laurentine, à mon éternel pote de pêche et bien plus encore, Ludo, ainsi qu'Heidi, au nordiste Paulo (parce que oui, La Rochelle, c'est le Noooooord!), et à mon partenaire de foot-ski-muscu-squash-natation-grimpe-rando-et-j'en-passe Mpapa. Merci également à tous mes potes bouillarguais, manduellois, riansais, et setec-iens (merci Mat' pour le coup!), à ceux de la prépa, et ceux de l'Ecole Centrale, notamment tous les footeux avec lesquels j'ai pu partagé de supers moments riches en émotions. Merci enfin à Magali, pour m'avoir soutenu une bonne partie de cette thèse et de ma vie, et à Florent, mon fillot préféré et p'tit frère adoptif.

Je voudrais également remercier sincèrement toute la famille Valdaperez, et en particulier Jean-Claude et Marie-Noël qui, depuis l'âge de trois mois, ont toujours été présents pour moi, et pour lesquels je voue une reconnaissance sans limite. Je vous dois en grande partie la réussite de cette thèse...

Pour finir, mes derniers remerciements, et non les moindres, vont à toute ma famille, et plus particulièrement à mes parents et ma sœur préférée. Vous avez toujours été là pour moi, et avez largement contribué à ce que je suis devenu aujourd'hui. En cela, je suis très heureux et fier des valeurs que vous avez su me transmettre. MERCI pour tout...

TABLE DES MATIÈRES

TABLE DES MATIÈRES	viii
NOTATIONS	xi
ACRONYMES	xiii
LISTE DES FIGURES	xiv
1 INTRODUCTION	1
1.1 CONTEXTE	1
1.2 LIENS AVEC LE PROJET REI HyPLITT	3
1.3 PRÉSENTATION DES TRAVAUX	5
2 MODÈLES DE TRANSFERT RADIATIF DANS LA COLONNE D'EAU	7
2.1 INTRODUCTION SUR LES GRANDEURS CONSIDÉRÉES	7
2.2 MODÈLE GÉNÉRAL	8
2.2.1 Réflectance en profondeur infinie	8
2.2.2 Réflectance en profondeur finie	11
2.3 MODÉLISATION DES PARAMÈTRES PHYSIQUES	15
2.3.1 Absorption	16
2.3.2 Rétro-diffusion	19
2.3.3 Variabilité du coefficient d'atténuation diffuse ($k_d(\lambda)+k_u(\lambda)$)	21
2.4 CONCLUSIONS	22
3 ÉTAT DE L'ART SUR L'ESTIMATION DES PARAMÈTRES DE LA COLONNE D'EAU	25
3.1 INTRODUCTION	25
3.2 MÉTHODES D'ESTIMATION D'UNE PARTIE DES PARAMÈTRES . .	25
3.2.1 Méthodes d'estimation de la bathymétrie	25
3.2.2 Méthodes d'estimation de la qualité de l'eau	26
3.3 MÉTHODES D'ESTIMATION GLOBALES	27
3.3.1 Inversion par optimisation d'une fonction d'erreur	27
3.3.2 Inversion par Look-Up Table	29
3.4 MÉTHODE UTILISÉE POUR LA COMPARAISON	30
4 ESTIMATION DE PARAMÈTRES DE L'EAU ET BATHYMÉTRIE : MÉTHODES STATISTIQUES	31
4.1 INTRODUCTION	31
4.2 PRÉSENTATION DE LA MÉTHODE DU MAXIMUM DE VRAISEMBLANCE	31
4.2.1 Introduction	31
4.2.2 Méthode d'estimation	32

4.2.3	Algorithme d'optimisation choisi	36
4.2.4	Présentation des données de test	38
4.2.5	Résultats	40
4.3	MODÉLISATION D'UNE PROFONDEUR VARIABLE	44
4.3.1	Introduction	44
4.3.2	Méthode d'estimation à une pente	45
4.3.3	Méthode d'estimation à deux pentes	50
4.3.4	En résumé	54
4.4	ETUDE DU MAXIMUM A POSTERIORI (MAP)	55
4.4.1	Introduction et méthode	55
4.4.2	Résultats	57
4.5	ETUDE DES BORNES DE CRAMÉR-RAO	62
4.5.1	Introduction et méthode	62
4.5.2	Résultats	64
4.6	CONCLUSIONS	67
5	MÉTHODES DE DÉTECTION DE CIBLES IMMERGÉES	69
5.1	INTRODUCTION	69
5.2	QUELQUES FILTRES CLASSIQUES ADAPTÉS À LA DÉTECTION DE CIBLES IMMERGÉES	71
5.3	LE GBF (GLRT-BASED BATHYMETRIC FILTER), UNE SOLUTION LORSQUE LES PARAMÈTRES DE LA COLONNE D'EAU SONT INCONNUS	72
5.4	RÉSULTATS	73
5.5	CONCLUSIONS	77
6	CONCLUSIONS ET PERSPECTIVES	79
6.1	CONCLUSIONS	79
6.2	PERSPECTIVES	83
A	ANNEXES	85
A.1	UN NOUVEAU FILTRE BATHYMÉTRIQUE BASÉ SUR LE GLRT (GBF)	87
A.2	MÉTHODE D'ESTIMATION LORSQUE LE FOND EST INCONNU	89
A.3	CALCUL DES BORNES DE CRAMÉR-RAO	92
A.3.1	Introduction	92
A.3.2	Cas scalaire	93
A.3.3	Cas multivarié	95
	LISTE DES PRODUCTIONS SCIENTIFIQUES	97
	BIBLIOGRAPHIE	99

NOTATIONS

SCALAIRES, VECTEURS, ET MATRICES

\mathbb{R}	Ensemble des réels
$x \in \mathbb{R}, X \in \mathbb{R}$	Variable scalaire
$x \in \mathbb{R}^L$	Vecteur de taille L
$X \in \mathbb{R}^{L \times M}$	Matrice de taille (L, M)

VECTEURS ET MATRICES SPÉCIALES

$\text{diag}[f(l)]_{l \in \llbracket 1; L \rrbracket}$	Matrice diagonale de taille L dont l'élément diagonal à la ligne l est $f(l)$
Id_L	Matrice identité de taille L
θ	Vecteur de paramètres à estimer

OPÉRATIONS SUR LES VECTEURS ET MATRICES

$ \cdot $	Déterminant
$\text{tr}(\cdot)$	Trace
$(\cdot)^t$	Transposition

FONCTIONS

$\frac{\partial}{\partial \theta_i}$	Dérivée partielle par rapport au paramètre θ_i
$\mathbb{E}[\cdot]$	Espérance mathématique

GRANDEURS PHYSIQUES

a	Coefficient d'absorption total
a_{CDOM}	Coefficient d'absorption du CDOM
a_{PHY}	Coefficient d'absorption du phytoplancton
a_w	Coefficient d'absorption de l'eau pure
b_b	Coefficient de rétro-diffusion total
$b_{b,PHY}$	Coefficient de rétro-diffusion du phytoplancton
$b_{b,NAP}$	Coefficient de rétro-diffusion des particules non algales
$b_{b,w}$	Coefficient de rétro-diffusion de l'eau pure
C_{CDOM}	Concentration en CDOM
C_{NAP}	Concentration en particules non algales
C_{PHY}	Concentration en phytoplancton
H	Profondeur

k	Coefficient d'atténuation (modèle M1)
k_d	Coefficient d'atténuation dans le sens descendant (modèle M2)
k_u^b, k_u^c	Coefficients d'atténuation dans le sens ascendant (modèle M2)
r	Réflectance en radiance
r_B	Réflectance en radiance du fond
r_∞	Réflectance en radiance d'une colonne d'eau infinie
R	Réflectance en irradiance
θ_s	Angle zénithal solaire
θ_v	Angle de visée

DIVERS

$\text{dist}_x(A, B)$	Distance entre les abscisses des points A et B
$[\mathbf{X}]_{ij}$	Élément de la matrice \mathbf{X} situé à la i^{me} ligne et j^{me} colonne
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$	Loi normale de vecteur moyen $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Gamma}$
$\mathbf{x} \sim \mathcal{F}$	\mathbf{x} suit la loi \mathcal{F}
$\hat{\mathbf{x}}, \hat{\mathbf{X}}$	Estimation d'un vecteur ou d'une matrice
$\underset{H_0}{\geq} \underset{H_1}{\leq}$	Test d'hypothèses

ACRONYMES

ACE	Estimateur angulaire adapté (Adaptive Cosine/coherence Estimator)
AMF	Filtre adapté adaptatif (Adaptive Matched Filter)
ATCOR	Modèle de correction atmosphérique et topographique (Airborne Atmospheric and Topographic Correction Model)
BACE	Estimateur angulaire adapté bathymétrique (Bathymetric Adaptive Cosine/coherence Estimator)
BAMF	Filtre adapté adaptatif bathymétrique (Bathymetric Adaptive Matched Filter)
BMF	Filtre adapté bathymétrique (Bathymetric Matched Filter)
BCR	Borne de Cramér-Rao
CDOM	Matière organique dissoute colorée (Colored Dissolved Organic Matter)
EMV	Estimateur (ou Estimation) du Maximum de Vraisemblance
GBF	Filtre bathymétrique issu du GLRT estimant tous les paramètres inconnus (GLRT-based Bathymetric Filter)
GLRT	Test du rapport de vraisemblance généralisé (Generalized Likelihood Ratio Test)
LEVACE	Méthode de détection alliant estimation des paramètres inconnus par LEVMAR et filtrage par ACE
LEVAMF	Méthode de détection alliant estimation des paramètres inconnus par LEVMAR et filtrage par AMF
LEVMAR	Méthode d'inversion des paramètres utilisant l'algorithme de Levenberg-Marquardt
MAP	Maximum <i>A Posteriori</i>
MF	Filtre adapté (Matched Filter)
MV	Maximum de Vraisemblance
NAP	Particules non-algales (Non algal Particles)
REQM	Racine de l'Erreur Quadratique Moyenne
ROC	Caractéristique de performance du détecteur (Receiver Operating Characteristic)
SNR	Rapport Signal à Bruit (Signal-to-Noise Ratio)

LISTE DES FIGURES

1.1	Cube hyperspectral.	2
1.2	Carte bathymétrique estimée par ActiMar à partir des données HypLitt.	3
1.3	Evolution du spectre du sable mouillé en fonction de la nature de la couche d'eau qui le recouvre.	4
1.4	Localisation des trois zones imagées S_1 , S_2 , et S_3	5
1.5	Exemples d'images HypLitt.	6
2.1	Schéma représentant les interactions entre la lumière incidente et l'atmosphère, la colonne d'eau et le fond.	11
2.2	Différents spectres d'absorption de l'eau pure.	16
2.3	Spectres d'absorption de l'eau pure salée, du phytoplancton, du CDOM, et des particules non algales.	19
2.4	Spectres de rétro-diffusion de l'eau pure salée, du phytoplancton, et des particules non algales.	20
2.5	Evolution du coefficient d'atténuation diffuse $k_d(\lambda) + k_u^b(\lambda)$ en fonction des concentrations.	22
3.1	Evolution de la réflectance entre H et $H + \Delta H$	29
4.1	Fonction de log-vraisemblance calculée sur une image HypLitt.	34
4.2	Influence du bruit sur la forme de la log-vraisemblance (données synthétiques).	35
4.3	Log-vraisemblance avec θ paramétré en fonction de t	37
4.4	Schéma d'optimisation à une dimension utilisé dans la méthode de relaxation.	38
4.5	Images en pseudo-vraies couleurs créées à partir de données hyperspectrales simulées, et spectres de réflectance correspondants.	39
4.6	Comparaison entre les spectres subsurfaces réels et le spectre simulé à partir des paramètres estimés et du spectre du fond.	42
4.7	Comparaison entre les spectres subsurfaces réels et le spectre simulé à partir des paramètres estimés avec ou sans estimation de f et σ^2	43
4.8	Estimation de la pente sur une image HypLitt : exemple 1.	47
4.9	Estimation de la pente sur une image HypLitt : exemple 2.	47
4.10	Estimation de la pente sur une image HypLitt : exemple 3.	48
4.11	Estimation de la pente sur une image HypLitt : exemple 4.	49

4.12	Erreur sur l'estimation des concentrations en fonction de la pente.	49
4.13	Variations discontinues ou continues de la hauteur.	50
4.14	Choix de la bande permettant d'obtenir la dynamique $r_{MAX}(\lambda) - r_{MIN}(\lambda)$ maximale.	50
4.15	Evolution de la dynamique $r_{MAX}(\lambda) - r_{MIN}(\lambda)$ en fonction de la longueur d'onde.	51
4.16	Méthode de localisation de la droite de rupture.	51
4.17	Représentation de la contrainte $H_1^- = H_1^+$ et $H_2^- = H_2^+$	52
4.18	Estimation d'une pente linéaire par morceaux : exemple sur simulations.	52
4.19	Estimation d'une pente linéaire par morceaux : premier exemple sur données HypLitt.	53
4.20	Estimation d'une pente linéaire par morceaux : deuxième exemple sur données HypLitt.	54
4.21	Effet de la régularisation sur la vraisemblance.	57
4.22	Erreurs sur l'estimation des paramètres, obtenues avec ou sans régularisation, en fonction de la profondeur (données simulées).	58
4.23	Erreurs sur l'estimation des paramètres, obtenues avec ou sans régularisation, en fonction du SNR (données simulées).	58
4.24	Localisation de la zone de mesure de la concentration en phytoplancton et de la turbidité.	59
4.25	Erreurs sur l'estimation de H , obtenues avec ou sans régularisation, en fonction de la profondeur (données HypLitt).	60
4.26	Evolution du biais expérimental en fonction de la taille de l'échantillon.	64
4.27	Comparaison entre les BCRs et les variances expérimentales des estimateurs en fonction de la taille de l'échantillon.	65
4.28	Evolution des BCRs en fonction des valeurs des paramètres.	66
5.1	Courbes ROC illustrant la détection du métal galvanisé.	73
5.2	Probabilité de détection du métal galvanisé en fonction du SNR.	74
5.3	Exemple d'image HypLitt utilisée pour l'évaluation des détecteurs bathymétriques.	75
5.4	Courbes ROC illustrant la détection de bâches blanches et noires sur les images HypLitt (modèle de Brando et al. (2009) et Bricaud et al. (1995)).	76
5.5	Courbes ROC illustrant la détection de bâches blanches et noires sur les images HypLitt (modèle de Lee (1994)).	77
A.1	Evolution en fonction de l'échantillon de la différence entre la profondeur estimée sous les hypothèses H_0 ou H_1 , et la profondeur estimée uniquement sur l'ensemble d'apprentissage.	87

INTRODUCTION



1.1 CONTEXTE

L'IMAGERIE hyperspectrale est actuellement en plein essor, tant au niveau de son champ d'application que des développements algorithmiques associés. Elle est en effet utilisée dans des domaines aussi variés que la défense, le domaine médical ou la surveillance de l'environnement. Cette popularité provient certainement de la richesse de l'information contenue dans une seule image, offrant ainsi de nombreuses possibilités de traitement.

Une image hyperspectrale peut être considérée comme l'extension d'une image couleur classique. Cette dernière caractérise la réponse spectrale de la scène étudiée (ou réflectance), à seulement trois longueurs d'onde (le rouge pour $\lambda = 700$ nm, le vert pour $\lambda = 546$ nm, et le bleu pour $\lambda = 436$ nm), contre une ou plusieurs centaines pour une image hyperspectrale. Sur la Figure 1.1, on représente ce type de données sous la forme d'un cube hyperspectral, avec deux dimensions spatiales et une dimension spectrale. On peut voir que chaque pixel est un vecteur de dimension L qui, une fois corrigé de l'illumination incidente, caractérise la réponse spectrale du ou des matériaux qui le constituent : il est ainsi possible de retrouver à distance sa composition.

De nombreuses applications sont développées, de la détection supervisée ou non supervisée, à la classification en zones spectralement homogènes, en passant par des problématiques d'estimation de paramètres caractéristiques de la scène imagée.

De nos jours, l'imagerie hyperspectrale est également beaucoup utilisée comme outil de cartographie en milieu aquatique. A partir de telles données, on peut retrouver la composition des fonds marins (Louchard et al. 2003, Mobley et al. 2005, Fearn et al. 2011) ou établir des cartes de bathymétrie en zones côtières (Adler-Golden et al. 2005, Brando et al. 2009, Heege et al. 2007). Pour illustrer cela, on montre sur la Figure 1.2 un exemple de cette dernière application, réalisé par la société ActiMar à partir des données HypLitt présentées dans la Section 1.2. On peut également identifier les concentrations de divers constituants de l'eau (phytoplankton, matière organique dissoute, etc) (Kutser et al. 2005, Gitelson et al. 2007, Le et al. 2009). L'utilisation de telles images aériennes pouvant couvrir des étendues très importantes permet d'améliorer considérablement

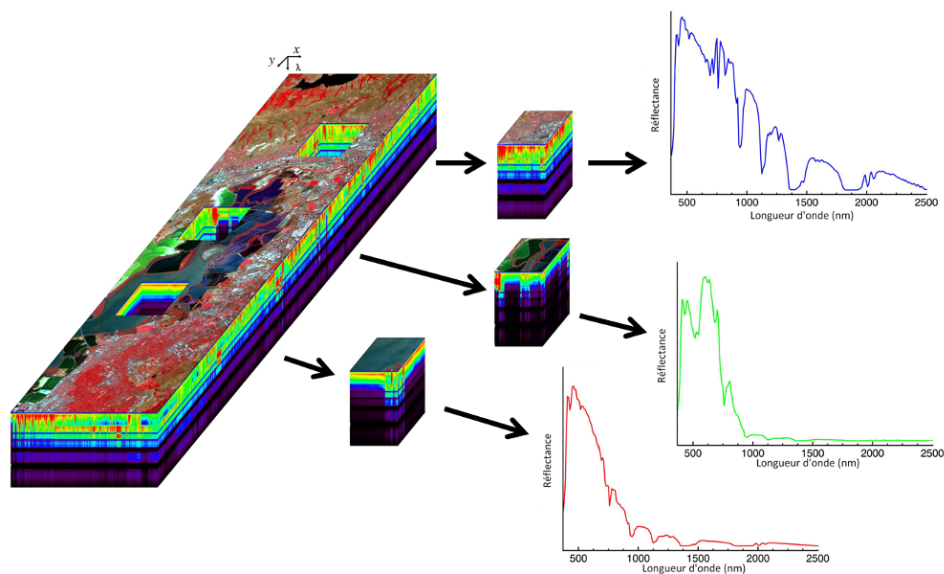


FIGURE 1.1 – Cube hyperspectral (Christophe 2006).

l'efficacité des méthodes traditionnelles de mesure de la composition de l'eau, basées sur des prélèvements effectués sur un faible nombre de sites. La portée de ce type d'imagerie représente donc un avantage très intéressant par rapport à d'autres techniques de cartographie.

Dans cette thèse, nous développons de nouvelles méthodes permettant d'estimer les caractéristiques de la colonne d'eau. Pour cela, nous utilisons un modèle bathymétrique de réflectance, qui tient compte de la profondeur et des concentrations en constituants dits "optiquement actifs" de l'eau. Nous abordons le problème d'un point de vue statistique, en considérant que chaque pixel est une réalisation d'un vecteur aléatoire, paramétré grâce au modèle bathymétrique. Par rapport aux techniques d'inversion classiques, cela représente une approche novatrice dans la résolution de ce type de problème. De plus, la précision obtenue et les temps de calcul sont très intéressants et figurent parmi les atouts majeurs de ces méthodes.

Par ailleurs, si plusieurs techniques d'estimation des paramètres caractéristiques de la colonne d'eau ont déjà été développées dans la littérature, peu d'auteurs se sont penchés sur la problématique de la détection de cibles immergées. En effet, l'implémentation des filtres traditionnellement utilisés en milieu aérien n'est pas satisfaisante dès lors qu'elle est faite en milieu aquatique. Par définition, les détecteurs supervisés nécessitent de connaître le spectre de la cible que l'on recherche afin de pouvoir le comparer au pixel test ; or, placée sous une certaine couche d'eau, la cible peut potentiellement avoir une réflectance très différente du spectre observé à l'air libre. On peut observer cela sur la Figure 1.3, où l'on représente l'évolution de cette déformation en fonction de la profondeur et de la turbidité de l'eau. Pour conserver de bonnes performances, il est nécessaire de corriger l'effet de la colonne d'eau avant le processus de détection, et dans nos travaux, le modèle bathymétrique est donc inséré dans le développement des filtres. Les détecteurs ainsi obtenus s'avèrent très intéressants vis-à-

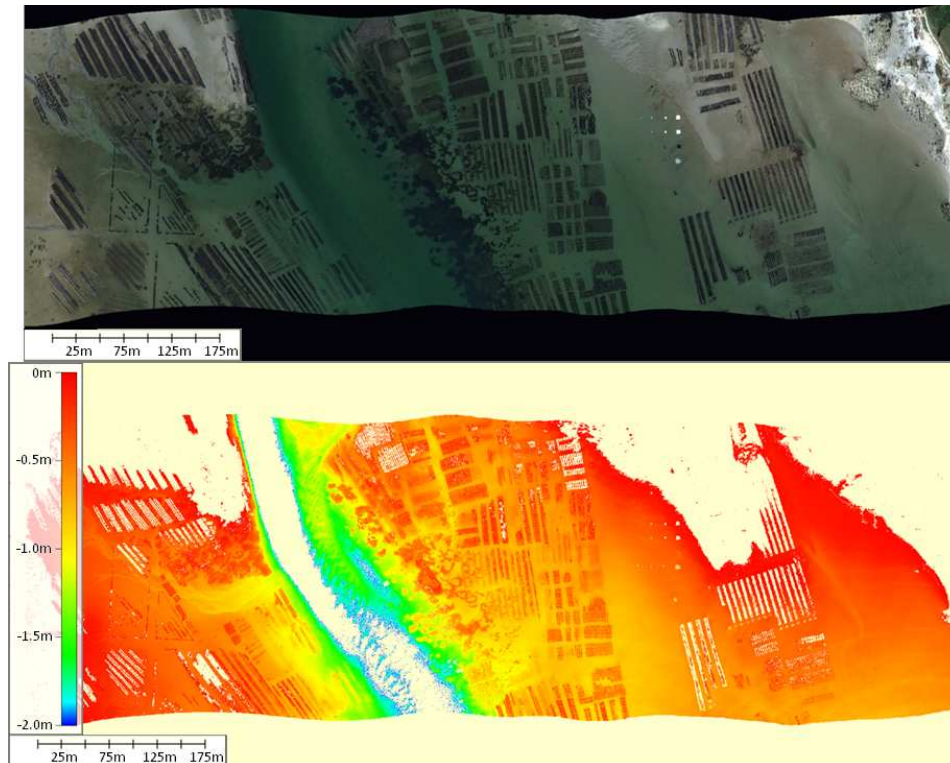


FIGURE 1.2 – Carte de profondeur estimée (en bas) réalisée par la société ActiMar, à partir des données HypLitt (en haut).

vis des méthodes consistant à supprimer l'effet de la colonne d'eau avec le modèle de réflectance et les paramètres estimés grâce aux méthodes usuelles, et appliquer les algorithmes de détection classiques.

1.2 LIENS AVEC LE PROJET REI HYP LITT

A partir de la deuxième année, cette thèse a été menée en parallèle avec le projet HypLitt (Evaluation des capacités de la télédétection HYPerspectrale et développement de méthodes innovantes de traitement d'images pour des applications Défense en zone LITTorale), qui est un contrat REI (Recherche Exploratoire et Innovation) financé par la Mission pour la Recherche et l'Innovation Scientifique (MRIS) de la Direction Générale de l'Armement (DGA). Il regroupe la société ActiMar et l'Institut Fresnel. Une campagne de télédétection hyperspectrale comprenant l'acquisition de données aéroportées à l'aide d'un capteur imageur VNIR Hypspx et de données de spectroradiométrie de terrain associées, a été réalisée par ActiMar en septembre 2010. Nos algorithmes ayant été testés jusqu'alors sur des images synthétiques, ce nouvel ensemble de données a été d'un apport considérable sur l'évaluation et la valorisation de nos méthodes. La campagne a été menée en Bretagne, à proximité de la presqu'île de Quiberon : trois sites S_1 , S_2 , et S_3 ont été définis afin d'imager la gamme bathymétrique la plus large possible (voir Figure 1.4). Les images ont été prises à différentes altitudes et suivant plusieurs caps. La campagne ayant été faite sur trois jours, nous avons pu avoir accès à différents instants du cycle de marée (et donc différentes profondeurs pour le même

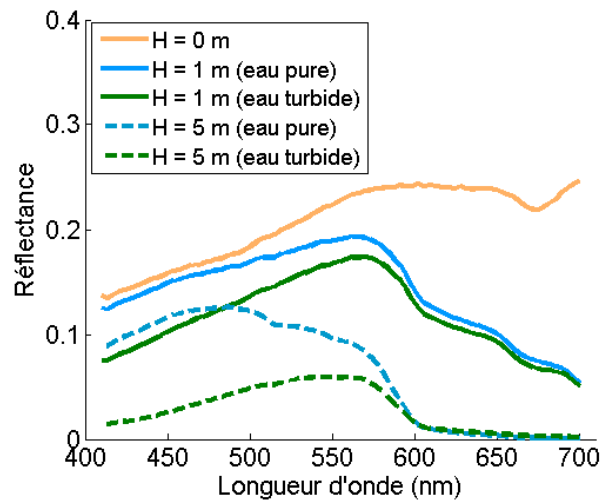


FIGURE 1.3 – Evolution du spectre du sable mouillé en fonction de la nature de la couche d'eau qui le recouvre.

Longueurs d'onde (nm)	Nombre de bandes spectrales	Résolution spectrale FWHM (nm)	Résolution spatiale (m)
400-1000	160	4.5	0.4, 0.5, 1, 2
400-1000	80	7.4	0.5

TABLE 1.1 – Caractéristiques des données HypLitt.

site) et des qualités d'eau sensiblement distinctes. Les caractéristiques des données sont présentées dans le Tableau 1.1. Sur ces images, comme sur les simulations, nous avons restreint l'intervalle de longueurs d'onde à 400 nm-700 nm. Ce choix est dicté par le fait que le rayonnement proche infra-rouge (au-delà de 700 nm) pénètre très peu dans l'eau : on peut d'ailleurs voir sur la Figure 1.3 que la lumière rouge est déjà considérablement atténuée. Pour cette raison, les études menées dans ce domaine ne prennent en compte que les longueurs d'onde du visible. Par ailleurs, nous avons choisi de travailler avec les images à 160 bandes ayant la résolution spatiale minimum (0.4 m et 0.5 m). En effet, la plupart des méthodes d'estimation développées (en dehors des méthodes d'estimation de pente) requièrent de posséder un échantillon d'apprentissage où les données sont identiquement distribuées, et ne sont donc pas adaptées à des résolutions spatiales trop importantes (pour lesquelles les paramètres caractéristiques de la colonne d'eau varieraient trop).

Afin de tester les méthodes de détection, une série de bâches a été installée sur chaque zone. Dans cette thèse, nous utilisons les images du site S_3 pour évaluer les performances des filtres, car cela correspond aux profondeurs les plus importantes, et donc aux situations les plus difficiles à gérer. Sur cette zone, cinq jeux de cibles de 1 m^2 chacune (une blanche et une noire) ont été déposés au fond, à des profondeurs allant de 4.70 m à près de 22 m (sites I_1 , à I_5). Par ailleurs, les prélèvements d'eau ont été réalisés uniquement dans la zone S_3 .

En ce qui concerne les prétraitements, les images ont été corrigées géométriquement, projetées et géoréférencées dans le système UTM 30N, grâce

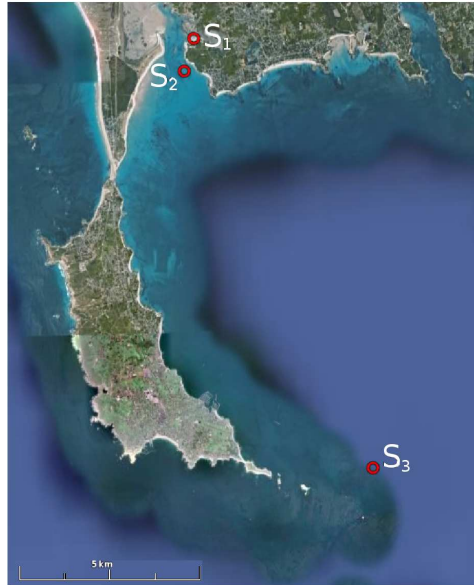


FIGURE 1.4 – Localisation des trois zones imagées S_1 , S_2 , et S_3 .

aux données de trajectographie. La précision géométrique est de l'ordre du pixel. Les données ont ensuite été corrigées atmosphériquement à l'aide du modèle ATCOR, présenté par Richter (2011), et ajustées grâce aux données de spectroradiométrie de terrain. Par ailleurs, les réflexions spéculaires présentes essentiellement sur les images du site S_3 ont été corrigées par l'algorithme de Lyzenga (1978).

Une partie des images sélectionnées sont présentées sur la Figure 1.5.

1.3 PRÉSENTATION DES TRAVAUX

Cette thèse se découpe en six chapitres.

Après cette introduction, nous présentons dans le deuxième chapitre les différents modèles de transfert radiatif dans la colonne d'eau qui ont été utilisés tout au long de ces dernières années. Nous y évoquons deux types de modélisations de la réflectance, selon que l'on se place en profondeur finie ou infinie. Puis nous décrivons les liens avec la nature de la couche d'eau, à travers l'expression des coefficients d'atténuation et de rétro-diffusion.

Dans le troisième chapitre, nous présentons les techniques d'estimation classiques des paramètres du modèle de réflectance. Dans les méthodes globales d'inversion, deux approches sont utilisées : dans la première, une fonction d'erreur entre la mesure et le modèle est minimisée en fonction des paramètres à estimer. Dans la deuxième, une bibliothèque de spectres est constituée avant tout traitement ; puis l'observation est comparée à chaque élément de cette librairie afin de trouver l'ensemble de paramètres permettant de reproduire le plus fidèlement possible le spectre mesuré.

Dans le quatrième chapitre, nous présentons les méthodes statistiques d'estimation étudiées au cours de cette thèse. Nous développons la méthode d'estimation du maximum de vraisemblance, en considérant dans un premier temps, que les paramètres inconnus sont constants dans l'image. Ensuite, nous insérons un modèle de pente dans l'expression de

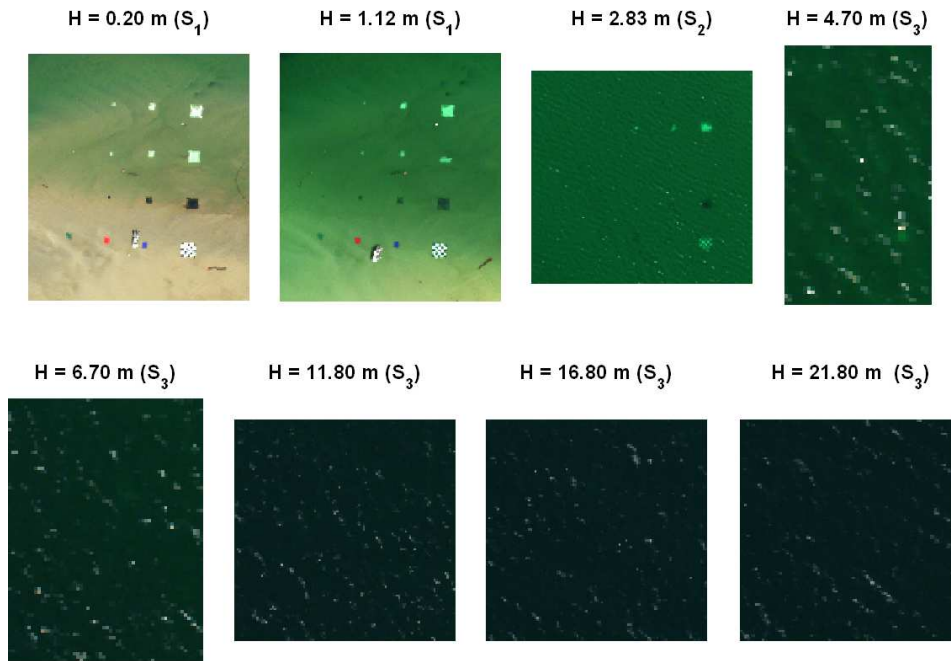


FIGURE 1.5 – Images en pseudo-vraies couleurs créées à partir des données HypLitt.

la vraisemblance, afin de pouvoir gérer des profondeurs variables. Puis, nous étudions l'apport de connaissances *a priori* sur la distribution des paramètres à estimer. Enfin, nous comparons nos résultats aux bornes de Cramér-Rao, outils théoriques nous donnant la précision maximale pouvant être atteinte par toute méthode d'estimation. Toutes ces méthodes sont testées à la fois sur les données synthétiques et sur les données HypLitt.

Dans le cinquième chapitre, nous insérons les modèles bathymétriques dans les détecteurs classiquement utilisés (MF, AMF, ACE). Dans le cas où les paramètres caractéristiques de la colonne d'eau sont inconnus, nous développons le GBF, un filtre issu du test de vraisemblance généralisé, permettant de détecter des cibles immergées même si l'on ne dispose d'aucune connaissance *a priori*. Ces filtres, dits bathymétriques, sont testés sur les données synthétiques et sur les données HypLitt.

Enfin, dans le sixième chapitre, nous concluons sur l'apport de nos méthodes d'estimation statistiques et de détection dans le contexte de l'analyse de scènes côtières.

MODÈLES DE TRANSFERT RADIATIF DANS LA COLONNE D'EAU

2.1 INTRODUCTION SUR LES GRANDEURS CONSIDÉRÉES

UNE image hyperspectrale est une mesure du rayonnement émis ou réfléchi par la scène étudiée dans un grand nombre de bandes spectrales contiguës. La grandeur considérée peut être simplement la radiance, ou la réflectance.

La réflectance en irradiance est donnée par :

$$R(\lambda) = \frac{E_u(\lambda)}{E_d(\lambda)} \quad (2.1)$$

où $E_u(\lambda)$ et $E_d(\lambda)$ sont respectivement les irradiances ascendante et descendante (en W.m^{-2}).

La grandeur mesurée par le capteur hyperspectral est une mesure de radiance, i.e. une mesure de l'irradiance dans une direction donnée. La réflectance en radiance (remote sensing reflectance) est ainsi donnée par :

$$r_{rs}(\lambda) = \frac{L_u(\lambda)}{E_d(\lambda)} \quad (2.2)$$

où $L_u(\lambda)$ est la radiance ascendante (en $\text{W.sr}^{-1}.\text{m}^{-2}$). Le facteur $Q = \frac{E_u(\lambda)}{L_u(\lambda)} = \frac{R(\lambda)}{r_{rs}(\lambda)}$ reliant les deux grandeurs ci-dessus est une mesure caractéristique de l'anisotropie de la radiance ascendante. Pour une distribution isotrope, on a $Q = \pi$ sr. Dans l'eau, Albert et Mobley (2003) évoquent une valeur de 5 sr.

Dans cette thèse, nous nous sommes concentrés sur l'étude des interactions de la lumière avec l'eau, et par conséquent, lors des simulations, nous ne modélisons pas les effets liés à l'atmosphère ou à la traversée de l'interface air/eau. Les réflectances avec lesquelles nous travaillons sont des réflectances dites subsurfaces, c'est-à-dire mesurées juste sous la surface de l'eau.

Sur les images HypLitt, la correction atmosphérique est réalisée avec le modèle ATCOR (Richter 2011).

En ce qui concerne la traversée de la surface, elle peut être représentée

par différentes relations plus ou moins similaires présentées dans la littérature ces dernières années. Certains auteurs lient ainsi la réflectance juste au-dessus de la surface r_{rs}^+ à la réflectance juste en-dessous de la surface r_{rs}^- par une relation linéaire :

$$r_{rs}^+(\lambda) = F \times r_{rs}^-(\lambda). \quad (2.3)$$

Lee et al. (1994) et Giardino et al. (2007) prennent $F = 0.533$, tandis que Kutser (2004) choisit $F = 0.544$.

Lee et al. (1998a) et Lee et Carder (2002) décrivent la traversée de cette interface par le rapport suivant :

$$r_{rs}^+(\lambda) = \frac{0.52 r_{rs}^-(\lambda)}{1 - 1.56 r_{rs}^-(\lambda)}. \quad (2.4)$$

Lee et al. (1999), Cannizzaro et Carder (2006), Gerardino-Neira et al. (2008), Bertels et al. (2008), Xiu et al. (2009), Volpe et al. (2011) et Fearn et al. (2011) utilisent :

$$r_{rs}^+(\lambda) = \frac{0.5 r_{rs}^-(\lambda)}{1 - 1.5 r_{rs}^-(\lambda)}. \quad (2.5)$$

L'application numérique de la formule donnée par Gege (2004) donne (en supposant ici qu'il n'y a pas de réflexions spéculaires) :

$$r_{rs}^+(\lambda) = \frac{0.54 r_{rs}^-(\lambda)}{1 - 2.7 r_{rs}^-(\lambda)}. \quad (2.6)$$

Enfin, Lee et al. (2002) et Lee (2009) utilisent :

$$r_{rs}^+(\lambda) = \frac{0.52 r_{rs}^-(\lambda)}{1 - 1.7 r_{rs}^-(\lambda)}. \quad (2.7)$$

La modélisation de Lee et al. (1999) ayant été très utilisée jusqu'à présent, nous l'avons choisie dans nos travaux.

Dans cette thèse, nous travaillons avec des réflectances en radiance car le capteur hyperspectral mesure une valeur de radiance. Toutefois, il est utile de garder en mémoire la relation $R(\lambda) = Q \times r_{rs}(\lambda)$ évoquée plus haut, pour passer d'une grandeur à l'autre. Par simplicité, on notera par la suite $r(\lambda) = r_{rs}(\lambda)$.

2.2 MODÈLE GÉNÉRAL

Les interactions entre la lumière incidente et la colonne d'eau sont multiples et complexes, comme illustré sur la Figure 2.1. Nous détaillons ici deux types de modélisation de la réflectance, selon que le fond soit ou non pris en compte.

2.2.1 Réflectance en profondeur infinie

Dans le cas de zones dites "optiquement profondes", quand la hauteur d'eau et la turbidité sont telles que le fond n'a quasiment aucune influence sur la réflectance subsurface, on peut représenter cette dernière

par la réflectance d'une colonne d'eau infinie $r_\infty(\lambda)$. Typiquement, cela correspond à ce que l'on peut retrouver en haute mer ou en fleuve. Le signal reçu caractérise alors uniquement la colonne d'eau, et dépend notamment de ses spectres d'absorption $a(\lambda)$ et de rétro-diffusion $b_b(\lambda)$.

Une formulation très utilisée est décrite par Gordon et al. (1975) et reprise notamment par Simis et al. (2005), Morel et al. (2007), et Randolph et al. (2008). Elle suppose que $R_\infty(\lambda)$ est proportionnel au rapport $\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}$:

$$R_\infty(\lambda) = G \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \quad (2.8)$$

où G est un coefficient variable dépendant notamment de la structure du champ de lumière incident. Forget et al. (1997), Cococcioni et al. (2007), Hakvoort et al. (2002) choisissent $G = 0.33$, tandis que Dekker et al. (2001) prennent $G = 0.39 \pm 0.03$. Strömbeck et Pierson (2001) choisissent $G = 0.083$, mais travaillent avec des réflectances en radiance. Vahtmäe et al. (2008) et Kutser (2004) paramétrisent G en fonction du cosinus de l'angle d'incidence de la lumière sous la surface $\mu_d \approx \cos \theta_s$, où θ_s est l'angle zénithal solaire : $G = (-0.629\mu_d + 0.975)$ (à noter qu'une autre relation entre μ_d et θ_s est proposée par McKee et al. (2003)). Pour Giardino et al. (2007), ce coefficient dépend de la longueur d'onde et s'exprime par $G(\lambda) = \frac{1}{1 + \frac{\mu_d(\lambda)}{\mu_u(\lambda)}}$, où $\mu_d(\lambda)$ (respectivement $\mu_u(\lambda)$) sont les cosinus moyens des angles entre la normale à la surface et la lumière descendante (respectivement ascendante).

En pratique, lorsque la turbidité de l'eau est principalement due au phytoplancton (Morel et Prieur (1977) qualifient ces eaux comme étant de type I, contrairement aux eaux de type II, riches en particules non organiques et en matière organique dissoute colorée), le terme d'absorption est très supérieur au terme de rétro-diffusion, et par conséquent, il est courant de faire l'approximation $\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)} \approx \frac{b_b(\lambda)}{a(\lambda)}$. Les paramétrisations de $r_\infty(\lambda)$ présentées plus haut peuvent alors être reprises avec ce nouveau rapport (Morel et Prieur 1977, Kirk 1984, Sathyendranath et al. 1989). D'autres modélisations faisant intervenir cette fraction sont également utilisées par Morel et Gentili (1993), Jerome et al. (1996) ou McKee et al. (2003).

Par ailleurs, de nombreux auteurs considèrent que le facteur G de l'Equation 2.8 est en réalité une fonction des coefficients d'absorption et de rétro-diffusion à travers la fraction $\frac{b_b(\lambda)}{a(\lambda)+b_b(\lambda)}$, afin de tenir compte de ses variations causées par les diffusions multiples (Lee et al. (2004)). Ainsi, grâce à des méthodes de simulation de Monte-Carlo, Gordon et al. (1988) aboutissent à l'expression suivante, reprise plus tard par Keller (2001) :

$$r_\infty(\lambda) = \left(0.0949 + 0.0794 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right) \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}. \quad (2.9)$$

En se basant sur cette dernière paramétrisation, Lee et al. (1998a) proposent une nouvelle expression construite à partir du modèle numérique

de transfert radiatif Hydrolight (Mobley et al. 1993, Mobley 1994), qui est très utilisé dans ce domaine (Albert et Mobley 2003, McKee et al. 2003, Matthews et al. 2010, Giardino et al. 2007, Van Der Woerd et Pasterkamp 2007, Lee et al. 2004) :

$$r_{\infty}(\lambda) = \left(0.070 + 0.155 \left(\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.752} \right) \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}. \quad (2.10)$$

Cette expression est ensuite actualisée par Lee et al. (1999), et demeure à ce jour, au même titre que l'article dans sa globalité, comme l'une des références dans le domaine :

$$r_{\infty}(\lambda) = \left(0.084 + 0.17 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right) \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}. \quad (2.11)$$

Depuis, Lee et al. (2002) et Lee (2009) remplacent le couple de valeurs (0.084;0.17), convenant davantage pour les eaux côtières que le couple (0.0949;0.0794) proposé par Gordon et al. (1988) destiné plutôt aux eaux de haute mer, par (0.089;0.125), simple moyenne des deux précédents.

Toutefois, une des limitations de cette modélisation vient du fait que différentes valeurs de $a(\lambda)$ et de $b_b(\lambda)$ peuvent donner le même rapport $\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}$, et donc la même valeur de $G(\lambda)$. Pourtant, étant donné que ce paramètre $G(\lambda)$ dépend en partie de la distribution angulaire de la lumière, et que celle-ci est différente selon que la lumière ait été rétro-diffusée par les molécules d'eau ou les particules (on entend par là tout type de particules, organiques ou non organiques), la dépendance de $G(\lambda)$ en $a(\lambda)$ et $b_b(\lambda)$ à travers la seule fraction $\frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}$ est insuffisante pour décrire ses variations. Lee et al. (2004) proposent ainsi de scinder en deux cette fonction :

$$G(\lambda) = g_w \frac{b_{bw}(\lambda)}{b_b(\lambda)} + g_p(\lambda) \frac{b_{bp}(\lambda)}{b_b(\lambda)} \quad (2.12)$$

où g_w est une constante et $g_p(\lambda)$ une fonction de $a(\lambda)$ et $b_b(\lambda)$. On voit bien ici que la nouvelle fonction G est la somme d'un terme modélisant l'influence de la diffusion par les molécules d'eau, et d'un terme représentant l'influence de la diffusion par les particules. En revanche, les auteurs ne fournissent pas d'expressions analytiques pour g_w et $g_p(\lambda)$: la constante g_w ainsi que celles intervenant dans l'expression semi-analytique de $g_p(\lambda)$ sont retrouvées en minimisant l'erreur de modélisation calculée à partir des spectres générés par le logiciel Hydrolight. Ce modèle a été repris depuis par Cannizzaro et Carder (2006).

Enfin, Albert et Mobley (2003) ont développé une nouvelle paramétrisation à l'aide d'Hydrolight, fonction de $x(\lambda) = \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}$, du cosinus de l'angle zénithal solaire $\cos \theta_s$, du cosinus de l'angle de visée $\cos \theta_v$, et de la vitesse du vent u :

$$r_{\infty}(\lambda) = p_1 \left(1 + p_2 x(\lambda) + p_3 x(\lambda)^2 + p_4 x(\lambda)^3 \right) \left(1 + p_5 \frac{1}{\cos \theta_s} \right) \times (1 + p_6 u) \left(1 + p_7 \frac{1}{\cos \theta_v} \right) x(\lambda). \quad (2.13)$$

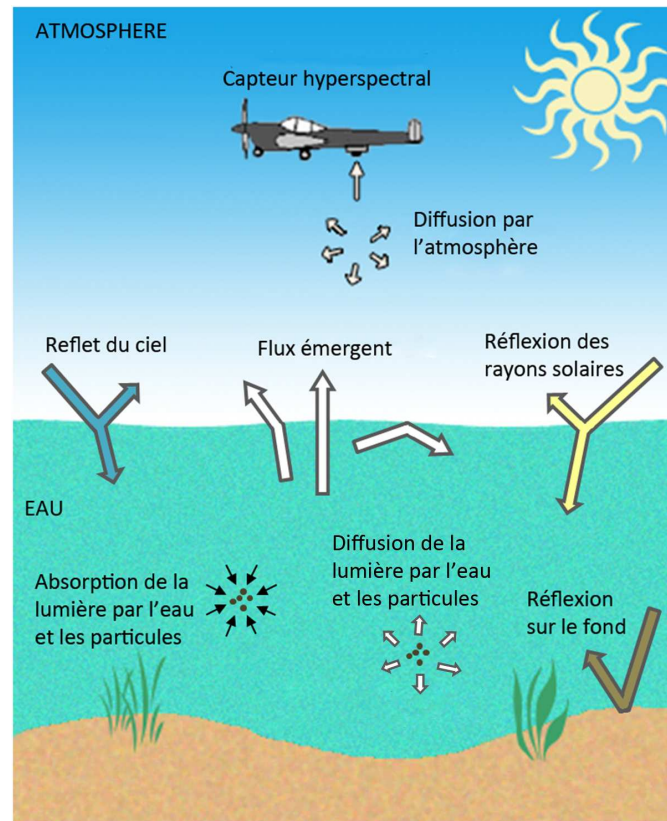


FIGURE 2.1 – Schéma représentant les interactions entre la lumière incidente et l'atmosphère, la colonne d'eau et le fond.

Les valeurs numériques de p_1 , p_2 , p_3 , p_4 , p_5 , p_6 , et p_7 sont données dans l'article original.

Dans cette thèse, nous avons choisi de travailler avec la modélisation de Lee et al. (1999) (Equation 2.11), celle-ci ayant été utilisée et validée dans de nombreux articles ces dernières années.

2.2.2 Réflectance en profondeur finie

Modélisation de la réflectance

En zones côtières, lorsque le fond est visible et influence la réflectance subsurface, une autre modélisation est utilisée. Celle-ci est décrite sur la Figure 2.1. Une partie des photons entrant dans l'eau sont réfléchis par la colonne d'eau elle-même et ressortent sans avoir interagi avec le fond. Cela correspond au terme $r_\infty(\lambda)$ évoqué précédemment. A cela s'ajoutent les photons ayant été réfléchis par le fond et qui vont également influencer la réflectance subsurface. Dans les deux cas, ces flux d'énergie sont atténués par la colonne d'eau à travers des processus d'absorption et de rétro-diffusion. Quatre paramètres physiques dits "optiquement actifs" sont reconnus comme principaux responsables de la déformation des spectres (en plus des molécules d'eau) : la profondeur H , et les concentrations en phytoplancton C_{PHY} , en matière organique dissoute colorée (ou substance jaune, CDOM) C_{CDOM} et en particules non algales C_{NAP} .

Un premier modèle de réflectance (en irradiance), nommé SSI (Single Scattered Irradiance), est développé par Philpot (1987) :

$$R(\lambda) = R_\infty(\lambda) + (R_{0,B}(\lambda) - R_\infty(\lambda)) e^{-(k_d(\lambda) + k_u(\lambda))H} \quad (2.14)$$

où $R_{0,B}(\lambda)$ représente la réflectance en irradiance du fond, et $k_d(\lambda)$ (resp. $k_u(\lambda)$) le coefficient d'atténuation dans le sens descendant (resp. ascendant). Par construction du modèle SSI, un photon ne peut être rétro-diffusé qu'une seule fois (même si les valeurs des paramètres $R_\infty(\lambda)$, $k_d(\lambda)$, et $k_u(\lambda)$ prennent en compte les diffusions multiples), et par conséquent, l'atténuation de la lumière rétro-diffusée vers la surface est ensuite supposée uniquement causée par le phénomène d'absorption. Ajouté au fait que les structures géométriques des flux descendant et ascendant sont différentes, cela explique que les coefficients $k_d(\lambda)$ et $k_u(\lambda)$ ne sont pas égaux.

Maritorea et al. (1994) complexifient ce modèle en distinguant les photons ascendants ayant interagi ou non avec le fond ; en effet, la structure géométrique de ces deux flux d'énergie étant différente, ils ne sont pas atténués de la même façon :

$$R(\lambda) = R_\infty(\lambda) + \left(R_{0,B}(\lambda) e^{-k_u^b(\lambda)H} - R_\infty(\lambda) e^{-k_u^c(\lambda)H} \right) e^{-k_d(\lambda)H} \quad (2.15)$$

où $k_u^b(\lambda)$ (resp. $k_u^c(\lambda)$) représente le coefficient d'atténuation des photons ayant interagi (resp. n'ayant pas interagi) avec le fond.

Toutefois, en pratique, $k_u^b(\lambda)$ et $k_u^c(\lambda)$ sont difficilement mesurables (Philpot 1989), et Maritorea et al. (1994) proposent de n'utiliser qu'un seul coefficient d'atténuation $k(\lambda)$, choisi égal à $k_d(\lambda)$:

$$R(\lambda) = R_\infty(\lambda) + (R_{0,B}(\lambda) - R_\infty(\lambda)) e^{-2k(\lambda)H}. \quad (2.16)$$

Ce modèle simplifié est repris notamment par Vahtmäe et al. (2008), Bertels et al. (2008), Minghelli-Roman et al. (2007), et Lee (2003).

Lee et al. (1998a) actualisent le modèle de l'Equation 2.15 en développant un modèle semi-analytique inversible, permettant de retrouver la profondeur et les propriétés optiques de la colonne d'eau à partir de la réflectance. En générant des spectres de réflectance grâce au logiciel Hydrolight, les auteurs paramétrisent empiriquement les coefficients $k_d(\lambda)$, $k_u^b(\lambda)$ et $k_u^c(\lambda)$ en fonction de l'absorption, de la rétro-diffusion, et de l'angle zénithal solaire. Ils expriment ainsi la réflectance en radiance $r(\lambda)$ telle que :

$$r(\lambda) = r_\infty(\lambda) + \left(A_1 R_{0,B}(\lambda) e^{-k_u^b(\lambda)H} - A_0 r_\infty(\lambda) e^{-k_u^c(\lambda)H} \right) e^{-k_d(\lambda)H} \quad (2.17)$$

où $A_0 = 1.03$ et $A_1 = 0.31$.

Ce modèle est ensuite actualisé par Lee et al. (1999), qui prend notamment en compte des angles de visée non nuls :

$$r(\lambda) = r_\infty(\lambda) + \left(\frac{1}{\pi} R_{0,B}(\lambda) e^{-k_u^b(\lambda)H} - r_\infty(\lambda) e^{-k_u^c(\lambda)H} \right) e^{-k_d(\lambda)H}. \quad (2.18)$$

Ici, le fond est supposé lambertien, et par conséquent, le facteur d'anisotropie associé est choisi égal à π . Jusqu'à présent, ce dernier modèle a été très utilisé et beaucoup d'auteurs le considèrent comme référence (Brando et al. 2009, Albert et Mobley 2003, Fearn et al. 2011, Cannizzaro et Carder 2006, Lee et Carder 2002, Xiu et al. 2009, Gerardino-Neira et al. 2008, Volpe et al. 2011).

Plus récemment et toujours à l'aide d'Hydrolight, Albert et Mobley (2003) ont développé de nouvelles paramétrisations pour $r_\infty(\lambda)$, $k_u^b(\lambda)$, $k_u^c(\lambda)$ et $k_d(\lambda)$, plus complexe et utilisée avec le modèle général de Lee et al. (1999). Celles-ci sont optimisées pour les eaux de type II, riches en CDOM et en particules en suspension (Morel et Prieur 1977), et font notamment intervenir la vitesse du vent.

Dans nos travaux, nous avons tour à tour utilisé le modèle simplifié de Maritorena et al. (1994) (Equation 2.16) que l'on nommera M1, puis celui de Lee et al. (1999) (Equation 2.18) que l'on nommera M2.

Modélisation des coefficients d'atténuation

Ces grandeurs, caractéristiques de la composition de la colonne d'eau, ont été sujettes à nombre de paramétrisations différentes ces dernières décennies.

Etant donnée la structure géométrique différente des flux d'énergie ascendant et descendant, il est communément admis que l'atténuation ne soit pas identique dans les deux sens. Souvent, c'est la difficulté d'accès au coefficient d'atténuation dans la direction montante $k_u(\lambda)$ qui conduit les auteurs à considérer que les deux paramètres sont égaux (Philpot 1987; 1989, Maritorena et al. 1994). En revanche, la structure du champ lumineux descendant étant moins complexe, l'évaluation du coefficient $k_d(\lambda)$ est plus facile, et par conséquent, davantage de paramétrisations ont pu être étudiées.

A l'aide de méthodes de simulation de Monte-Carlo, Kirk (1981) propose ainsi l'expression suivante en fonction du coefficient d'absorption $a(\lambda)$ et du coefficient de diffusion $b(\lambda)$:

$$k_d(z_m, \lambda) = (a(\lambda)^2 + 0.256 a(\lambda) b(\lambda))^{0.5} \quad (2.19)$$

où z_m est la profondeur située au milieu de la zone euphotique, c'est-à-dire la profondeur pour laquelle l'irradiance est égale à 10% de la valeur mesurée juste sous la surface.

Afin de traduire la dépendance de $k_d(\lambda)$ en la structure du champ incident, Kirk (1984) développe, de la même façon, une nouvelle expression faisant aussi intervenir le cosinus μ_d défini dans la Section 2.2.1 :

$$k_d(z_m, \lambda) = \frac{1}{\mu_d} (a(\lambda)^2 + (0.473\mu_d - 0.218) a(\lambda) b(\lambda))^{0.5}. \quad (2.20)$$

Cette formulation a notamment été reprise par Vahtmäe et al. (2008).

Gordon (1989) développe une expression reprise notamment par Albert et Mobley (2003) et Morel et al. (2007) :

$$k_d(\lambda) = \kappa \frac{a(\lambda) + b_b(\lambda)}{\mu_d}. \quad (2.21)$$

Différentes valeurs de κ sont utilisées : Gordon (1989), et Morel et al. (2007) prennent $\kappa = 1.0395$, tandis que Albert et Mobley (2003) et Gege (2004) utilisent $\kappa = 1.0546$. Pour McKee et al. (2003), κ varie entre 1.02 et 1.06. Enfin, Sathyendranath et Platt (1997), Lee et al. (1998a), et Fearn et al. (2011) choisissent de prendre $\kappa = 1$.

En ce qui concerne le coefficient d'atténuation dans le sens ascendant, Sathyendranath et Platt (1997) proposent de faire l'analogie avec l'expression de $k_d(\lambda)$ ci-dessus :

$$k_u(\lambda) = \frac{a(\lambda) + b_b(\lambda)}{\mu_u} \quad (2.22)$$

où μ_u est le cosinus moyen de l'angle fait par la lumière ascendante avec la normale à la surface. Toutefois, de par la structure différente des flux descendant et montant (ce dernier étant plus diffus), les auteurs reconnaissent que cette relation n'est valable que lorsque l'influence de la diffusion est faible par rapport celle de l'absorption, et que par conséquent, elle doit être utilisée avec prudence.

Comme évoqué dans la Section 2.2.2, Lee et al. (1998a) distinguent les photons ayant interagi ou non avec le fond (respectivement $k_u^b(\lambda)$ et $k_u^c(\lambda)$) :

$$k_u^b(\lambda) = 1.1(a(\lambda) + b_b(\lambda)) \left(1 + 4.9 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.5} \quad (2.23)$$

$$k_u^c(\lambda) = 1.2(a(\lambda) + b_b(\lambda)) \left(1 + 2.0 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.5}. \quad (2.24)$$

Lee et al. (1999) actualisent ensuite ces expressions, en faisant notamment intervenir l'angle de visée θ_v :

$$k_u^b(\lambda) = \frac{1}{\cos \theta_v} (1.04(a(\lambda) + b_b(\lambda))) \left(1 + 5.4 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.5} \quad (2.25)$$

$$k_u^c(\lambda) = \frac{1}{\cos \theta_v} (1.03(a(\lambda) + b_b(\lambda))) \left(1 + 2.4 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.5} \quad (2.26)$$

$$(2.27)$$

Albert et Mobley (2003) développent un modèle qui est optimisé pour leur zone d'étude, à savoir le lac de Constance en Allemagne, et qui est plutôt adapté aux eaux de type II. Ils expriment $k_u^c(\lambda)$ et $k_u^b(\lambda)$ tels que :

$$k_u^{b,c}(\lambda) = (a(\lambda) + b_b(\lambda)) \left(1 + \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{q_1} \left(1 + q_2 \frac{1}{\cos \theta_s} \right). \quad (2.28)$$

Deux ensembles $\{q_1, q_2\}$ sont calculés pour $k_u^b(\lambda)$ et $k_u^c(\lambda)$. Cette modélisation est notamment reprise par Matthews et al. (2010).

Pour les mêmes raisons que pour la réflectance d'une colonne d'eau infinie, nous avons choisi la modélisation de Lee et al. (1999) pour représenter les trois coefficients d'atténuation.

En ce qui concerne μ_d , comme une majorité d'auteurs, on le choisit égal à $\cos \theta_s$, où θ_s est l'angle zénithal solaire. Ce dernier correspond à l'angle entre la direction du soleil et la verticale du lieu étudié. Il s'exprime en fonction de la latitude Φ , de la déclinaison solaire δ , et de l'heure t , et est donné par la relation suivante (Campbell et Norman 1998) :

$$\cos \theta_s = \sin \Phi \sin \delta + \cos \Phi \cos \delta \cos [15(t - 12)] \quad (2.29)$$

La déclinaison solaire est égal à l'angle entre la direction du soleil et le plan de l'équateur, et dépend uniquement du jour de l'année (Campbell et Norman 1998). Pour les images HypLitt ayant été prises entre le 14 et le 17 septembre, cela correspond à une déclinaison d'environ -4° . La latitude est elle égale à $47^\circ 30'$.

Enfin, dans tous nos tests, l'angle de visée θ_v est pris égal à 0. On suppose ainsi que les images ont été prises à la verticale de la scène.

2.3 MODÉLISATION DES PARAMÈTRES PHYSIQUES

Les phénomènes d'absorption et de rétro-diffusion, responsables de l'atténuation dépendent bien évidemment de la composition de l'eau. De nombreuses modélisations ont été utilisées dans la littérature, et diffèrent essentiellement par le nombre et le type de constituants pris en compte, et les spectres d'absorption et de rétro-diffusion spécifiques associés à chacun d'entre eux. Communément, en plus des molécules d'eau elles-même, trois constituants sont reconnus pour jouer un rôle important sur la déformation du spectre : le phytoplancton, le CDOM, et les particules non algales (ou tripton). Certains auteurs décrivent ainsi l'effet de chaque composant séparément, tandis que d'autres regroupent certains d'entre eux en un seul terme. D'autres encore décomposent par exemple l'influence du phytoplancton en la somme des influences de chaque type de phytoplancton. La taille des particules peut également être un facteur de différenciation.

Pour cette raison, il est donc nécessaire de prendre certaines précautions avant de regrouper sous la même expression plusieurs termes issus de modèles différents, afin, par exemple, de ne pas doubler la modélisation de l'influence d'un constituant.

Les coefficients d'absorption et de rétro-diffusion s'écrivent comme la somme des contributions de chaque composant i :

$$a(\lambda) = \sum_i a_i(\lambda) \quad (2.30)$$

$$b_b(\lambda) = \sum_i b_{b,i}(\lambda). \quad (2.31)$$

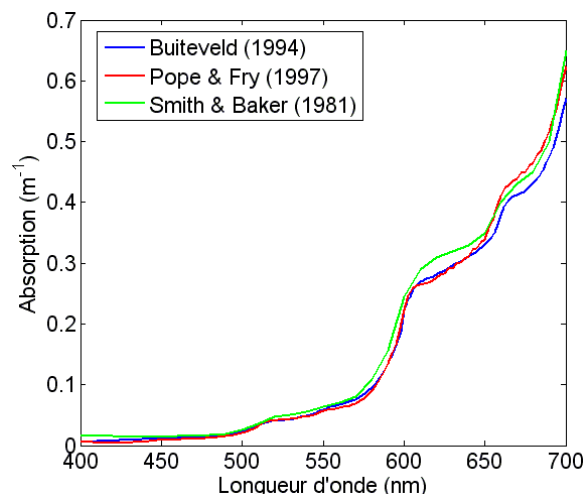


FIGURE 2.2 – Différents spectres d'absorption de l'eau pure.

2.3.1 Absorption

Influence de l'eau pure

Le coefficient d'absorption de l'eau de mer pure a fait l'objet de plusieurs études et de nombreux auteurs ont proposé des spectres qui diffèrent malgré tout quelque peu. Nous en avons sélectionné trois qui interviennent fréquemment dans les articles les plus importants :

- le modèle de Pope et Fry (1997), utilisé notamment par Lee et al. (1998a; 1999), Brando et al. (2009), Giardino et al. (2007), Cannizzaro et Carder (2006), Carder et al. (2004), Matthews et al. (2010), Fearn et al. (2011), Xiu et al. (2009),
- le modèle de Smith et Baker (1981), utilisé par Giardino et al. (2007), Vahtmäe et al. (2008), Forget et al. (1997), Kutser (2004), et
- le modèle de Buiteveld et al. (1994) utilisé par Albert et Mobley (2003).

Ces trois spectres sont représentés sur la Figure 2.2.

En théorie, l'absorption dépend de la salinité et de la température (Gege 2004, Sullivan et al. 2006), mais en pratique, peu d'auteurs prennent en compte ces influences.

Dans cette thèse, nous utilisons le spectre moyen, calculé à partir des trois modèles présentés.

Influence du phytoplancton

La difficulté réside ici dans le fait que chaque espèce de phytoplancton a un effet différent sur l'absorption, comme on peut le voir sur l'étude menée par Ahn et al. (1992). Sans aucune connaissance *a priori* sur la composition du phytoplancton, il est donc délicat de modéliser précisément l'absorption $a_{PHY}(\lambda)$, chaque représentation étant adaptée à la zone d'étude.

Deux types de modélisation sont utilisées dans la littérature. La première fait intervenir le spectre d'absorption spécifique du phytoplancton

$a_{PHY}^*(\lambda)$ et sa concentration, et exprime l'absorption telle que :

$$a_{PHY}(\lambda) = C_{PHY} a_{PHY}^*(\lambda) \quad (2.32)$$

$a_{PHY}(\lambda)$ étant exprimée en m^{-1} , C_{PHY} en $\mu g.L^{-1}$, et $a_{PHY}^*(\lambda)$ en $L.m^{-1}.\mu g^{-1}$. Le spectre $a_{PHY}^*(\lambda)$ est en général déterminé pour un type de phytoplancton, et par conséquent, il est différent selon les lieux d'étude. Par exemple, Brando et al. (2009) utilisent ainsi un spectre adapté à la Baie de Moreton (Australie), Giardino et al. (2007) au Lac de Garde (Italie), Keller (2001), Albert et Mobley (2003), et Vahtmäe et al. (2008) au Lac de Constance (Allemagne), et Hakvoort et al. (2002) au Veluwemeer (Pays-Bas). Bricaud et al. (1995) ont étudié l'évolution de ce spectre dans différentes régions du monde et couvert une large gamme de concentrations.

Enfin, Gege (2004) propose de prendre en compte la contribution de chaque espèce i :

$$a_{PHY}(\lambda) = \sum_i C_{PHY,i} a_{PHY,i}^*(\lambda). \quad (2.33)$$

Par ailleurs, une autre modélisation est introduite par Lee (1994). Celle-ci fait intervenir la valeur de l'absorption $a_{PHY}(\lambda)$ à 440 nm, et deux coefficients $a_0(\lambda)$ et $a_1(\lambda)$ dérivés empiriquement :

$$a_{PHY}(\lambda) = \{a_0(\lambda) + a_1(\lambda) \ln [a_{PHY}(440)]\} a_{PHY}(440). \quad (2.34)$$

Le terme $a_{PHY}(440)$ peut s'exprimer en fonction de C_{PHY} (Prieur et Sathyendranath 1981) :

$$a_{PHY}(440) = 0.06 C_{PHY}^{0.65}. \quad (2.35)$$

Cette formulation a été reprise maintes fois (Lee et al. 1998a; 1999, Lee et Carder 2002, Cannizzaro et Carder 2006, Lee 2009). Xiu et al. (2009) l'a également utilisée en recalculant les spectres $a_0(\lambda)$ et $a_1(\lambda)$.

Dans cette thèse, nous utilisons deux modélisations : la formulation 2.32 avec l'absorption spécifique de Bricaud et al. (1995), et celle de Lee (1994) (Equation 2.34).

Influence de la matière organique dissoute colorée (CDOM)

De par sa forme dissoute, ce constituant n'intervient sur l'atténuation qu'à travers l'absorption. Dans la quasi-totalité des articles concernés, cette influence se matérialise par un terme exponentiel en λ , dont la pente S_{CDOM} varie selon les études :

$$a_{CDOM}(\lambda) = a_{CDOM}(\lambda_0) e^{-S_{CDOM}(\lambda-\lambda_0)}. \quad (2.36)$$

Par analogie avec les autres constituants, comme un certain nombre d'auteurs, nous posons $C_{CDOM} = a_{CDOM}(\lambda_0)$, C_{CDOM} s'exprimant ainsi en m^{-1} . En ce qui concerne la longueur d'onde de référence λ_0 , elle est souvent prise égale à 440 nm (Lee et al. 1999, Brando et al. 2009, Giardino et al. 2007, Albert et Mobley 2003, Gege 2004, Fearn et al. 2011), voire à 400 nm (Pierson et Strömbeck 2001, Carder et al. 2004, Vahtmäe et al. 2008, Kutser

et al. 2005). Enfin, la valeur de la pente S_{CDOM} peut varier du simple au double selon le lieu d'étude : par exemple, Albert et Mobley (2003) prennent $S_{CDOM} = 0.014 \text{ nm}^{-1}$ pour l'étude du Lac de Constance (Allemagne), tandis que Keller (2001) choisit $S_{CDOM} = 0.027 \text{ nm}^{-1}$ pour le Lac de Zurich (Suisse). Toutefois, il semble que la valeur la plus représentative pour une grande variété d'eaux (Gege 2004) soit $S_{CDOM} \approx 0.015 \text{ nm}^{-1}$ (Forget et al. 1997, Lee et al. 1998a, Brando et al. 2009, Fearn et al. 2011). Par ailleurs, il convient de remarquer que la valeur de la pente dépend aussi du type de constituant pris en compte. En effet, le CDOM ayant le même effet sur l'absorption que la matière organique particulaire (Cannizzaro et Carder 2006, Carder et al. 2004), certains auteurs les regroupent sous le même terme a_{CDOM} et utilisent une valeur moyenne pour S_{CDOM} . Lee et al. (1999), Cannizzaro et Carder (2006), Lee et Carder (2003) s'accordent sur $S_{CDOM} = 0.015 \text{ nm}^{-1}$, Lee (2009) utilise $S_{CDOM} = 0.019 \text{ nm}^{-1}$, et Xiu et al. (2009) choisissent $S_{CDOM} = 0.012 \text{ nm}^{-1}$. Du fait de la difficulté liée au choix de cette valeur, Schwarz et al. (2002) proposent ainsi, lorsque C_{CDOM} est grand, de considérer S_{CDOM} comme une inconnue, et de l'estimer au même titre que C_{CDOM} .

Dans cette thèse, nous utilisons la modélisation de Brando et al. (2009), où $\lambda_0 = 440 \text{ nm}$ et $S_{CDOM} = 0.0157 \text{ nm}^{-1}$.

Influence des particules non algales

Bien que l'influence de ce constituant sur l'absorption soit négligée dans de nombreux articles (Lee et al. 1999, Cannizzaro et Carder 2006, Fearn et al. 2011, Lee 2009, Xiu et al. 2009, Gege 2004), elle est tout de même prise en compte régulièrement. La décroissance exponentielle de ce terme peut être représentée par deux modélisations différentes.

Dans la première, l'absorption s'exprime comme le produit de la concentration en particules et de l'absorption spécifique des particules (Sathyendranath et al. 1989, Vahtmäe et al. 2008, Hakvoort et al. 2002, Kutser 2004) :

$$a_{NAP}(\lambda) = C_{NAP} a_{NAP}^*(\lambda). \quad (2.37)$$

Le spectre $a_{NAP}^*(\lambda)$ est notamment donné par Kutser (1997) ou Hakvoort et al. (2002).

L'absorption des particules peut également s'exprimer par un terme exponentiel centré en $\lambda_0 = 440 \text{ nm}$ (Brando et al. 2009, Giardino et al. 2007) :

$$a_{NAP}(\lambda) = C_{NAP} a_{NAP}^*(\lambda_0) e^{-S_{NAP}(\lambda-\lambda_0)}. \quad (2.38)$$

Dans notre étude, on utilise la modélisation exponentielle de Brando et al. (2009), avec $a_{NAP}^*(\lambda_0) = 0.0048 \text{ m}^2 \cdot \text{g}^{-1}$ et $S_{NAP} = 0.0106 \text{ nm}^{-1}$.

En résumé

Nous prenons donc en compte l'influence des trois constituants présentés précédemment dans la modélisation du coefficient d'absorption

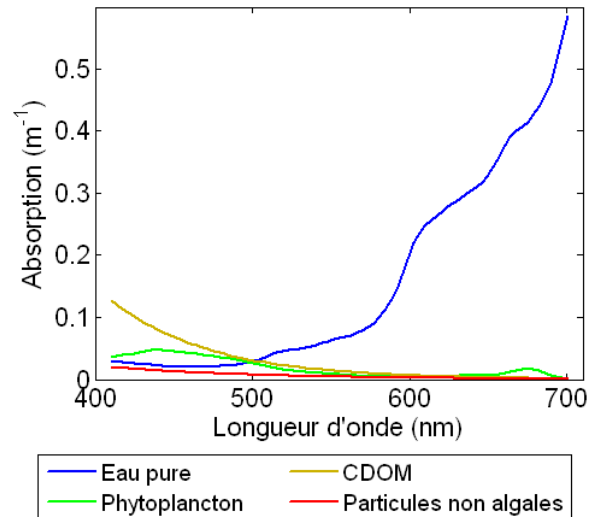


FIGURE 2.3 – Spectres d'absorption de l'eau pure salée (Pope et Fry 1997, Smith et Baker 1981, Buiteveld et al. 1994), du phytoplancton ($C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, Lee (1994)), du CDOM ($C_{CDOM} = 0.08 \text{ m}^{-1}$, Brando et al. (2009)), et des particules non algales ($C_{NAP} = 3 \text{ mg.L}^{-1}$, Brando et al. (2009)).

$a(\lambda)$:

$$a(\lambda) = a_w(\lambda) + a_{PHY}(\lambda) + a_{CDOM}(\lambda) + a_{NAP}(\lambda). \quad (2.39)$$

$a_w(\lambda)$ est donné par le spectre moyen des spectres donnés par Pope et Fry (1997), Smith et Baker (1981), Buiteveld et al. (1994). Pour $a_{PHY}(\lambda)$, on utilise les deux modèles présentés par Lee (1994) d'une part, et par Brando et al. (2009) d'autre part (avec le spectre d'absorption spécifique donné par Bricaud et al. (1995)). $a_{CDOM}(\lambda)$ et $a_{NAP}(\lambda)$ sont quant à eux décrits par Brando et al. (2009).

Afin d'illustrer les influences de chaque élément, nous représentons leur spectre d'absorption sur la Figure 2.3.

2.3.2 Rétro-diffusion

Influence de l'eau pure

Le coefficient de rétro-diffusion de l'eau pure $b_{b,w}(\lambda)$ peut s'exprimer en fonction du coefficient de diffusion de l'eau pure. Une probabilité de rétro-diffusion de 50% est couramment admise (Sathyendranath et al. 1989, Forget et al. 1997, Albert et Mobley 2003, Kutser 2004, Vahtmäe et al. 2008), c'est-à-dire que les molécules d'eau distribuent également les photons dans les demi-espaces avant et arrière par rapport au rayonnement incident. L'effet de la salinité peut aussi pris en compte avec l'ajout d'un facteur multiplicatif de 1.3 (Morel et al. 2007). Le coefficient de diffusion de l'eau pure est ensuite donné par Smith et Baker (1981) ou Buiteveld et al. (1994).

Une autre formulation développée par Morel (1974) est également très utilisée (Richardson et LeDrew 2006, Lee et al. 1998a; 1999; 1998b, Brando et al. 2009, Giardino et al. 2007, Keller 2001, Cannizzaro et Carder 2006, Carder et al. 2004, Lee 2009, Xiu et al. 2009, Lee et Carder 2002, Gege 2004).

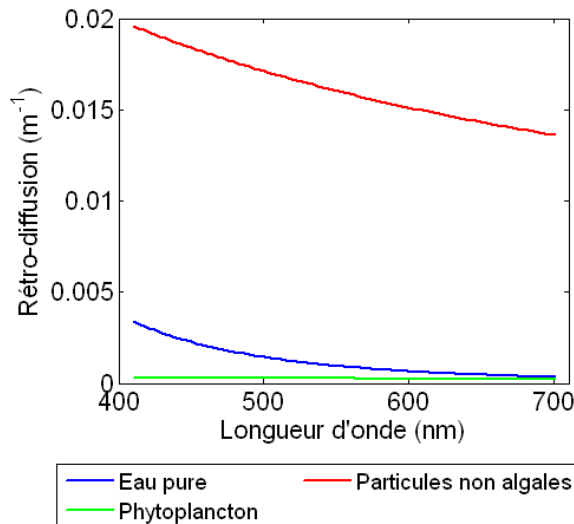


FIGURE 2.4 – Spectres de rétro-diffusion de l'eau pure salée (Morel 1974), du phytoplancton ($C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, Brando et al. (2009)), et des particules non algales ($C_{NAP} = 3 \text{ mg.L}^{-1}$, Brando et al. (2009)).

Elle exprime l'absorption de l'eau pure sous la forme suivante :

$$b_{b,w}(\lambda) = b_1 \times \left(\frac{\lambda}{\lambda_1} \right)^{-4.32} \quad (2.40)$$

avec $\lambda_1 = 500 \text{ nm}$. Le coefficient b_1 peut être ajusté en fonction de la salinité : pour l'eau douce, on a $b_1 = 0.00111 \text{ m}^{-1}$, tandis que pour de l'eau de mer avec une salinité de 35-38%, on a $b_1 = 0.00144 \text{ m}^{-1}$.

Dans cette thèse, nous utilisons le modèle de Morel (1974) ci-dessus (Equation 2.40).

Influence du phytoplancton et des particules non algales

On regroupe ici ces deux constituants dans la même section car il est courant de représenter leur effet sur la rétro-diffusion à l'aide d'un seul terme. Comme on peut le voir sur la Figure 2.4, les particules non algales ont une importance prépondérante, tant et si bien que quelques auteurs négligent l'influence du phytoplancton, voire même de l'eau pure sur la rétro-diffusion (Volpe et al. 2011).

De nombreux articles décrivent ainsi l'effet de l'ensemble des particules (chlorophylliennes et non-chlorophylliennes) sans distinction. La majorité d'entre eux l'exprime sous la forme d'une fonction puissance négative de λ :

$$b_{b,PHY+NAP}(\lambda) = b_{b,PHY+NAP}(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{PHY+NAP}} \quad (2.41)$$

Le choix de la longueur d'onde de référence λ_2 est très variable selon les articles : 550 nm pour Lee et al. (1998a), Fearn et al. (2011), Cannizzaro et Carder (2006), 532 nm pour Xiu et al. (2009), ou encore 440 nm pour Lee (2009). L'exposant $Y_{PHY+NAP}$ varie également et dépend notamment de la

taille des particules : ainsi, si certains utilisent une valeur fixée (Lee et al. 1998a, Lee 2009, Fearn et al. 2011), d'autres l'estiment empiriquement à partir de valeurs de réflectance (Lee et al. 1999, Cannizzaro et Carder 2006, Xiu et al. 2009).

Par ailleurs, quelques auteurs décrivent cette influence comme le produit de la concentration en particules et de leur spectre d'absorption spécifique, considéré comme constant ou quasi-constant (Hakvoort et al. 2002, Albert et Mobley 2003).

D'autre part, il est tout de même fréquent que l'effet du phytoplancton soit décrit séparément de celui des particules non algales. Comme précédemment, on utilise souvent le produit des concentrations et des spectres d'absorption spécifique (Kutser et al. 2001, Vahtmäe et al. 2008, Giardino et al. 2007), ou la fonction puissance de λ (Brando et al. 2009, Giardino et al. 2007).

Dans cette thèse, dans la continuité des choix précédents, nous utilisons la modélisation proposée par Brando et al. (2009) :

$$b_{b,PHY}(\lambda) = C_{PHY} b_{b,PHY}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{PHY}} \quad (2.42)$$

$$b_{b,NAP}(\lambda) = C_{NAP} b_{b,NAP}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{NAP}} \quad (2.43)$$

avec $\lambda_2 = 542 \text{ nm}$, $b_{b,PHY}^*(\lambda_2) = 0.00038 \text{ m}^2.\text{mg}^{-1}$, $Y_{PHY} = 0.681$, $b_{b,NAP}^*(\lambda_2) = 0.0054 \text{ m}^2.\text{g}^{-1}$, et $Y_{NAP} = 0.681$.

En résumé

En plus des molécules d'eau, nous prenons donc en compte l'influence du phytoplancton et des particules non algales dans la modélisation du coefficient de rétro-diffusion $b_b(\lambda)$:

$$b_b(\lambda) = b_{b,w}(\lambda) + b_{b,PHY}(\lambda) + b_{b,NAP}(\lambda). \quad (2.44)$$

$b_{b,w}(\lambda)$ est donné par le modèle de Morel (1974). Quant à $b_{b,PHY}(\lambda)$ et $b_{b,NAP}(\lambda)$, leur effet est décrit par Brando et al. (2009).

Afin d'illustrer les influences de chaque élément, nous représentons leur spectre de rétro-diffusion sur la Figure 2.4.

2.3.3 Variabilité du coefficient d'atténuation diffuse ($k_d(\lambda) + k_u(\lambda)$)

Les paramètres $k_d(\lambda)$, $k_u^b(\lambda)$, et $k_u^c(\lambda)$ intervenant dans l'Equation 2.18 ayant été reliés à C_{PHY} , C_{CDOM} , et C_{NAP} , il est intéressant d'observer l'évolution de l'atténuation diffuse $k_d(\lambda) + k_u(\lambda)$ (en négligeant ici les différences entre $k_u^c(\lambda)$ et $k_u^b(\lambda)$) en fonction de ces concentrations. On trace ainsi sur la Figure 2.5 le terme $k_d(\lambda) + k_u^b(\lambda)$ pour plusieurs ensembles $\{C_{PHY}, C_{CDOM}, C_{NAP}\}$. On voit que chaque constituant agit différemment sur le spectre de réflectance. En repérant les régions spectrales où l'atténuation est la plus faible, on peut notamment remarquer que l'on retrouve

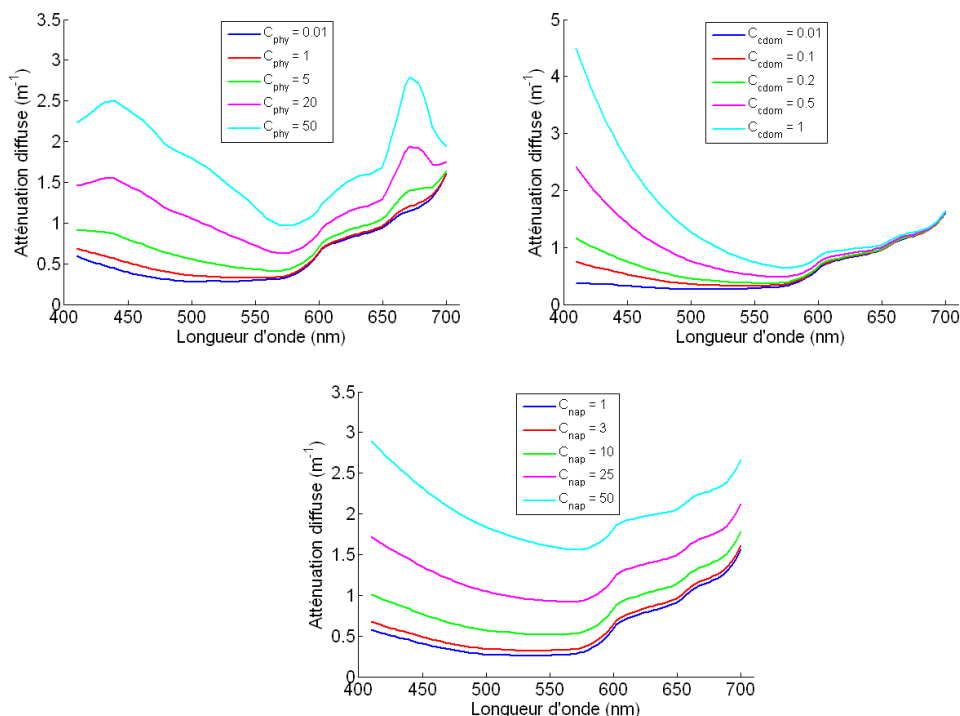


FIGURE 2.5 – Evolution du coefficient d'atténuation diffuse $k_d(\lambda) + k_u^b(\lambda)$ en fonction des concentrations. Pour chaque graphe, on fixe les valeurs des deux autres paramètres invariants à : $C_{\text{PHY}} = 0.7 \mu\text{g}\cdot\text{L}^{-1}$, $C_{\text{CDOM}} = 0.08 \text{ m}^{-1}$, et $C_{\text{NAP}} = 2.8 \text{ mg}\cdot\text{L}^{-1}$.

bien l'origine des teintes verte et jaune des eaux riches en phytoplancton et en CDOM respectivement.

2.4 CONCLUSIONS

Nos travaux portant sur l'étude de zones côtières, nous ne pouvons négliger l'influence du fond, et par conséquent, nous utilisons les deux modèles de réflectance suivants :

- Modèle M1 de Maritorena et al. (1994) :

$$r(\lambda) = r_{\infty}(\lambda) + (r_{0,B}(\lambda) - r_{\infty}(\lambda)) e^{-2k(\lambda)H}$$

- Modèle M2 de Lee et al. (1999) :

$$r(\lambda) = r_{\infty}(\lambda) + \left(\frac{1}{\pi} R_{0,B}(\lambda) e^{-k_u^b(\lambda)H} - r_{\infty}(\lambda) e^{-k_c(\lambda)H} \right) e^{-k_d(\lambda)H}.$$

La réflectance en profondeur infinie est donnée par le modèle de Lee et al. (1999) :

$$r_{\infty}(\lambda) = \left(0.084 + 0.17 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right) \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)}.$$

Les coefficients d'atténuation sont également donnés par les modèles de Lee et al. (1999) :

$$\begin{aligned} k_d(\lambda) &= \frac{a(\lambda) + b_b(\lambda)}{\mu_d} \\ k_u^b(\lambda) &= \frac{1}{\cos \theta_v} (1.04(a(\lambda) + b_b(\lambda))) \left(1 + 5.4 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.5} \\ k_u^c(\lambda) &= \frac{1}{\cos \theta_v} (1.03(a(\lambda) + b_b(\lambda))) \left(1 + 2.4 \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \right)^{0.5} . \end{aligned}$$

Enfin, les spectres d'absorption et de rétro-diffusion sont reliés aux paramètres caractéristiques de la transparence de l'eau via les expressions suivantes :

- Spectre d'absorption total utilisant le modèle d'absorption du phytoplancton de Brando et al. (2009) (utilisé avec le spectre d'absorption spécifique de Bricaud et al. (1995)) :

$$\begin{aligned} a(\lambda) &= a_w(\lambda) + C_{PHY} a_{PHY}^*(\lambda) \\ &\quad + C_{CDOM} e^{-S_{CDOM}(\lambda-\lambda_0)} + C_{NAP} a_{NAP}^*(\lambda_0) e^{-S_{NAP}(\lambda-\lambda_0)} \end{aligned}$$

- Spectre d'absorption total utilisant le modèle d'absorption du phytoplancton de Lee (1994) :

$$\begin{aligned} a(\lambda) &= a_w(\lambda) + \left\{ a_0(\lambda) + a_1(\lambda) \ln \left[0.06 C_{PHY}^{0.65} \right] \right\} 0.06 C_{PHY}^{0.65} \\ &\quad + C_{CDOM} e^{-S_{CDOM}(\lambda-\lambda_0)} + C_{NAP} a_{NAP}^*(\lambda_0) e^{-S_{NAP}(\lambda-\lambda_0)} \end{aligned}$$

- Spectre de rétro-diffusion :

$$\begin{aligned} b_b(\lambda) &= b_{b,w}(\lambda) + C_{PHY} b_{b,PHY}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{PHY}} \\ &\quad + C_{NAP} b_{b,NAP}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{NAP}} . \end{aligned}$$

Par ailleurs, il est important de noter que les modèles utilisés ne prennent pas en compte la fluorescence du phytoplancton et du CDOM, ainsi que la diffusion Raman de l'eau.

Enfin, le fond est supposé lambertien, et la colonne d'eau homogène.

ÉTAT DE L'ART SUR L'ESTIMATION DES PARAMÈTRES DE LA COLONNE D'EAU

3.1 INTRODUCTION

QUE ce soit en imagerie multispectrale ou hyperspectrale, une multitude de méthodes ont été développées pour estimer les paramètres caractéristiques de la colonne d'eau. Hormis la bathymétrie et les concentrations en phytoplancton, en CDOM et en particules non algales, ces techniques permettent également d'obtenir le spectre de réflectance du fond (Mobley et al. 2005), la turbidité (Kallio et al. 2001), ou encore la transparence de l'eau (Thiemann et Kaufmann 2002).

Dans nos travaux, nous nous concentrons sur l'estimation des quatre paramètres évoqués dans le Chapitre 2, à savoir la profondeur H , la concentration en phytoplancton C_{PHY} , la concentration en CDOM ou matière organique dissoute colorée C_{CDOM} , et la concentration en particules non algales C_{NAP} . Tous les autres paramètres intervenant dans le modèle de réflectance subsurface des Equations 2.16 et 2.18 sont supposés connus *a priori*.

Nous présentons ici deux ensembles de méthodes, selon qu'elles permettent d'estimer tout ou partie des paramètres.

3.2 MÉTHODES D'ESTIMATION D'UNE PARTIE DES PARAMÈTRES

Il est à noter que le même type de méthode peut être utilisé pour estimer des paramètres différents.

3.2.1 Méthodes d'estimation de la bathymétrie

Plusieurs approches peuvent être utilisées pour estimer la profondeur. La première d'entre elles consiste à développer une relation empirique reliant les valeurs de radiance ou de réflectance, et la bathymétrie. Ainsi, Philpot (1989) développe une variable fonction de la radiance, et variant linéairement avec la profondeur. Minghelli-Roman et al. (2007) adaptent cette méthode aux images hyperspectrales issues du capteur MeRIS.

D'autres auteurs utilisent une approche par Look-Up Table : grâce au modèle numérique de transfert radiatif Hydrolight, Louchard et al. (2003) et Mobley et al. (2005) créent ainsi une librairie de spectres tabulée en fonction de la hauteur, et choisissent celui qui minimise l'erreur résiduelle entre la mesure et le modèle.

Heege et al. (2007) utilisent également une méthode d'inversion, où une fonction d'erreur entre le spectre de réflectance mesuré et le spectre de réflectance modélisé est minimisée en fonction de la profondeur.

Enfin, Lee et al. (1998b) proposent une approche par réseau de neurones, où l'apprentissage du réseau est effectué sur un ensemble de spectres simulés par le modèle de Lee et al. (1998a).

3.2.2 Méthodes d'estimation de la qualité de l'eau

En ce qui concerne l'estimation des concentrations, des méthodologies similaires à celles présentées ci-dessus ont été développées.

Les algorithmes empiriques reliant certaines concentrations avec une ou plusieurs valeurs de réflectance sont très utilisés. En général, ces techniques sont employées lorsque le fond n'est pas visible et que la réflectance peut être représentée par la réflectance d'une colonne d'eau infinie. Elles s'appuient sur le fait qu'à certaines longueurs d'onde, le signal est principalement influencé par le constituant recherché. Concernant l'estimation de C_{PHY} , O'Reilly et al. (1998) présentent un ensemble de relations du type $C_{PHY} = f\left(\frac{r(\lambda_1)}{r(\lambda_2)}\right)$, où la fonction f est une fonction exponentielle en un polynôme de $\frac{r(\lambda_1)}{r(\lambda_2)}$ (avec λ_1 et λ_2 sont respectivement des longueurs d'onde dans le domaine du bleu et du vert). Cependant, ces algorithmes sont surtout valables dans les eaux de type I à dominance chlorophyllienne (Dall'Olmo et al. 2005); afin d'éviter les influences du CDOM et des particules non algales dans des eaux plus turbides, on utilise davantage des longueurs d'onde proches du rouge, autour de 700 nm (Kallio et al. 2001, Strömbeck et Pierson 2001, Dall'Olmo et al. 2005, Simis et al. 2005, Zimba et Gitelson 2006, Gitelson et al. 2007, Randolph et al. 2008, Le et al. 2009, Matthews et al. 2010). Enfin, dans le cas où le fond est visible, Cannizzaro et Carder (2006) proposent d'adapter l'algorithme d'estimation de C_{PHY} et les bandes prises en compte, selon que le pixel considéré ait été classé auparavant comme "optiquement peu profond" ou "optiquement profond".

Ce type de méthode peut également permettre d'estimer les autres concentrations (Dekker et al. 2001, Kutser et al. 2005, Petus et al. 2010).

Par ailleurs, les mêmes méthodes d'inversion par Look-Up Table ou par minimisation de l'erreur entre la mesure et le modèle que celles appliquées à l'estimation de la bathymétrie, peuvent être aussi utilisées pour retrouver les différentes concentrations (Kutser 2004, Forget et al. 1997, Pierson et Strömbeck 2001, Van Der Woerd et Pasterkamp 2007).

Dans le cas où le fond est supposé infini, la réflectance est modélisée par le terme $r_\infty(\lambda)$, et sous certaines hypothèses sur la paramétrisation des spectres d'absorption et de rétro-diffusion (il faut qu'ils varient linéairement en fonction des paramètres C_{PHY} , C_{CDOM} , et C_{NAP}), et de $r_\infty(\lambda)$ (qui doit être modélisé par l'Equation 2.8), il est possible d'écrire un système linéaire du type $Ax = b$, où $x = (C_{PHY}, C_{CDOM}, C_{NAP})^t$, qui peut se résoudre aisément par $x = (A^t A)^{-1} A^t b$. Cette approche par inversion matricielle a notamment été employée par Keller (2001) et Hakvoort et al. (2002).

Enfin, d'autres techniques utilisant l'Analyse en Composantes Principales ou la dérivée première de la réflectance sont présentées par Sathyendranath et al. (1989) et Han (2005) respectivement.

3.3 MÉTHODES D'ESTIMATION GLOBALES

Pour estimer l'ensemble $\theta = \{H, C_{PHY}, C_{CDOM}, C_{NAP}\}$, deux approches sont utilisées. Dans la première, une fonction d'erreur entre le spectre mesuré et le modèle est minimisée en fonction des paramètres, et la valeur de θ sélectionnée est celle permettant d'obtenir l'erreur minimale. La deuxième approche est une approche par Look-Up Table : on génère une librairie de spectres en faisant varier les différents paramètres du modèle, et chaque observation est comparée à l'ensemble de cette base de données. Une mesure de distance est alors utilisée pour choisir quel ensemble θ permet de construire le spectre le plus proche du signal mesuré.

Cinq variantes de ces méthodes sont décrites et comparées par Dekker et al. (2011).

3.3.1 Inversion par optimisation d'une fonction d'erreur

La méthode originale est développée par Lee et al. (1999). Elle s'appuie sur les modèles de réflectance subsurface de l'Equation 2.18, de réflectance $r_\infty(\lambda)$ de l'Equation 2.11, et des coefficients d'atténuation des Equations 2.21, 2.25 et 2.26.

Le coefficient d'absorption est donné par :

$$a(\lambda) = a_w(\lambda) + a_{PHY}(\lambda) + a_{CDOM}(\lambda) \quad (3.1)$$

$$= a_w(\lambda) + [a_0(\lambda) + a_1(\lambda) \ln(P)] P + G e^{-S_{CDOM}(\lambda-440)} \quad (3.2)$$

où $P = a_{PHY}(440)$, et $G = a_{CDOM}(440)$.

Le coefficient de rétro-diffusion est donné par :

$$b_b(\lambda) = b_{b,w}(\lambda) + b_{b,PHY+NAP}(\lambda) \quad (3.3)$$

$$= b_{b,w}(\lambda) + X \left(\frac{400}{\lambda} \right)^{Y_{PHY+NAP}} \quad (3.4)$$

où $X = b_{b,PHY+NAP}(400)$.

Enfin, la forme de la réflectance du fond est supposée semblable à celle du sable, à un facteur multiplicatif près :

$$R_{0,B}(\lambda) = B \times R_{sable}(\lambda). \quad (3.5)$$

La réflectance dépend alors des seuls paramètres inconnus H , P , G , X , et B . L'algorithme proposé par Lee et al. (1999) consiste à minimiser une fonction objectif comparant le spectre mesuré r et le spectre modélisé \hat{r} .

$$err = \frac{\left[\sum_{\lambda=400}^{675} (r(\lambda) - \hat{r}(\lambda))^2 + \sum_{\lambda=750}^{830} (r(\lambda) - \hat{r}(\lambda))^2 \right]^{0.5}}{\sum_{\lambda=400}^{675} \hat{r}(\lambda) + \sum_{\lambda=750}^{830} \hat{r}(\lambda)}. \quad (3.6)$$

La coupure entre 675 nm et 750 nm permet d'éviter l'influence de la fluorescence du phytoplancton, qui peut apparaître sur les observations, mais qui n'est pas prise en compte par le modèle. Cependant, en eaux peu profondes, cette fenêtre doit être utilisée pour l'inversion (Dekker et al. 2011).

L'initialisation des paramètres est différente pour chaque pixel, et s'appuie sur des expressions empiriques, fonction des valeurs de réflectance à différentes longueurs d'onde (Lee et al. 1999, Dekker et al. 2011).

Ce problème non linéaire d'optimisation par moindres carrés s'effectue par un algorithme de type Levenberg-Marquardt (Levenberg 1944, Marquardt 1963).

Depuis, des variantes de cette méthode ont été développées.

Klonowski et al. (2007) proposent par exemple de modifier la modélisation du fond, en le représentant sous la forme d'une combinaison linéaire de spectres issus de trois classes de fond, les sédiments, la végétation sub-aquatique, et le corail :

$$R_{0,B}(\lambda) = B_{sed,i} R_{sed,i}(\lambda) + B_{veg,j} R_{veg,j}(\lambda) + B_{cor,k} R_{cor,k}(\lambda), \quad (3.7)$$

$$B_{sed,i} + B_{veg,j} + B_{cor,k} = 1$$

où $R_{sed,i}(\lambda)$ (resp. $R_{veg,j}(\lambda)$ et $R_{cor,k}(\lambda)$) est le i -ème spectre dans la classe "Sédiment" (resp. "Végétation" et "Corail").

Brando et al. (2009) proposent également un autre algorithme inspiré de celui de Lee et al. (1999), et appelé SAMBUCA (Semi-Analytical Model for Bathymetry, Unmixing, and Concentration Assessment). La paramétrisation des coefficients d'absorption et de rétro-diffusion change quelque peu par rapport à celle de Lee et al. (1999), et est donnée par les Equations 2.32, 2.36, 2.38, 2.42, et 2.43. Le fond est représenté par une combinaison linéaire de deux spectres, chacun étant issu d'une librairie spectrale. Après optimisation du produit du SAM (Spectral Angle Mapper, Kruse et al. (1993)) et d'une fonction des moindres carrés, cette méthode calcule un indicateur de la visibilité du fond (SDI pour Substrate Detectability Index), donné par l'écart maximal sur les longueurs d'onde entre le spectre de réflectance modélisé et le spectre de réflectance d'une colonne d'eau infinie. A partir d'une certaine hauteur d'eau H^* , on a $SDI=0$, et l'estimation de H devient impossible car on ne peut mesurer aucun signal en provenance du fond (toutes les profondeurs supérieures à H^* pouvant être solution du problème). Le SDI nous permet donc d'avoir une idée sur la fiabilité de l'estimation de la profondeur.

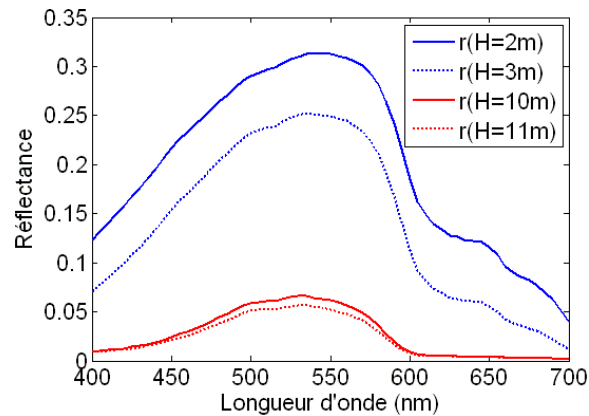


FIGURE 3.1 – Evolution de la réflectance entre H et $H + \Delta H$ ($C_{PHY} = 0.7 \mu\text{g}\cdot\text{L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$, $C_{NAP} = 2.8 \text{ mg}\cdot\text{L}^{-1}$)

3.3.2 Inversion par Look-Up Table

Mobley et al. (2005) proposent de construire la librairie spectrale à l'aide du logiciel Hydrolight. Une large gamme de profondeurs, de types de fond, d'absorption, et de rétro-diffusion notamment est utilisée afin de couvrir le maximum de situations possibles. Chaque pixel est alors comparé à cette base de données via une mesure des moindres carrés, et l'ensemble de paramètres qui permet d'obtenir la distance minimale est choisi.

Un des points importants de la méthode est le choix de la gamme de valeurs et de l'échantillonnage : plus l'intervalle sera large et finement découpé, et meilleure sera la précision de l'estimation. En revanche, le temps de calcul se retrouvera considérablement augmenté. Ce compromis dépend notamment du type d'application.

Hedley et al. (2009) apportent une autre solution à ce problème d'échantillonnage : au lieu d'adopter un pas fixe, ils raisonnent à l'inverse en subdivisant l'espace des paramètres de manière à ce que l'espace des réflectances soit régulièrement échantillonné. En guise d'exemple, on constate qu'une variation de la hauteur de 1 m entraîne une variation de la réflectance beaucoup plus importante pour $H = 2 \text{ m}$ que pour $H = 10 \text{ m}$ (voir Figure 3.1). Dans la méthode de Hedley et al. (2009), le pas de l'échantillonnage est donc bien plus faible pour les faibles profondeurs que pour les fortes profondeurs.

Un des avantages de ces techniques d'inversion par Look-Up Table par rapport aux méthodes d'inversion par optimisation est que l'algorithme permet d'obtenir la solution globale, tandis que la méthode de Levenberg-Marquardt converge vers un minimum local (d'où l'importance de l'initialisation). En revanche, la précision ne peut être inférieure à la borne donnée par le pas de discrétisation, contrairement aux techniques d'inversion par optimisation.

3.4 MÉTHODE UTILISÉE POUR LA COMPARAISON

Dans nos travaux, nous comparons nos résultats d'estimation avec ceux obtenus par une méthode globale inspirée de celle de Brando et al. (2009), et donc de Lee et al. (1999). En effet, la précision obtenue conduit notamment à ce que cette approche, et plus généralement celle de Lee et al. (1999), soit largement utilisée dans le domaine. De plus, la méthode de comparaison est dite inspirée de celle de Brando et al. (2009) car, si les paramètres inconnus sont identiques (c'est-à-dire H , C_{PHY} , C_{CDOM} , et C_{NAP}), en revanche, nous supposons ici que le spectre du fond est connu *a priori*. Le calcul du SDI n'est pas effectué ici. Comme dans la méthode originale, la fonction d'erreur est optimisée avec un algorithme de type Levenberg-Marquardt.

ESTIMATION DE PARAMÈTRES DE L'EAU ET BATHYMÉTRIE : MÉTHODES STATISTIQUES

4.1 INTRODUCTION

Nous présentons ici des méthodes statistiques d'estimation. Dans un premier temps, nous abordons le problème de manière simple, en considérant notamment une colonne d'eau uniforme dans la zone étudiée. Dans un deuxième temps, nous introduisons un modèle linéaire de profondeur afin d'améliorer les résultats d'estimation lorsque nous avons affaire à des fonds variables. Puis, nous étudions l'intérêt d'apporter une information *a priori* sur la distribution des paramètres caractérisant la couche d'eau, en développant l'estimateur du maximum *a posteriori*. Enfin, nous calculons les bornes de Cramér-Rao, outils statistiques donnant une borne inférieure théorique pour la variance de nos estimateurs, afin d'étudier le comportement de notre méthode vis-à-vis de ces bornes.

4.2 PRÉSENTATION DE LA MÉTHODE DU MAXIMUM DE VRAI-SEMBLANCE

4.2.1 Introduction

Même si l'hypothèse de gaussianité pour représenter les données hyperspectrales n'est pas toujours respectée, elle est néanmoins utilisée avec un certain succès, et nous supposons donc ici que le pixel \mathbf{r} suit une loi normale de spectre moyen $\boldsymbol{\mu}$ et de matrice de covariance $\boldsymbol{\Gamma}$. Nous considérons uniquement des spectres purs, et par conséquent, cette variabilité modélise uniquement la variabilité intra-classe du matériau présent dans le pixel. Par ailleurs, il est intéressant de remarquer que, étant donné la linéarité des modèles bathymétriques utilisés, les abondances d'un mélange linéaire de spectres de fond sont conservées après atténuation par la couche d'eau.

Les deux modèles M_1 et M_2 présentés dans le Chapitre 2 sont utilisés lors des différents tests. On peut les réécrire sous une forme matricielle, en faisant apparaître le pixel spectral $\mathbf{r} = [r(\lambda_1) \dots r(\lambda_L)]^t$, où L est le nombre

de bandes spectrales. Le modèle M1 de Maritorena et al. (1994) s'écrit ainsi :

$$\mathbf{r} = \mathbf{r}_\infty + \mathbf{K}(\mathbf{r}_{0,B} - \mathbf{r}_\infty) \quad (4.1)$$

où \mathbf{r}_∞ est la réflectance d'une colonne d'eau infinie donnée par l'Equation 2.11, $\mathbf{r}_{0,B}$ la réflectance du fond, et \mathbf{K} la matrice d'atténuation de l'eau. On a $\mathbf{K} = \text{diag}\left[e^{-2k(\lambda_i)H}\right]_{i \in \llbracket 1;L \rrbracket}$ où $k(\lambda)$ est le coefficient d'atténuation donné par l'Equation 2.21, et H la profondeur.

En ce qui concerne le modèle M2 de Lee et al. (1999), il peut se réécrire de la façon suivante :

$$\mathbf{r} = \mathbf{r}_\infty + \mathbf{K}_B \mathbf{r}_{0,B} - \mathbf{K}_C \mathbf{r}_\infty \quad (4.2)$$

où les matrices \mathbf{K}_B et \mathbf{K}_C représentent respectivement l'atténuation des photons ayant interagi ou non avec le fond. On a

$\mathbf{K}_B = \text{diag}\left[e^{-(k_d(\lambda_i) + k_u^b(\lambda_i))H}\right]_{i \in \llbracket 1;L \rrbracket}$, et $\mathbf{K}_C = \text{diag}\left[e^{-(k_d(\lambda_i) + k_u^c(\lambda_i))H}\right]_{i \in \llbracket 1;L \rrbracket}$, où $k_d(\lambda)$ est le coefficient d'atténuation dans la direction descendante donné par l'Equation 2.21, tandis que $k_u^c(\lambda)$ et $k_u^b(\lambda)$ sont les coefficients d'atténuation dans la direction ascendante donnés par les Equations 2.26 et 2.25.

Dans nos travaux, le modèle M1 est utilisé dans les simulations. Le modèle M2 est utilisé sur les données synthétiques et réelles, notamment afin d'évaluer l'impact d'un modèle plus réaliste sur les résultats d'estimation.

4.2.2 Méthode d'estimation

On suppose ici que l'on dispose d'un ensemble d'apprentissage $\chi_N = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$, constitué de N réalisations indépendantes et identiquement distribuées du vecteur aléatoire \mathbf{r} . Ce dernier est paramétré par l'un des deux modèles M1 ou M2, et comme mentionné précédemment, est distribué tel que :

$$\mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma}) \quad (4.3)$$

où $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ et $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}(\boldsymbol{\theta})$. $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_P\}$ représente l'ensemble des paramètres inconnus du modèle (parmi H , C_{PHY} , C_{CDOM} et C_{NAP}).

Dans la méthode d'estimation du maximum de vraisemblance (MV) présentée, le but est d'estimer cet ensemble en choisissant la valeur de $\boldsymbol{\theta}$ la plus vraisemblable, c'est-à-dire en optimisant la fonction de vraisemblance $P(\chi_N|\boldsymbol{\theta})$ en fonction des paramètres inconnus :

$$\hat{\boldsymbol{\theta}}_{MV}(\chi_N) = \underset{\boldsymbol{\theta}}{\text{argmax}} P(\chi_N|\boldsymbol{\theta}). \quad (4.4)$$

La vraisemblance associée à cet échantillon χ_N s'exprime de la façon suivante :

$$\begin{aligned} P(\chi_N|\boldsymbol{\theta}) &= \prod_i^N \frac{1}{\sqrt{(2\pi)^L |\boldsymbol{\Gamma}|}} e^{-\frac{1}{2}(\mathbf{r}_i - \boldsymbol{\mu})^t \boldsymbol{\Gamma}^{-1} (\mathbf{r}_i - \boldsymbol{\mu})} \\ &= \frac{1}{\sqrt{(2\pi)^L |\boldsymbol{\Gamma}|^N}} e^{-\frac{1}{2} \sum_i^N (\mathbf{r}_i - \boldsymbol{\mu})^t \boldsymbol{\Gamma}^{-1} (\mathbf{r}_i - \boldsymbol{\mu})}. \end{aligned} \quad (4.5)$$

Première approche simplifiée

Il existe de multiples sources de variabilité spectrale : en particulier, on peut citer les variabilités dues au fond, au capteur, aux inhomogénéités de la colonne d'eau, ou encore aux perturbations affectant la surface et qui ne sont pas modélisées (vagues, réflexions spéculaires du soleil ou du ciel, etc).

Dans un premier temps, on cherche les expressions analytiques des estimateurs solutions de l'Equation 4.4. Pour cela, on simplifie le problème en ne conservant que les variabilités dues au fond et au capteur, et, en supposant que celles-ci soient gaussiennes, le pixel spectral \mathbf{r} peut s'écrire, avec le modèle M_1 , tel que :

$$\mathbf{r} = \mathbf{r}_\infty + \mathbf{K}(\mathbf{r}_{0,B} - \mathbf{r}_\infty) + \mathbf{n}_S \quad (4.6)$$

où $\mathbf{r}_{0,B} \sim \mathcal{N}(\boldsymbol{\mu}_{0,B}, \sigma_B^2 \mathbf{I}_{d_L})$ est un vecteur aléatoire décrivant la variabilité du fond, et $\mathbf{n}_S \sim \mathcal{N}(\mathbf{o}, \sigma_S^2 \mathbf{I}_{d_L})$ un vecteur aléatoire représentant le bruit de capteur.

Le vecteur spectral \mathbf{r} , somme de deux vecteurs aléatoires gaussiens indépendants, suit également une loi normale multivariée de moyenne $\boldsymbol{\mu}$, et de matrice de covariance $\boldsymbol{\Gamma}$. On a :

$$\begin{aligned} \boldsymbol{\mu} &= \mathbb{E} [\mathbf{r}_\infty + \mathbf{K}(\mathbf{r}_{0,B} - \mathbf{r}_\infty) + \mathbf{n}_S] \\ \boldsymbol{\mu} &= \mathbf{r}_\infty + \mathbf{K}(\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty). \end{aligned} \quad (4.7)$$

De même, $\boldsymbol{\Gamma}$ est donné par :

$$\begin{aligned} \boldsymbol{\Gamma} &= \mathbb{E} \left\{ \left[(\mathbf{r}_\infty + \mathbf{K}(\mathbf{r}_{0,B} - \mathbf{r}_\infty) + \mathbf{n}_S) - (\mathbf{r}_\infty + \mathbf{K}(\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty)) \right] \right. \\ &\quad \left. \times \left[(\mathbf{r}_\infty + \mathbf{K}(\mathbf{r}_{0,B} - \mathbf{r}_\infty) + \mathbf{n}_S) - (\mathbf{r}_\infty + \mathbf{K}(\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty)) \right]^t \right\} \\ \boldsymbol{\Gamma} &= \mathbf{K} \mathbb{E} \left[(\mathbf{r}_{0,B} - \boldsymbol{\mu}_{0,B})(\mathbf{r}_{0,B} - \boldsymbol{\mu}_{0,B})^t \right] \mathbf{K} + \mathbb{E} [\mathbf{n}_S \mathbf{n}_S^t] \\ \boldsymbol{\Gamma} &= \sigma_B^2 \mathbf{K}^2 + \sigma_S^2 \mathbf{I}_{d_L}. \end{aligned} \quad (4.8)$$

On remplace alors dans la vraisemblance 4.5 les paramètres statistiques $\boldsymbol{\mu}$ et $\boldsymbol{\Gamma}$ par leur expression 4.7 et 4.8 respectivement. Une fois que l'on connaît ces dépendances, trouver une solution analytique de $\hat{\theta}_{i,MV}(\chi_N)$ revient à trouver la valeur de θ_i qui annule la dérivée de la vraisemblance. Malheureusement, les calculs sont délicats, et nous n'avons pu résoudre cette équation pour les quatre paramètres. Par conséquent, nous nous sommes réorientés vers une optimisation numérique de la vraisemblance. Cette modélisation simplifiée de $\boldsymbol{\Gamma}$ sera utilisée par la suite pour la simulation des images.

Deuxième approche

Ici, on conserve l'expression du spectre moyen $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ de l'Equation 4.7.

En revanche, afin d'être moins restrictif sur la variabilité de \mathbf{r} , nous ne faisons plus d'hypothèse sur la forme diagonale de la matrice de covariance.

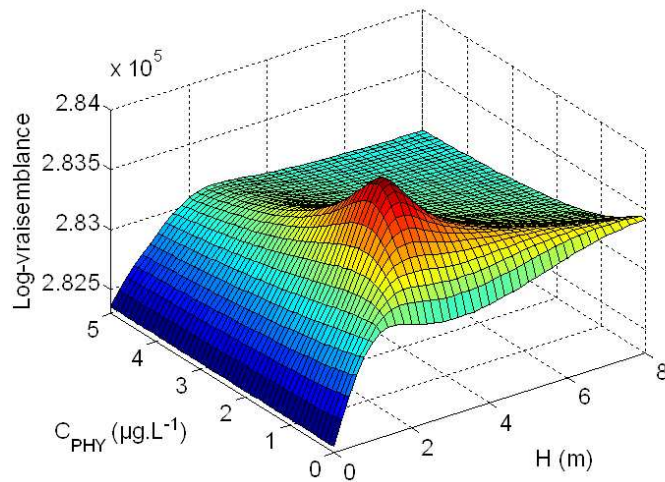
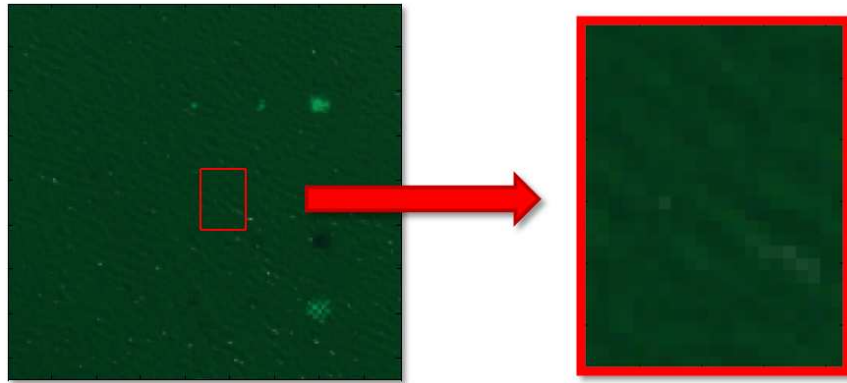


FIGURE 4.1 – Exemple de fonction de vraisemblance en fonction de H et C_{PHY} , calculée à partir d'une image HypLitt, pour une profondeur de 2.83 m.

Celle-ci est donc supposée inconnue, et est remplacée par son estimateur du maximum de vraisemblance (EMV) :

$$\hat{\Gamma}_{MV}(\chi_N; \theta) = \frac{1}{N} \sum_{i=1}^N (r_i - \mu(\theta))(r_i - \mu(\theta))^t. \quad (4.9)$$

On remplace alors Γ par $\hat{\Gamma}_{MV}(\chi_N; \theta)$ dans l'expression de la vraisemblance 4.5 et on obtient :

$$P(\chi_N | \theta) = \frac{1}{\sqrt{(2\pi)^L |\hat{\Gamma}_{MV}(\chi_N; \theta)|}} e^{-\frac{L \times N}{2}}. \quad (4.10)$$

Le vecteur de paramètres inconnu θ est alors estimé en maximisant la fonction $P(\chi_N | \theta)$, ou de façon équivalente, la fonction suivante :

$$\hat{\theta}_{MV}(\chi_N) = \underset{\theta}{\operatorname{argmax}} |\hat{\Gamma}_{MV}(\chi_N; \theta)|^{-1}. \quad (4.11)$$

De la même façon que dans l'approche simplifiée précédente, il est très délicat de retrouver les expressions analytiques des estimateurs en annulant la dérivée de cette fonction, et par conséquent, une méthode d'optimisation numérique est mise en place afin d'estimer leur valeur.

En guise d'exemple, sur la Figure 4.1, on trace la log-vraisemblance en

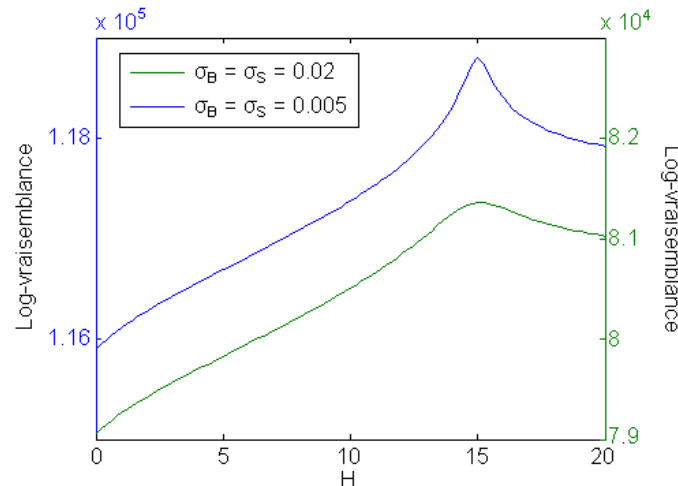


FIGURE 4.2 – Influence du bruit sur la forme de la log-vraisemblance (données synthétiques, $H = 15$ m, $C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$ et $C_{NAP} = 2.8 \text{ mg.L}^{-1}$).

fonction de H et de C_{PHY} . Celle-ci est évaluée ici sur une image HypLitt pour laquelle la profondeur est de 2.83 m.

Par ailleurs, toujours pour illustrer la forme de cette fonction, on trace également sur la Figure 4.2 son évolution en fonction de H (les autres paramètres étant supposés connus) et de la valeur des bruits σ_B^2 et σ_S^2 modélisant la variabilité du fond et du capteur. Ici, la log-vraisemblance est calculée sur des données synthétiques. On note bien l'influence du bruit, qui rend la courbe moins "piquée", donc l'estimation plus difficile à effectuer.

Sauf indication particulière, cette méthode sera utilisée par la suite pour l'estimation des paramètres.

Troisième approche semi-supervisée

Une dernière approche semi-supervisée a enfin été testée afin d'apporter une solution lorsque le fond est inconnu. En effet, la méthode évoquée dans le paragraphe précédent suppose que l'on connaît ce spectre *a priori*. En pratique, il peut être difficile de posséder cette information, et par conséquent, au même titre que la méthode d'inversion par optimisation d'une fonction d'erreur de Lee et al. (1999) présentée dans la Section 3.3.1, on peut simplement supposer connaître le spectre du fond à un facteur multiplicatif près. On a ainsi :

$$\boldsymbol{\mu} = \mathbf{r}_\infty + \mathbf{K}(f\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty) \quad (4.12)$$

où f est un paramètre multiplicatif inconnu. Typiquement, la forme du spectre $\boldsymbol{\mu}_{0,B}$ peut être semblable à celle du sable.

Ainsi, en reprenant l'expression simplifiée de la matrice de covariance diagonale de l'Equation 4.8, il est possible de retrouver l'expression analytique de l'EMV de f , en égalant à zéro la dérivée par rapport à f de la fonction de log-vraisemblance, elle-même obtenue à partir de l'Equa-

tion 4.5. On trouve alors :

$$\widehat{f}_{MV}(\chi_N; \boldsymbol{\theta}) = \frac{\boldsymbol{\mu}_{0,B}^t \mathbf{K} \boldsymbol{\Gamma}(\boldsymbol{\theta})^{-1} (\bar{\mathbf{r}}(\chi_N) + (\mathbf{K} - \mathbf{Id}_L) \mathbf{r}_\infty)}{\boldsymbol{\mu}_{0,B}^t \mathbf{K} \boldsymbol{\Gamma}(\boldsymbol{\theta})^{-1} \mathbf{K} \boldsymbol{\mu}_{0,B}} \quad (4.13)$$

où $\bar{\mathbf{r}}(\chi_N)$ est la moyenne empirique du vecteur aléatoire \mathbf{r} , calculée sur l'échantillon χ_N . Tous les détails des calculs, y compris ce qui suit, sont présentés en Annexe A.2.

Par ailleurs, en remplaçant f par l'expression 4.13 de son EMV dans l'Equation 4.5, et en faisant l'hypothèse que les variabilités du fond et du bruit de capteur sont égales, i.e. $\sigma_B^2 = \sigma_C^2 = \sigma^2$, on peut de la même façon retrouver l'expression analytique de l'EMV de σ^2 :

$$\widehat{\sigma}_{MV}^2(\chi_N; \boldsymbol{\theta}) = \frac{1}{N \times L} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})) \quad (4.14)$$

où $\mathbf{G} = \text{diag} \left[1 + e^{-4k(\lambda_i)H} \right]_{l \in \llbracket 1; L \rrbracket}$.

Dans le cas semi-supervisé, en considérant f et σ^2 inconnus, la log-vraisemblance $l(\chi_N | \boldsymbol{\theta})$ à optimiser est donc égale à :

$$\begin{aligned} l(\chi_N | \boldsymbol{\theta}) = & -\frac{N}{2} \ln \left[(2\pi)^L |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})| \right] \\ & - \frac{1}{2} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})) \end{aligned} \quad (4.15)$$

où $\boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}) = \mathbf{r}_\infty + \mathbf{K} \left(\widehat{f}_{MV}(\chi_N; \boldsymbol{\theta}) \boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty \right)$, et

$\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta}) = \widehat{\sigma}_{MV}^2(\chi_N; \boldsymbol{\theta}) (\mathbf{K}^2 + \mathbf{Id}_L)$. En développant l'expression 4.15, on obtient :

$$l(\chi_N | \boldsymbol{\theta}) = -\frac{N \times L}{2} [1 + \ln(2\pi)] - \frac{N}{2} \ln |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})|. \quad (4.16)$$

Comme précédemment, le vecteur de paramètres inconnus $\boldsymbol{\theta}$ est ensuite retrouvé par un algorithme d'optimisation numérique.

On peut remarquer qu'à une fonction croissante près, on obtient la même dépendance en $\boldsymbol{\theta}$ que dans l'Equation 4.11, à savoir que l'on doit optimiser l'inverse du déterminant de la matrice de covariance (qui est bien sûr différente dans les deux cas).

Bien qu'au même titre que la première, cette troisième approche fasse une hypothèse assez forte sur la forme de $\boldsymbol{\Gamma}$, elle est néanmoins intéressante lorsque le fond est mal connu, et dans ce cas, permet d'apporter une autre solution que la méthode présentée dans la deuxième approche.

4.2.3 Algorithme d'optimisation choisi

Devant la complexité supposée de la fonction de l'Equation 4.11, le choix de la méthode d'optimisation s'avère *a priori* important.

Dans un premier temps, afin d'éviter les maxima locaux, nous nous sommes ainsi orientés vers l'utilisation d'un algorithme d'optimisation

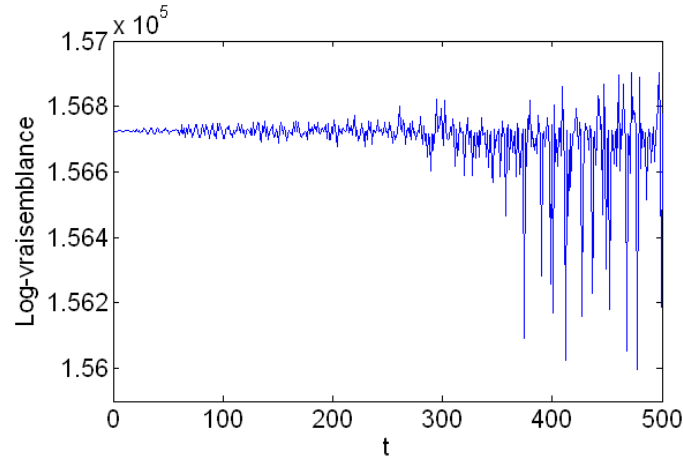


FIGURE 4.3 – Log-vraisemblance où θ est paramétré en fonction de t par la transformation d'Archimède.

global. Quelques tests ont été effectués avec la technique du recuit simulé (Kirkpatrick et al. 1983). Cependant, cette méthode présente l'inconvénient d'être peu rapide, et par conséquent, nous nous sommes par la suite intéressés aux courbes α -denses. Cette technique globale est présentée par Cherruault et Mora (2006) et consiste à transformer un problème d'optimisation à P dimensions en un problème à une dimension :

$$\operatorname{argmax}_{\theta=\{\theta_1, \dots, \theta_p\} \in \mathcal{S} \subset \mathbb{R}^P} P(\chi_N | \theta) \quad \rightarrow \quad \operatorname{argmax}_{t \in \mathcal{I} \subset \mathbb{R}} P(\chi_N | \{h_1(t), \dots, h_p(t)\}) \quad (4.17)$$

où la fonction $\mathbf{h}(t) = (h_1(t), \dots, h_p(t))$ définit une courbe α -dense de \mathbb{R} dans \mathbb{R}^P . Typiquement, cela peut être une spirale généralisée dans \mathbb{R}^P (transformation d'Archimède), où les deux premières variables θ_1 et θ_2 sont d'abord paramétrées de la façon suivante :

$$\theta_1 = \kappa t_1 \cos t_1 \quad (4.18)$$

$$\theta_2 = \kappa t_1 \sin t_1. \quad (4.19)$$

Puis, les variables t_1 et θ_3 sont également paramétrées par :

$$t_1 = \kappa t_2 \cos t_2 \quad (4.20)$$

$$\theta_3 = \kappa t_2 \sin t_2. \quad (4.21)$$

Ce processus est ainsi répété jusqu'à obtenir la paramétrisation complète de θ en fonction d'une seule variable réelle t . Chaque point θ^* de l'espace des paramètres \mathcal{S} est alors approché à $\sqrt{P}\pi\kappa$ près par le point de la courbe $\mathbf{h}(t^*) = (h_1(t^*), \dots, h_p(t^*))$. L'image de la pelote de laine est parfaite pour comprendre le principe de la méthode.

D'autres transformations permettant notamment d'introduire des bornes inférieures et supérieures sur les intervalles de variation des paramètres peuvent également être implémentées.

Nous avons donc mis en œuvre cette méthode pour l'optimisation de la vraisemblance, et sur la Figure 4.3, on trace la fonction $\ln P(\chi_N | \{h_1(t), \dots, h_p(t)\})$ en fonction de t . On se rend alors compte que celle-ci est fortement oscillante : en effet, la façon dont \mathbf{h} décrit \mathcal{S} est telle que deux points θ_1 et θ_2 proches dans \mathcal{S} , peuvent être approchés

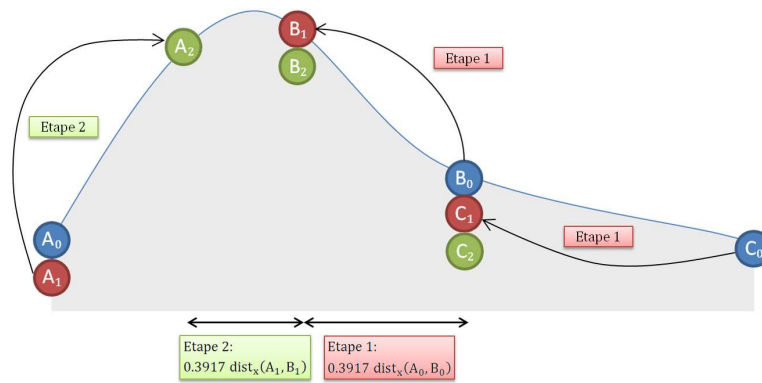


FIGURE 4.4 – Schéma d’optimisation à une dimension utilisé dans la méthode de relaxation.

par deux valeurs de t très éloignées. Bien que la forme de la vraisemblance présentée sur la Figure 4.1 soit plutôt régulière, l’optimisation globale de cette nouvelle fonction devient donc assez délicate, et il apparaît plus pertinent de conserver la structure originale à P dimensions de \mathcal{S} .

Après avoir effectué ce constat sur la forme de la fonction $P(\chi_N|\theta)$ sur plusieurs images, nous avons finalement décidé d’opter pour une technique d’optimisation simple à P dimensions, la méthode de relaxation, qui maximise alternativement chaque paramètre et détermine un maximum local. Dans chaque direction θ_i , ce dernier est retrouvé par une méthode sans gradient, très semblable à la méthode de la bisection pour trouver un zéro. Le principe est décrit sur la Figure 4.4. Il consiste à encadrer le maximum entre trois points (notés A, B, et C dans l’exemple), et de restreindre petit à petit l’écart entre ceux-ci jusqu’à ce qu’ils soient quasiment confondus au niveau du maximum. À l’initialisation, deux d’entre eux sont égaux aux bornes inférieure et supérieure de l’intervalle de variation du paramètre θ_i . A chaque itération i , on place un nouveau point X_i sur le plus grand des deux segments $[A_iB_i]$ ou $[B_iC_i]$. On peut montrer qu’il existe une manière optimale de choisir ce point : si, comme sur l’exemple à l’étape 1, le plus grand segment est $[A_iB_i]$, on le positionne à une distance telle que le rapport $\frac{\text{dist}_x(A_i, X_i)}{\text{dist}_x(X_i, B_i)}$ soit égal au nombre d’or $\frac{1+\sqrt{5}}{2}$. Selon la valeur des ordonnées, on déplace alors B_i en $B_{i+1} = X_i$ et C_i en $C_{i+1} = B_i$, ou seulement A_i en $A_{i+1} = X_i$.

La méthode de relaxation étant une technique d’optimisation locale, nous supposons donc qu’il n’y a qu’un seul maximum local (et global) dans les intervalles considérés. Ce choix est justifié si l’on se fie à la forme des vraisemblances tracées en fonction des différents paramètres (voir Figures 4.1 et 4.2).

4.2.4 Présentation des données de test

En ce qui concerne les données simulées, le fond est construit à partir d’un mélange linéaire de trois des principaux composants du sable, c’est-à-dire le quartz, le feldspath et le mica. Les spectres sont tirés de la librairie USGS développée par Clark et al. (2007). La variabilité intra-classe est modélisée par un bruit blanc gaussien de variance $\sigma_B^2 = 0.02^2$, et par

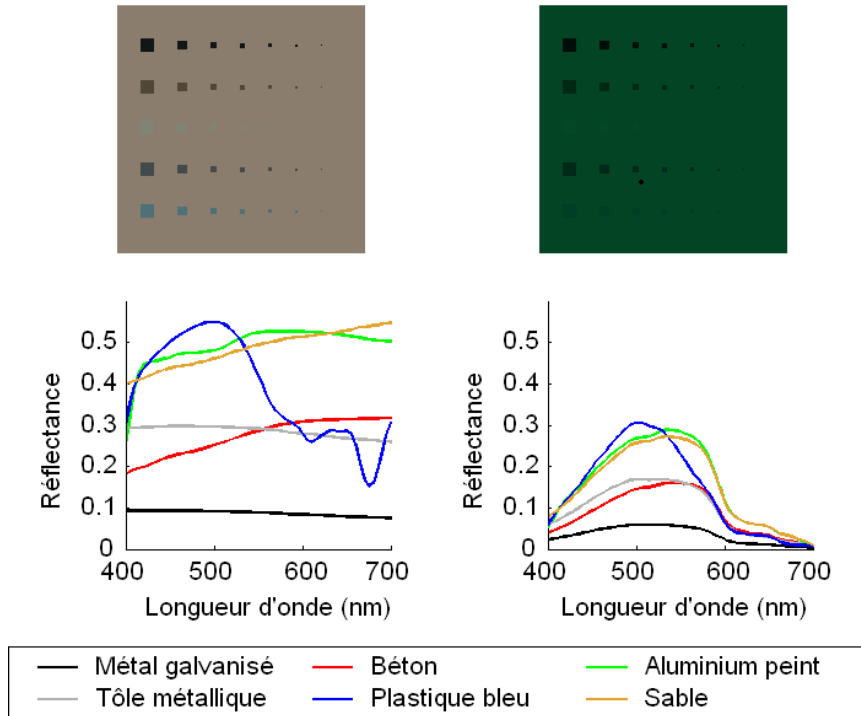


FIGURE 4.5 – Images en pseudo-vraies couleurs créées à partir de données hyperspectrales simulées, et spectres de réflectance correspondants. Colonne de gauche : fond sableux. Colonne de droite : fond sableux recouvert d'eau ($H = 3$ m, $C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$ et $C_{NAP} = 2.8 \text{ mg.L}^{-1}$).

conséquent, chaque pixel de fond est distribué de la façon suivante :

$$r_{0,B} \sim \mathcal{N}(\mu_{0,B}, \sigma_B^2 \mathbf{Id}_L) \quad (4.22)$$

Une fois atténué par la colonne d'eau, le même bruit est ajouté afin de représenter toute autre source de variabilité, telle que le bruit de capteur. Il aurait pu être intéressant d'introduire à la place un bruit de Poisson, de façon à prendre en compte le bruit de certains capteurs récents ; cependant, nous voulons aussi décrire la variabilité spectrale naturelle pouvant être due par exemple à une qualité d'eau inhomogène et qui peut être bien représentée par une distribution gaussienne. Cette modification n'améliorerait donc pas forcément les résultats.

En guise d'exemple, on montre sur la Figure 4.5, deux images simulées représentant un fond sableux recouvert ou non d'une couche d'eau turbide de 3 m de hauteur (à noter qu'ici, les concentrations d'une eau dite turbide sont tirées de l'article de Brando et al. (2009), c'est-à-dire $C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$ et $C_{NAP} = 2.8 \text{ mg.L}^{-1}$). Ici, différentes cibles ont été disposées sur le sable.

La gamme de longueurs d'onde est fixée à 400-700 nm, avec une résolution spectrale de 5 nm (61 bandes spectrales). L'échantillon d'apprentissage est un cube de 21×21 pixels (voir Section 5.4). Le rapport signal à bruit (SNR) en décibels est donné par : $\text{SNR} = 10 \times \ln \left(\frac{\sum_i r_i^2}{\sum_i n_i^2} \right)$ dans lequel n_i est le bruit simulé dans le pixel r_i .

En ce qui concerne les données HypLitt, nous avons testé notre mé-

	H = 0.1m	H = 5m	H = 10m	H = 20m	H = 30m
H	1.18	0.33	0.50	1.62	26.81
C_{PHY}	32.53	2.95	6.66	18.76	19.02
C_{CDOM}	9.93	1.24	3.87	10.69	10.66
C_{NAP}	34.94	2.97	3.56	3.77	3.35

TABLE 4.1 – Résultats de l'estimation par MV dans le cas covariance, profondeur et concentrations inconnues (données simulées) : REQMs relatives (%) en fonction de la profondeur (eau turbide) et calculées sur 500 tirages aléatoires (modèle M1).

	SNR = 1dB	SNR = 5dB	SNR = 10dB	SNR = 20dB
H	3.00	1.18	0.76	0.35
C_{PHY}	51.65	17.44	10.77	5.81
C_{CDOM}	27.26	10.84	6.93	3.63
C_{NAP}	13.46	5.54	3.55	1.77

TABLE 4.2 – Résultats de l'estimation par MV dans le cas covariance, profondeur et concentrations inconnues (données simulées) : REQMs relatives (%) en fonction du SNR ($H = 14$ m, eau turbide) et calculées sur 100 tirages aléatoires (modèle M1).

thode sur les images présentées sur la Figure 1.5, et dont la profondeur varie de 0.20 m à 21.80 m. Il convient de remarquer que, contrairement aux simulations, les fonds ne sont pas parfaitement plats, et ces petites variations de hauteur d'eau au sein de chaque image, augmentent donc l'erreur calculée (notamment pour les profondeurs les plus faibles où la variation relative est la plus forte). Par ailleurs, la composition de l'eau n'ayant pas pu être mesurée sur l'ensemble des données étudiées, la vérité terrain est donc constituée dans la Section 4.2.5 suivante des spectres de sable (mouillé et sec) et de cible, et de la profondeur. Les résultats d'estimation du MV concernant les autres paramètres sont présentés dans la Section 4.4.

Nos résultats d'estimation sont comparés avec ceux obtenus par la méthode présentée dans la Section 3.4. Pour ce faire, l'implémentation Matlab® de l'algorithme de Levenberg-Marquardt est utilisée.

4.2.5 Résultats

Les résultats dépendent des connaissances *a priori* que l'on possède sur les paramètres, et ainsi, la précision est différente selon que l'on estime la covariance et la profondeur, ou la covariance, la profondeur et les concentrations. Dans nos simulations, on se place dans le cas le plus difficile, à savoir covariance, profondeur et concentrations inconnues.

Les résultats sont obtenus avec un échantillon $\chi_N = \{r_1, \dots, r_N\}$ de pixels de fond. En pratique, cela semble tout à fait plausible d'avoir accès à cet ensemble puisque, pour les images HypLitt, cela revient à considérer une zone de 90 m² autour du pixel central. En zone côtière, la profondeur ne varie pas beaucoup sur de telles surfaces. Malgré tout, si besoin est, il est toujours possible de réduire la taille de χ_N sans dégrader considérablement les résultats.

	H = 1m		H = 5m		H = 20m	
	EMV _{σ}	EMV _{$\hat{\sigma}, \hat{f}$}	EMV _{σ}	EMV _{$\hat{\sigma}, \hat{f}$}	EMV _{σ}	EMV _{$\hat{\sigma}, \hat{f}$}
H	22.71	0.22	4.28	0.94	5.91	34.36
C_{PHY}	99.99	7.71	17.07	8.15	39.70	38.88
C_{CDOM}	28.17	6.34	8.72	2.13	12.71	12.81
C_{NAP}	99.99	35.70	70.51	12.62	3.87	4.96

TABLE 4.3 – Résultats de l'estimation par MV dans le cas où le fond supposé connu est en réalité égal à 80% du fond réel (données simulées) : REQMs relatives (%) avec (EMV _{$\hat{\sigma}, \hat{f}$}) ou sans (EMV _{σ}) estimation de f et de σ^2 , en fonction de la profondeur (eau turbide) et calculées sur 100 tirages aléatoires (modèle M2).

Images simulées

Les résultats sont présentés dans le Tableau 4.1 : on montre l'évolution de la racine de l'erreur quadratique moyenne (REQM) sur l'estimation des paramètres H , C_{PHY} , C_{CDOM} , et C_{NAP} , en fonction de la profondeur. L'EQM sur l'estimation de Y est donnée par : $EQM_Y = \mathbb{E} [(\hat{Y} - Y)^2]$.

Excepté en zone très peu profonde (0.1 m), dans lesquelles les constituants de l'eau n'ont que peu d'influence sur la réflectance subsurface, on voit sur les données simulées, que plus la hauteur d'eau est importante, et moins l'estimation de H , C_{PHY} et C_{CDOM} est précise. Seule la précision sur C_{NAP} semble rester constante dans cette gamme de profondeurs. En zone "optiquement profonde", on a $r \approx r_\infty$. Ainsi, r ne dépend quasiment plus de H et l'estimation de la profondeur devient très difficile. En revanche, la précision sur l'estimation de C_{PHY} , C_{CDOM} et C_{NAP} ne diminue plus au-delà de 20-30m, car la réflectance reste égale à r_∞ .

Par ailleurs, on montre également la robustesse au bruit de notre méthode sur le Tableau 4.2 pour $H = 14$ m. D'un point de vue général, on voit que les erreurs sont acceptables, même pour des niveaux de bruit élevés. Comme dans le Tableau 4.1, on remarque que C_{PHY} est le paramètre le plus sensible, tandis que H reste bien estimé, et ce même pour SNR = 1 dB.

Enfin, on observe dans le Tableau 4.3 l'influence d'une mauvaise connaissance *a priori* du spectre du fond. Ici, les résultats sont obtenus en utilisant la première approche simplifiée (EMV _{σ}) et la troisième approche semi-supervisée (EMV _{$\hat{\sigma}, \hat{f}$}) présentées dans la Section 4.2.2. Le fond supposé connu est en réalité égal à 80% du fond réel. On voit bien que pour des profondeurs faibles et intermédiaires, une mauvaise connaissance du fond entraîne une mauvaise estimation des paramètres. Dans ce cas, l'estimation du facteur multiplicatif f permet de retrouver le fond réel, et donc de conserver de bonnes performances sur les autres paramètres. En revanche, pour de fortes profondeurs, le fond n'a quasiment plus d'influence sur la réflectance subsurface, et l'erreur sur l'estimation f devient supérieure à l'erreur initiale de 20% sur la connaissance du spectre du fond. Dans ce cas, il n'est pas nécessaire d'estimer f puisqu'en outre, on a $r \approx r_\infty$.

	H = 0.20m		H = 1.12m		H = 2.83m		H = 4.70m	
	M1	M2	M1	M2	M1	M2	M1	M2
EMV	11.4	7.3	18.1	12.2	16.2	6.7	21.1	11.4
LEV	12.6	16.1	26.1	19.4	29.0	20.6	34.7	28.7
	H = 6.70m		H = 11.80m		H = 16.80m		H = 21.80m	
	M1	M2	M1	M2	M1	M2	M1	M2
EMV	22.1	16.0	26.0	31.2	392.8	47.2	229.0	55.1
LEV	12.9	14.4	54.4	63.5	453.0	67.1	244.1	25.5

TABLE 4.4 – Résultats de l'estimation de H par MV dans le cas covariance, profondeur et concentrations inconnues (données HypLitt) : REQMs relatives (%) en fonction de la profondeur, pour les deux méthodes et les deux modèles M1 et M2.

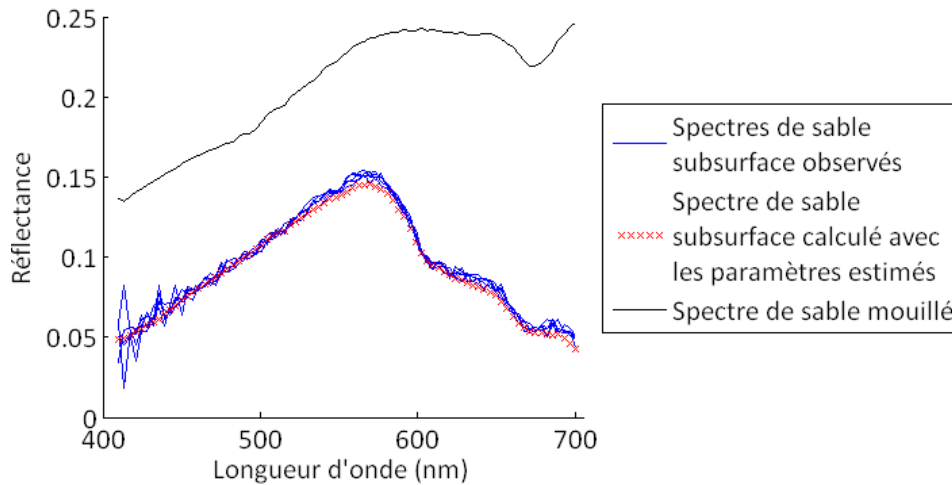


FIGURE 4.6 – Comparaison entre les spectres subsurfaces réels et le spectre simulé à partir des paramètres estimés et du spectre du fond (H = 1.12 m ; REQM = 12.2%)

Images HypLitt

En ce qui concerne les données HypLitt, nous considérons deux sous-images pour chaque profondeur, dont la taille varie entre 200 et 1000 pixels. Dans le Tableau 4.4, on montre les REQMs relatives (calculées sur ces deux images) pour la méthode d'estimation du MV (EMV), et pour la méthode classique d'inversion par minimisation d'une fonction d'erreur présentée dans la Section 3.4 (par abus de langage, on nommera cette technique LEVMAR), et ce pour les deux modèles M1 et M2. Tout d'abord, on peut remarquer que l'EMV permet de mieux estimer la profondeur que LEVMAR. Par exemple, pour $H = 2.83$ m et en utilisant le modèle M1, l'erreur de l'EMV est de 16.2% tandis qu'elle est de 29.0% avec LEVMAR. Pour des profondeurs allant jusqu'à 11.80 m, l'erreur d'estimation avec l'EMV reste inférieure à 26% (modèle M1). De plus, l'utilisation d'un modèle plus réaliste semble avoir un impact assez important, notamment pour l'EMV, et représente donc un bon moyen d'améliorer facilement les résultats. Par ailleurs, plusieurs raisons peuvent expliquer que les erreurs d'estimation soient supérieures à celles présentées dans le Tableau 4.1 : les inhomogénéités de type de fond, de profondeur ou de turbidité, ou encore des erreurs dues au modèle peuvent en être à l'origine.

En guise d'exemple, on compare sur la Figure 4.6 le spectre de sable subsurface simulé (obtenu à l'aide du spectre de sable corrigé avec les

	H = 0.20m			H = 4.70m		
	EMV $_{\sigma}$	EMV	EMV $_{\hat{\sigma},\hat{f}}$	EMV $_{\sigma}$	EMV	EMV $_{\hat{\sigma},\hat{f}}$
$\mu_B = \mu_{B,reel}$	23.00	12.72	19.83	26.04	9.24	18.73
$\mu_B \neq \mu_{B,reel}$	30.50	122.10	19.83	33.08	13.61	18.73

TABLE 4.5 – Résultats de l'estimation de H par MV pour les trois approches présentées en Section 4.2.2 (données HypLitt) : REQMs relatives (%) en fonction de la profondeur et de la connaissance a priori du fond (modèle M2). Pour $H = 0.20$ m, on a $\mu_B = 130\% \times \mu_{B,reel}$, tandis que pour $H = 4.70$ m, on a $\mu_B = 80\% \times \mu_{B,reel}$.

paramètres estimés avec l'EMV) avec un ensemble de spectres observés : on voit que l'ajustement est bien réalisé.

Enfin, comme sur les images simulées, on étudie l'influence d'une mauvaise connaissance a priori du fond sur les résultats d'estimation. On présente ces derniers dans le Tableau 4.5 en fonction de la méthode d'estimation utilisée, de la profondeur, et de la connaissance du fond. On constate tout d'abord que, lorsque le fond est connu, la méthode EMV présentée dans la deuxième approche (voir Section 4.2.2) et qui n'impose pas de matrice de covariance diagonale, donne de meilleurs résultats que les méthodes EMV $_{\sigma}$ (première approche simplifiée) et EMV $_{\hat{\sigma},\hat{f}}$ (troisième approche semi-supervisée), qui imposent justement cette contrainte. En revanche, lorsque le fond n'est connu qu'à un facteur multiplicatif près, les performances d'estimation des méthodes EMV et EMV $_{\sigma}$ sont dégradées, tandis que la méthode EMV $_{\hat{\sigma},\hat{f}}$, qui estime le fond simultanément avec les autres paramètres caractéristiques de la colonne d'eau, donne toujours les mêmes résultats. Par ailleurs, on voit que pour la méthode EMV, cette dégradation est beaucoup moins marquée lorsque la couche d'eau est plus importante, puisque dans ce cas-là, le fond a moins d'influence sur la réflectance subsurface.

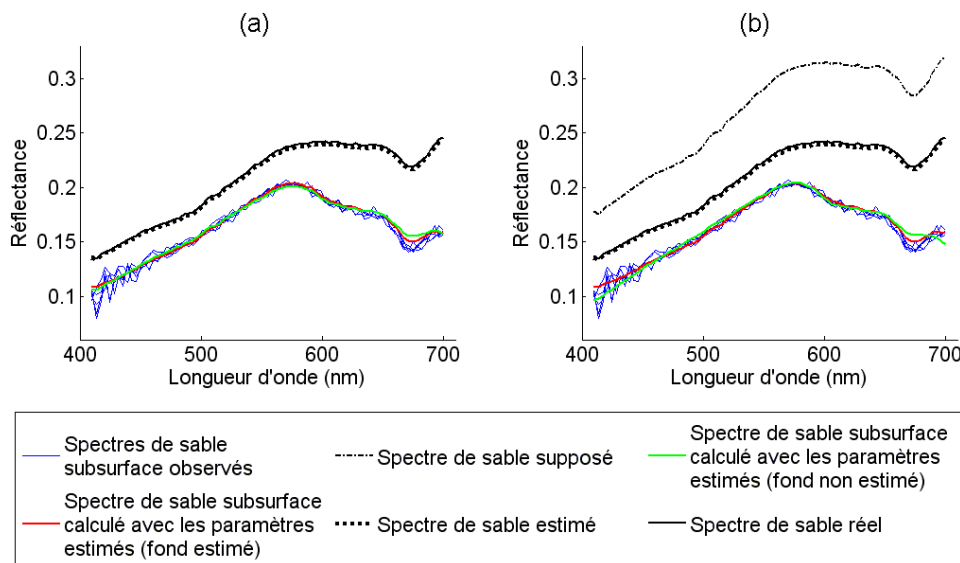


FIGURE 4.7 – Comparaison entre les spectres subsurfaces réels et les spectres simulés à partir des paramètres estimés avec (EMV $_{\hat{\sigma},\hat{f}}$) ou sans (EMV $_{\sigma}$) estimation de f et σ^2 ($H = 0.20$ m). Pour (a), le fond supposé est égal au spectre de fond réel, tandis que pour (b), il est égal à 130% du spectre réel.

Pour illustrer cela, on compare sur la Figure 4.7 les spectres de sable subsurface réels avec les spectres de sable subsurface reconstruits à partir des paramètres estimés par les méthodes EMV_{σ} et $EMV_{\hat{\sigma}, \hat{f}}$, et ce, lorsque le fond est connu ou non à un facteur multiplicatif près ($H = 0.20$ m). On constate que dans les deux cas, le fond estimé par la méthode $EMV_{\hat{\sigma}, \hat{f}}$ est identique et très proche du fond réel. Lorsque le fond supposé est différent du spectre réel, la méthode $EMV_{\hat{\sigma}, \hat{f}}$ conserve les mêmes résultats comme observé dans le Tableau 4.5, tandis que les performances de la méthode EMV_{σ} sont dégradées (en l'occurrence, la profondeur est surestimée).

Ainsi, l'ensemble de ces résultats nous montre que la méthode EMV fonctionne bien : cependant, en eau très peu profonde, l'atténuation étant principalement causée par l'eau pure, il est difficile d'estimer convenablement les concentrations des divers constituants. En eau très profonde, le fond n'a quasiment plus aucune influence sur la réflectance, et la précision sur l'estimation de la hauteur d'eau diminue. La gamme de profondeurs pour laquelle l'EMV donne les meilleurs résultats dépend de plusieurs facteurs, tels que la turbidité de l'eau, le niveau de bruit ou encore la taille de l'échantillon.

De plus, en l'absence d'une connaissance parfaite du spectre du fond, une approche semi-supervisée reste opérationnelle en faible et moyenne profondeurs.

4.3 MODÉLISATION D'UNE PROFONDEUR VARIABLE

4.3.1 Introduction

Dans la méthode EMV, nous ne prenons pas en compte les variations des différents paramètres H , C_{PHY} , C_{CDOM} et C_{NAP} du modèle bathymétrique au sein de l'image. Cette hypothèse reste valable dans une bonne partie des cas, lorsque l'on travaille sur une petite zone géographique en zone côtière. Si on se place dans les mêmes conditions que pour la détection, nous avons besoin d'un nombre de pixels environ égal à sept fois le nombre de bandes (voir Section 5.4), et cela revient donc à considérer des régions de 90 m^2 pour les images HypLitt à la résolution spatiale de 40 cm avec un nombre de bandes égal à 81. Sur de telles zones, on peut supposer dans un premier temps que les paramètres sont constants.

Néanmoins, afin d'affiner le processus d'estimation, il serait intéressant de pouvoir modéliser ces variations, notamment celles de la profondeur. En effet, ce paramètre est *a priori* celui qui peut évoluer le plus facilement au sein d'une image, la qualité de la couche d'eau pouvant davantage être considérée comme homogène sur une petite zone. Nous avons donc pris en compte ces variations, en les modélisant sous la forme d'un ou deux plans.

Au-delà d'améliorer bien évidemment la précision sur l'estimation de H , nous montrons dans cette partie que cela améliore également la précision sur l'estimation des concentrations pour des zones de faibles fonds. En revanche, nous verrons que pour des profondeurs plus importantes, la variation relative du fond étant plus faible, la scène peut être représentée

par un fond plat sans que cela ne nuise à la précision de l'estimation des autres paramètres.

En eau peu profonde, la modélisation de la pente a donc un double intérêt, à savoir d'augmenter la qualité de l'estimation de la hauteur ET des concentrations.

Ici, cette méthode est évaluée sur des simulations et des images réelles HypLitt.

4.3.2 Méthode d'estimation à une pente

Méthode

On utilise ici le modèle bathymétrique de réflectance M2 mis sous forme matricielle dans la Section 4.2.1. Lors du processus classique d'EMV, le spectre moyen et la matrice de covariance du fond sont supposés identiques sur toute l'image. En revanche, cela n'est plus vrai dans le cas d'un fond variant linéairement, puisque ces éléments dépendent entre autres, du paramètre H : on a donc autant de spectres moyens et de matrices de covariance qu'il y a de pixels.

Pour tout pixel (i,j) , $\boldsymbol{\mu} = \boldsymbol{\mu}_{ij}$, et $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{ij}$.

Le spectre moyen s'écrit tel que :

$$\boldsymbol{\mu}_{ij} = \mathbf{r}_\infty + \mathbf{K}_{B,ij}\boldsymbol{\mu}_{0,B} - \mathbf{K}_{C,ij}\mathbf{r}_\infty \quad (4.23)$$

où $\boldsymbol{\mu}_{0,B}$ représente le spectre moyen du fond (sans eau), et $\mathbf{K}_{B,ij}$ (resp. $\mathbf{K}_{C,ij}$) la matrice d'atténuation des photons qui ont interagi (resp. qui n'ont pas interagi) avec le fond et qui arrivent sur le pixel (i,j) .

La matrice de covariance est définie par :

$$\boldsymbol{\Gamma}_{ij} = \mathbb{E} \left[\left(\mathbf{r}_{ij} - \boldsymbol{\mu}_{ij} \right) \left(\mathbf{r}_{ij} - \boldsymbol{\mu}_{ij} \right)^t \right] \quad (4.24)$$

En utilisant le modèle bathymétrique de réflectance, on exprime le vecteur aléatoire \mathbf{r}_{ij} en fonction des paramètres \mathbf{r}_∞ , $\mathbf{K}_{B,ij}$, $\mathbf{K}_{C,ij}$, et d'un vecteur aléatoire \mathbf{r}_B représentant la réflectance du fond. On peut alors écrire :

$$\begin{aligned} \boldsymbol{\Gamma}_{ij} &= \mathbb{E} \left[\left(\left(\mathbf{r}_\infty + \mathbf{K}_{B,ij}\mathbf{r}_{0,B} - \mathbf{K}_{C,ij}\mathbf{r}_\infty \right) - \left(\mathbf{r}_\infty + \mathbf{K}_{B,ij}\boldsymbol{\mu}_{0,B} - \mathbf{K}_{C,ij}\mathbf{r}_\infty \right) \right) \right. \\ &\quad \left. \times \left(\left(\mathbf{r}_\infty + \mathbf{K}_{B,ij}\mathbf{r}_{0,B} - \mathbf{K}_{C,ij}\mathbf{r}_\infty \right) - \left(\mathbf{r}_\infty + \mathbf{K}_{B,ij}\boldsymbol{\mu}_{0,B} - \mathbf{K}_{C,ij}\mathbf{r}_\infty \right) \right)^t \right] \\ \boldsymbol{\Gamma}_{ij} &= \mathbb{E} \left[\mathbf{K}_{B,ij} \left(\mathbf{r}_{0,B} - \boldsymbol{\mu}_{0,B} \right) \left(\mathbf{r}_{0,B} - \boldsymbol{\mu}_{0,B} \right)^t \mathbf{K}_{B,ij} \right] \\ \boldsymbol{\Gamma}_{ij} &= \mathbf{K}_{B,ij}\boldsymbol{\Gamma}_0\mathbf{K}_{B,ij} \end{aligned} \quad (4.25)$$

où $\boldsymbol{\Gamma}_0$ désigne la matrice de covariance du fond qui est totalement indépendante des paramètres H , C_{PHY} , C_{CDOM} et C_{NAP} .

Les deux matrices $\mathbf{K}_{B,ij}$ et $\mathbf{K}_{C,ij}$ deviennent :

$$\mathbf{K}_{B,ij} = \text{diag} \left[e^{-(k_d(\lambda_l) + k_u^b(\lambda_l))H_{ij}} \right]_{l \in \llbracket 1;L \rrbracket} \quad (4.26)$$

$$\mathbf{K}_{C,ij} = \text{diag} \left[e^{-(k_d(\lambda_l) + k_u^c(\lambda_l))H_{ij}} \right]_{l \in \llbracket 1;L \rrbracket} \quad (4.27)$$

Pente	$H_0 = 0.10 \text{ m}$			$H_0 = 3 \text{ m}$		
	SP	AP	AP/SP	SP	AP	AP/SP
0%	0.56	0.72	1.29	0.13	0.14	1.08
5%	68.46	25.01	0.37	3.96	1.83	0.46
10%	115	63.40	0.55	7.47	5.47	0.73
20%	274	221	0.81	13.44	11.58	0.86
30%	385	333	0.86	18.40	16.98	0.92
Pente	$H_0 = 7 \text{ m}$					
	SP	AP	AP/SP			
0%	0.29	0.30	1.03			
5%	1.79	1.06	0.59			
10%	3.44	2.66	0.77			
20%	6.51	5.99	0.92			
30%	9.30	8.92	0.96			

TABLE 4.6 – REQMs (en %) sur l'estimation de la hauteur en fonction de la pente. Dans le cas AP (AP : Avec Pente), on estime cette pente, tandis que le fond est supposé plat dans le cas SP (SP : Sans Pente).

où H_{ij} désigne la profondeur sur le pixel (i,j) .

Dans un premier temps, on peut modéliser la carte de profondeur $(H_{ij})_{1 < i < N_i, 1 < j < N_j}$ par un plan paramétré par le triplet $\{A, B, H_0\}$, où la hauteur en chaque pixel (i,j) est décrit par :

$$H_{ij} = A \times i + B \times j + H_0. \quad (4.28)$$

Si l'on reprend le processus d'estimation, la nouvelle vraisemblance s'exprime alors comme :

$$P(\chi_N | \theta) = P(r_{11}, \dots, r_{N_i N_j}) = \prod_{i,j}^{N_i N_j} \frac{1}{\sqrt{(2\pi)^L |\Gamma_{ij}|}} e^{-\frac{1}{2} (r_{ij} - \mu_{ij})^t \Gamma_{ij}^{-1} (r_{ij} - \mu_{ij})} \quad (4.29)$$

où $N = N_i \times N_j$, et $\theta = \{A, B, H_0, C_{PHY}, C_{CDOM}, C_{NAP}\}$.

En pratique, la matrice de covariance du fond est estimée au préalable à partir de la relation $\hat{\Gamma}_0 = \mathbf{K}_B^{-1} \hat{\Gamma} \mathbf{K}_B^{-1}$, où la matrice $\hat{\Gamma}$ et les paramètres intervenant dans l'expression de \mathbf{K}_B sont estimés en surface par MV sur une petite zone de l'image en supposant le fond plat.

Résultats

Nous avons d'abord testé cette méthode sur les simulations. Ici, chaque réflectance au pixel (i,j) est simulée de la façon suivante :

$$r_{ij} = r_\infty + \mathbf{K}_{B,ij}(\mu_{0,B} + \mathbf{n}_B) - \mathbf{K}_{C,ij} r_\infty + \mathbf{n}_S \quad (4.30)$$

où \mathbf{n}_B et \mathbf{n}_S sont des réalisations de deux vecteurs aléatoires représentant des bruits blancs gaussiens de variance σ_B^2 et σ_S^2 qui modélisent la variabilité spectrale du fond et du capteur respectivement. Ici, on prend $\sigma_B^2 = \sigma_S^2 = 0.01^2$.

Sur le Tableau 4.6, on montre l'évolution de la racine de l'erreur quadratique moyenne relative en fonction de la pente, de la profondeur H_0 , et de

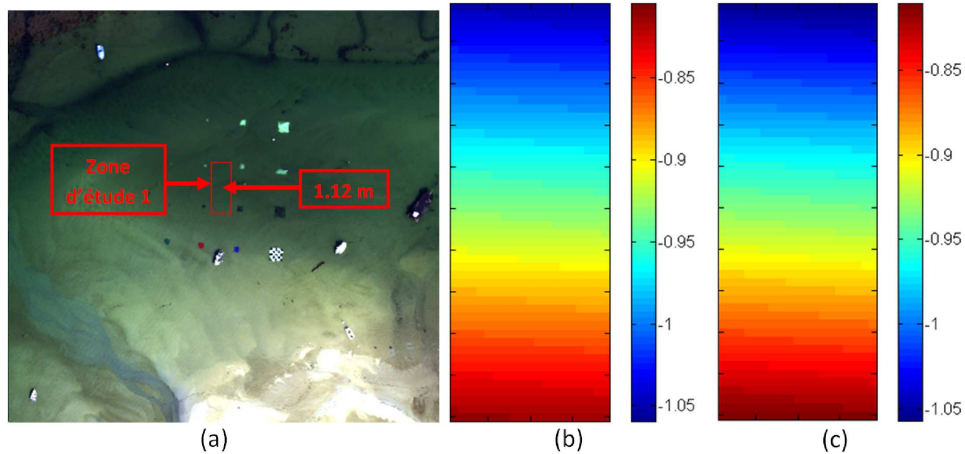


FIGURE 4.8 – (a) Image *H1-0500m-Lo1-090deg-20100917-10h20mn-atm* (zone 1), (b) relief estimé avec le modèle de Brando et al. (2009) et Bricaud et al. (1995), (c) relief estimé avec le modèle de Lee (1994).

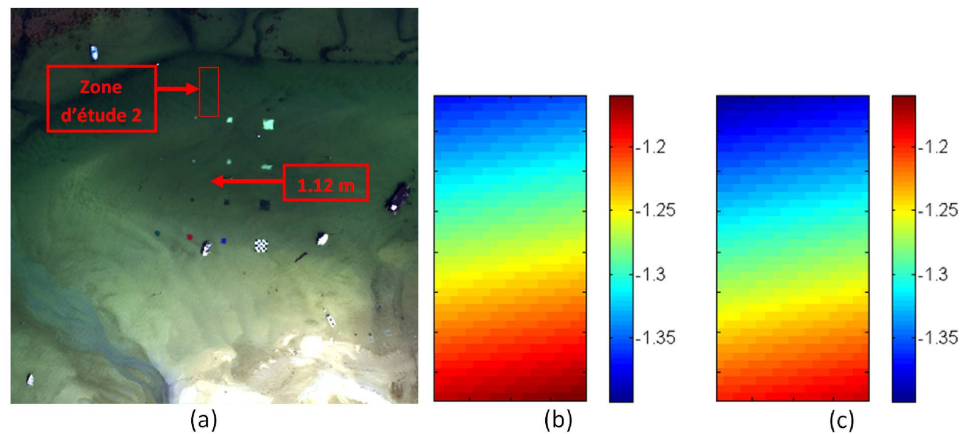


FIGURE 4.9 – (a) Image *H1-0500m-Lo1-090deg-20100917-10h20mn-atm* (zone 2), (b) relief estimé avec le modèle de Brando et al. (2009) et Bricaud et al. (1995), (c) relief estimé avec le modèle de Lee (1994).

la méthode utilisée. Globalement, on constate que la précision augmente lorsque l'on modélise cette pente. Toutefois, des différences sont à noter selon les cas. Tout d'abord, on remarque que pour un fond plat, l'ajout de deux inconnues supplémentaires engendre une hausse de l'erreur, hausse qui diminue cependant lorsque H augmente (30% pour $H_0 = 0.10$ m, 8% pour $H_0 = 3$ m, et 3% pour $H_0 = 7$ m). Dans un deuxième temps, on voit que pour des pentes intermédiaires comprises entre 5% et 10%, l'augmentation de la précision est importante puisqu'elle est comprise entre 23% et 63% selon la profondeur de base H_0 . Pour des pentes plus grandes ($\geq 20\%$), la différence de précision lorsque l'on modélise ou non la pente diminue car lorsque la profondeur augmente, il est plus difficile d'estimer convenablement les variations du relief. En effet, il y a beaucoup plus de similitudes entre deux spectres de réflectance d'une colonne d'eau de 7 m et 8 m, qu'entre deux réflectances d'une colonne d'eau de 0.10 m et 1.10 m. Le gain apporté lorsque l'on modélise la pente est donc plus important pour les faibles profondeurs (et *a fortiori* les faibles pentes).

Nous avons également testé notre méthode sur les données HypLitt.

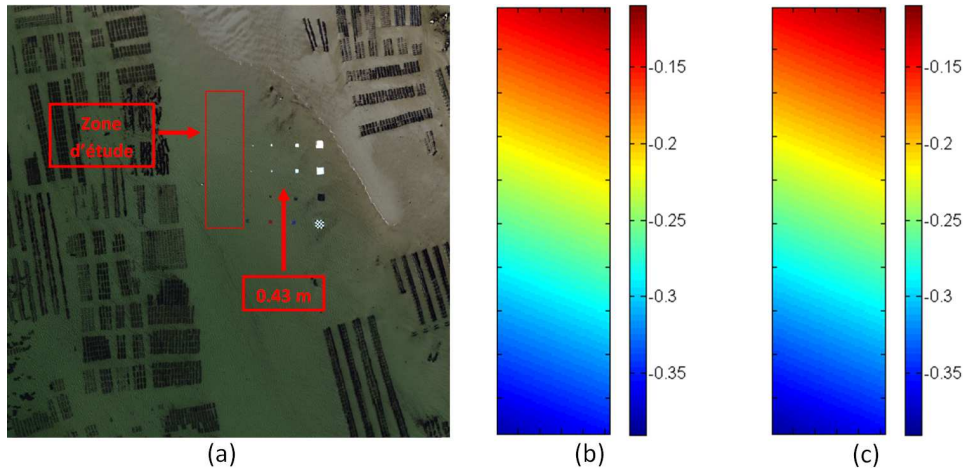


FIGURE 4.10 – (a) Image H2-0500m-Lo1-090deg-20100914-15h09mn-atm, (b) relief estimé avec le modèle de Brando et al. (2009) et Bricaud et al. (1995), (c) relief estimé avec le modèle de Lee (1994).

Dans ce cas, s'il est difficile d'avoir un calcul d'erreur fiable puisque nous ne connaissons la profondeur qu'en un seul point, il est facile d'évaluer visuellement la justesse de notre estimation, en observant l'image.

Ici, nous présentons sur les Figures 4.8, 4.9, 4.10, et 4.11 deux séries de résultats, obtenus avec les deux modèles d'absorption du phytoplancton présentés dans la Section 2.3.1 (celui de Brando et al. (2009) avec l'absorption spécifique de Bricaud et al. (1995), et celui de Lee (1994)). Globalement, on voit que la hauteur est bien estimée : la direction de la pente est bonne sur les quatre images, et la correspondance avec le point connu est plutôt réussie. Par exemple, pour la première image sélectionnée du site S_1 (voir Figure 4.8), la vraie hauteur au centre des cibles est de 1.12 m, alors que la hauteur estimée en ce même point est d'environ 0.95 m pour le modèle d'absorption du phytoplancton de Brando et al. (2009) et Bricaud et al. (1995), et de 0.98 m pour le modèle de Lee (1994). Pour l'image du site S_3 (voir Figure 4.11), la vraie hauteur dans le coin supérieur droit est de 4.70 m, tandis qu'elle est de 4.15 m pour le modèle de Brando et al. (2009) et Bricaud et al. (1995), et de 4.50 m pour le modèle de Lee (1994).

Dans l'ensemble, l'estimation réalisée avec le modèle de Lee (1994) paraît meilleure qu'avec le modèle de Brando et al. (2009) et Bricaud et al. (1995)

On voit donc que cette modélisation de la hauteur d'eau permet d'améliorer notablement l'estimation de H . Or, on sait que cette dernière est très liée à celle des trois autres concentrations : une augmentation de la précision sur H se traduira donc par une augmentation de la précision sur C_{PHY} , C_{CDOM} , et C_{NAP} . On peut d'ailleurs retrouver cela à travers les résultats de simulation présentés sur la Figure 4.12 : on voit bien que pour des zones de faibles fonds, la modélisation de la pente améliore également la précision sur l'estimation des concentrations. Sur le graphe de gauche, on voit par exemple, que pour une pente de 10%, l'erreur sur l'estimation de C_{PHY} passe de 13% à 7%, tandis que l'erreur sur C_{NAP} passe de 28% à 7%. En revanche, pour des profondeurs plus importantes (graphe de droite), la variation relative de la hauteur est plus faible : en guise

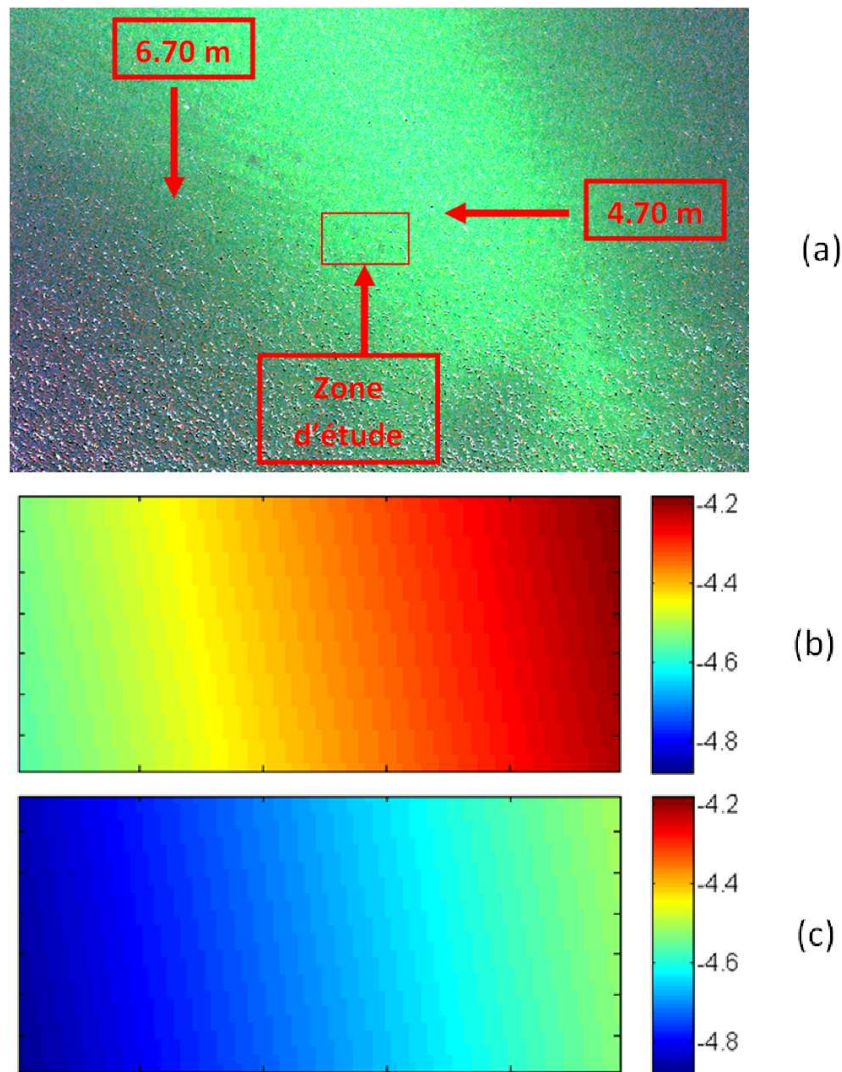


FIGURE 4.11 – (a) Image $H_3-0500m-Lo1-270deg-20100914-13h15mn-atm-1-deglintage-shoreline$, (b) relief estimé avec le modèle de Brando et al. (2009) et Bricaud et al. (1995), (c) relief estimé avec le modèle de Lee (1994).

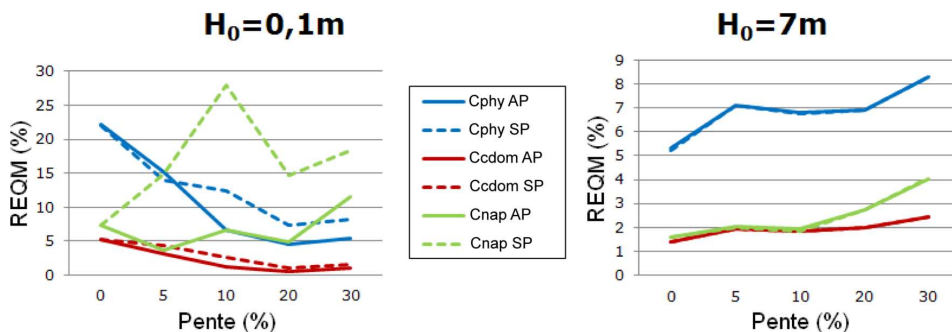


FIGURE 4.12 – REQMs calculées à partir de 100 tirages aléatoires sur l'estimation des concentrations (en %) en fonction la pente (AP : Avec Pente, la pente est modélisée ; SP : Sans Pente, le fond est supposé plat). A gauche, le point le moins profond se situe à 0.10 m, tandis qu'à droite, il se situe à 7 m.

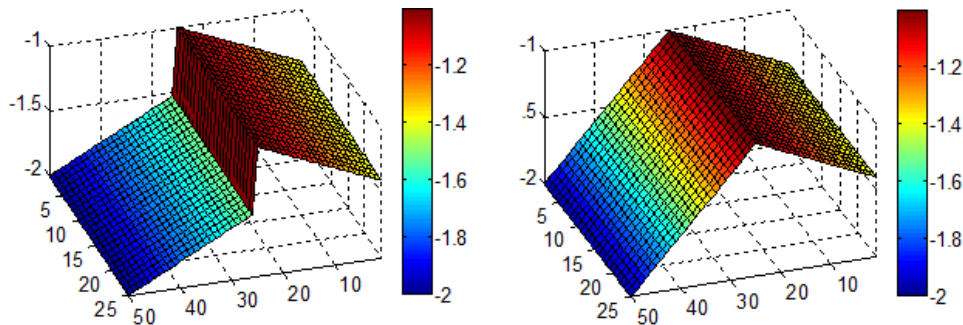
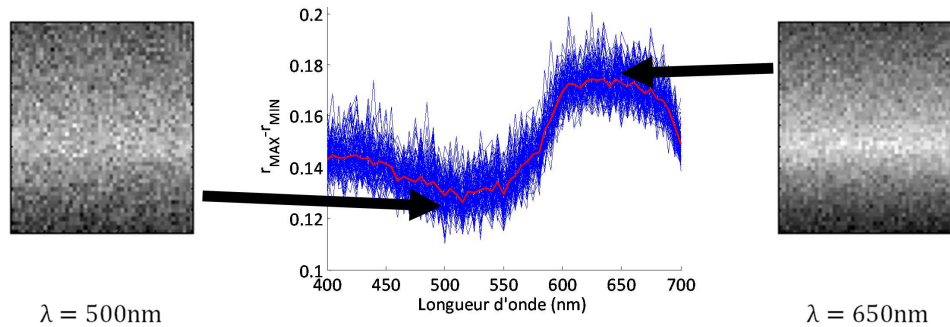


FIGURE 4.13 – Variations discontinues ou continues de la hauteur.

FIGURE 4.14 – Choix de la bande permettant d'obtenir la dynamique $r_{MAX}(\lambda) - r_{MIN}(\lambda)$ maximale. La profondeur sur la crête centrale du fond est de 1 m, tandis que la pente dans la partie supérieure (resp. inférieure) est de 5% (resp. 10%).

d'exemple, une pente de 10% entraîne un écart maximal par rapport à la valeur moyenne de H de 6% pour $H_0 = 7$ m contre 80% pour $H_0 = 0.1$ m. Dans ce cas, la représentation de la scène par un fond plat suffit alors pour avoir de bons résultats.

4.3.3 Méthode d'estimation à deux pentes

Méthode

Une fois que l'on sait estimer une profondeur variant linéairement dans l'image, on peut étendre cette méthode à plusieurs plans. Par la suite, on présente les travaux effectués pour deux plans, mais on pourrait très bien envisager de réaliser une autre étude, pour trois plans ou plus.

La technique développée ici se décompose en deux étapes : durant la première, on localise la zone de rupture de pente (crête ou vallée), puis, on estime de façon indépendante ou non la hauteur dans les deux régions extraites (voir Figure 4.13). Ici, le terme de rupture de pente correspond à un changement de signe de la dérivée. Dans ces travaux, on modélise cette zone par une droite dans le plan de l'image.

Lors de la première étape, on cherche donc à repérer cette droite. Pour ce faire, une phase préliminaire est de choisir une longueur d'onde pour laquelle la dynamique $r_{MAX}(\lambda) - r_{MIN}(\lambda)$ est maximale. Sur la Figure 4.14, on montre un exemple de ce choix : pour chacune des 100 images de l'échantillon, on trace l'évolution de la dynamique en fonction de la longueur d'onde (courbes bleues), puis on moyenne les résultats

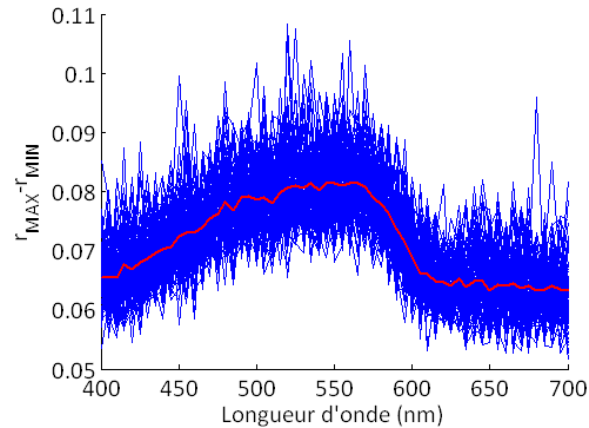


FIGURE 4.15 – Evolution de la dynamique $r_{MAX}(\lambda) - r_{MIN}(\lambda)$ en fonction de la longueur d'onde. La profondeur sur la crête centrale du fond est de 5 m, tandis que la pente dans la partie supérieure (resp. inférieure) est de 5% (resp. 10%).

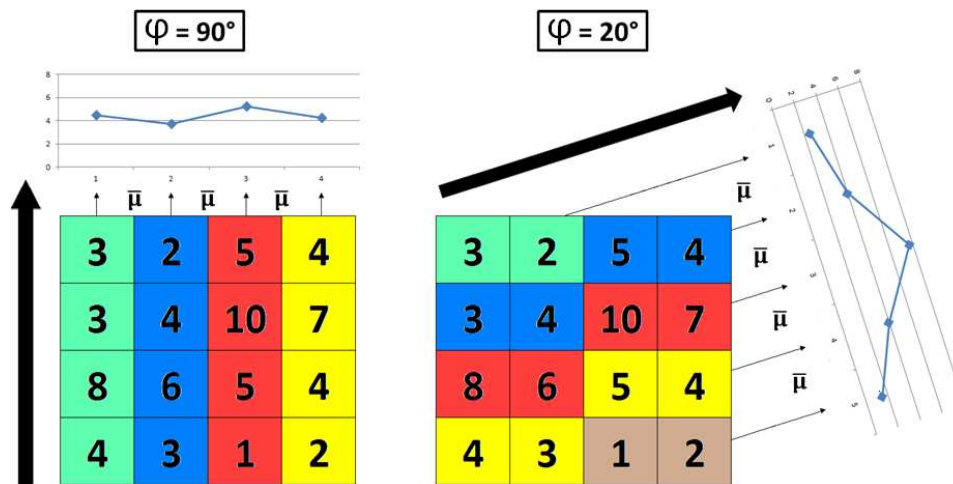


FIGURE 4.16 – Méthode de localisation de la droite de rupture.

(courbe rouge). Dans cette gamme de profondeurs, le domaine spectral approprié est compris entre 600 nm et 680 nm. En revanche, il se décalera vers le vert pour des hauteurs d'eau plus importantes, puisque, dans le rouge, le signal devient très atténué. En guise d'exemple, on trace les mêmes courbes sur la Figure 4.15, où la crête centrale du fond est à 5 m de profondeur, et on constate que le domaine spectral est ici compris entre 500 nm et 570 nm. En pratique, il faut donc retrouver cette longueur d'onde pour chaque image.

Passée cette étape préliminaire, il faut ensuite localiser la droite de rupture de pente sur cette bande.

La méthode est très proche de la transformée de Hough, et est illustrée sur la Figure 4.16, à l'aide d'une image 4×4 pixels. On construit un vecteur représentant la valeur moyenne des pixels situés sur chaque droite de direction φ (on fixe le pas de variation de l'angle φ à 1°). La taille de ce vecteur est égale au nombre de droites de direction φ (dans l'exemple, elle est de quatre pour $\varphi=90^\circ$, et de cinq pour $\varphi=20^\circ$). La rupture correspond alors à la droite de valeur moyenne minimum (vallée) ou maximum (crête). Dans l'exemple suivant, la crête se situe sur la droite rouge pour

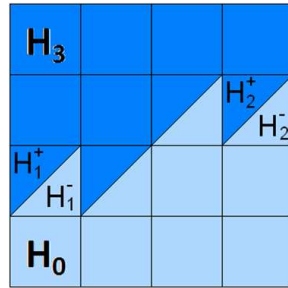


FIGURE 4.17 – Représentation de la contrainte $H_1^- = H_1^+$ et $H_2^- = H_2^+$.

$\varphi = 20^\circ$, tandis que la vallée se trouve sur la droite beige pour $\varphi = 20^\circ$.

Une fois que l'on connaît les deux sous-parties de l'image, on estime la pente sur chacune d'entre elles avec la méthode décrite dans la section précédente. En pratique, dans le cadre d'une approche aveugle, sans connaissance *a priori*, on peut calculer les deux pentes pour chaque scénario (vallée et crête), et choisir la description la plus vraisemblable.

L'estimation dans les deux régions peut se faire de manière dépendante ou indépendante, c'est-à-dire qu'il est possible de contraindre ou non $H_1^- = H_1^+$ et $H_2^- = H_2^+$ (voir Figure 4.17).

Si l'on estime séparément les deux triplets $\{A_1, B_1, H_{0,1}\}$ et $\{A_2, B_2, H_{0,2}\}$ correspondant à chaque région, il est possible de modéliser un relief discontinu tel qu'il est représenté sur la Figure 4.13 à gauche. En revanche, les pixels sont répartis en deux groupes, et la taille réduite de chaque échantillon a pour effet de diminuer la précision de l'estimation.

Si l'on impose la contrainte $H_1^- = H_1^+$ et $H_2^- = H_2^+$, on estime la profondeur sur toute l'image en une seule fois et on conserve donc la taille originale. Il est nécessaire d'effectuer le changement de variable suivant :

$$\{A_1, B_1, H_{0,1}, A_2, B_2, H_{0,2}\} \rightarrow \{H_0, H_1, H_2, H_3\} \quad (4.31)$$

où H_0, H_1, H_2 et H_3 sont les profondeurs à deux des quatre coins de l'image, et aux extrémités de la droite de rupture de pente.

Au-delà du fait de pouvoir travailler avec un échantillon plus grand, cette contrainte permet, par construction, de modéliser un relief continu. Elle permet également d'estimer une seule fois chaque concentration, au lieu

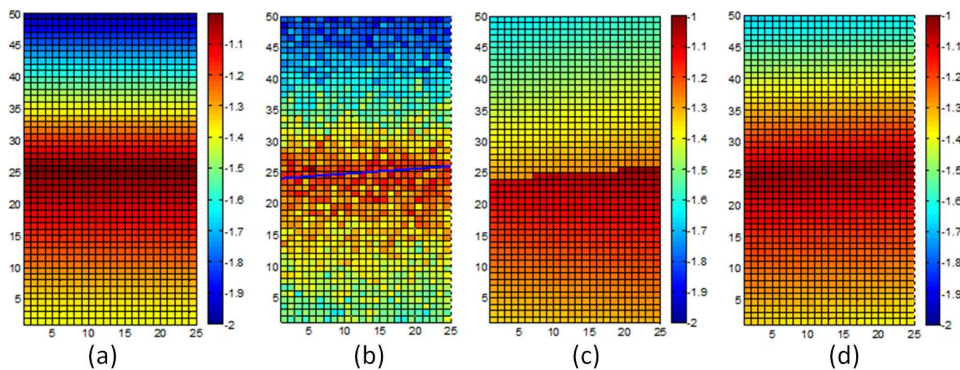


FIGURE 4.18 – (a) Vraie carte du relief, (b) localisation de la rupture de pente, (c) relief estimé sans la contrainte de continuité (EQM=0.0209), (d) relief estimé avec la contrainte de continuité (EQM=0.0083).

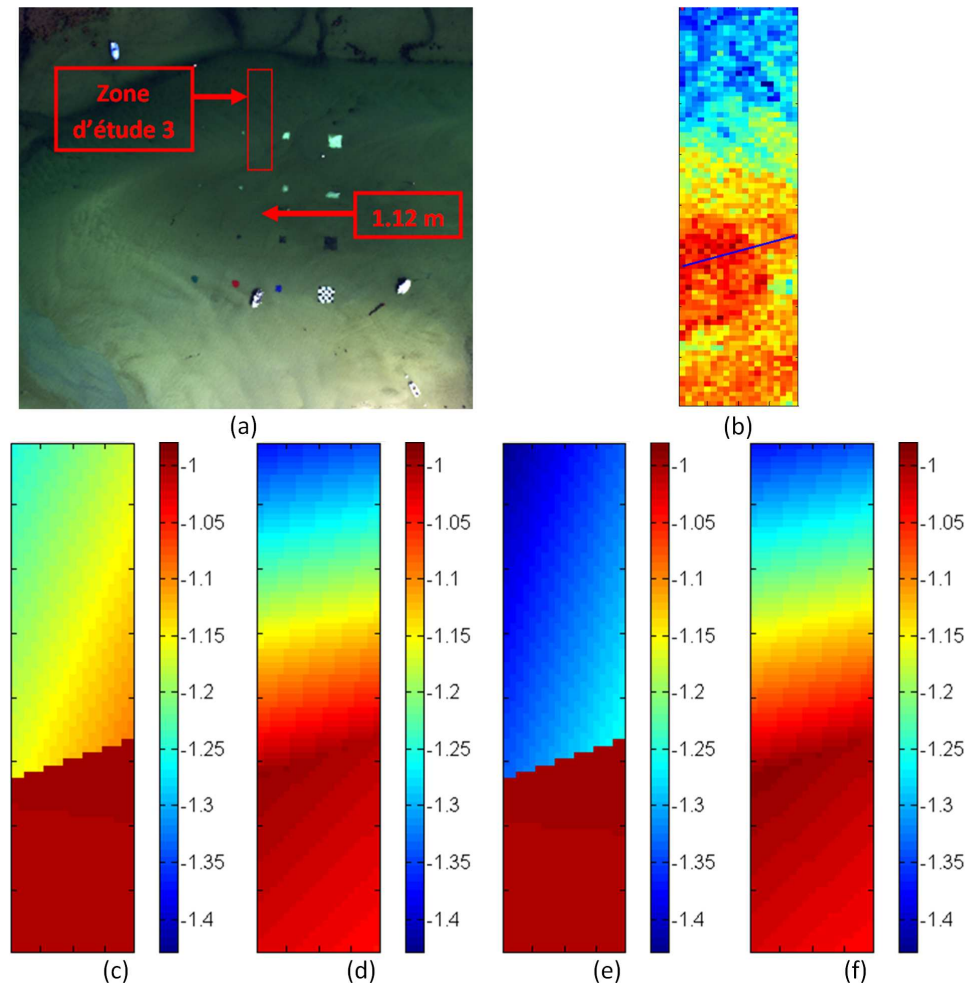


FIGURE 4.19 – (a) Image H1-0500m-Lo1-09odeg-20100917-10h20mn-atm, (b) localisation de la rupture de pente, (c) relief estimé sans la contrainte de continuité (modèle de Brando et al. (2009) et Bricaud et al. (1995) pour a_{PHY}), (d) relief estimé avec la contrainte de continuité (modèle de Brando et al. (2009) et Bricaud et al. (1995) pour a_{PHY}), (e) relief estimé sans la contrainte de continuité (modèle de Lee (1994) pour a_{PHY}), (f) relief estimé avec la contrainte de continuité (modèle de Lee (1994) pour a_{PHY}).

de deux dans le cas non contraint. Cependant, l'augmentation de la taille de l'échantillon est à mettre en parallèle avec l'ajout d'une inconnue supplémentaire par rapport à l'estimation d'un plan.

Résultats

Dans un premier temps, nous présentons des résultats obtenus en simulations. Sur la Figure 4.18, on compare la vraie carte du relief 4.18.a avec les deux cartes 4.18.c et 4.18.d estimées avec chaque méthode. De même que lors des tests sur l'estimation d'une seule pente, on prend $\sigma_B^2 = \sigma_S^2 = 0.01^2$; la hauteur varie entre 1 m (sur la crête) et 2 m (sur un bord). La sous-figure 4.18.b représente la bande choisie pour localiser la zone de rupture et la droite bleue la rupture estimée. Ici, on peut voir que la hauteur est mieux estimée par la méthode avec contraintes de continuité puisque l'erreur quadratique est réduite de 60%.

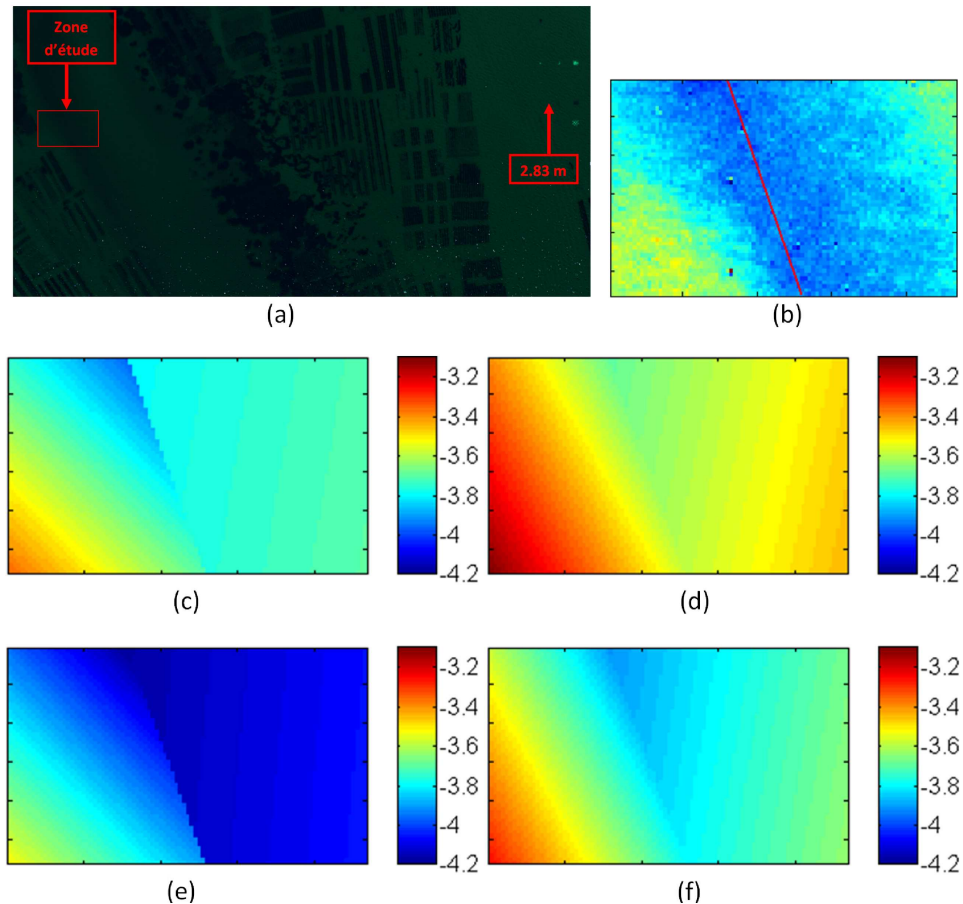


FIGURE 4.20 – (a) Image H2-0650m-Lo2-090deg-20100918-13h23mn-atm-1-deglintage-shoreline, (b) localisation de la rupture de pente, (c) relief estimé sans la contrainte de continuité (modèle de Brando et al. (2009) et Bricaud et al. (1995) pour a_{PHY}), (d) relief estimé avec la contrainte de continuité (modèle de Brando et al. (2009) et Bricaud et al. (1995) et al. pour a_{PHY}), (e) relief estimé sans la contrainte de continuité (modèle de Lee (1994) pour a_{PHY}), (f) relief estimé avec la contrainte de continuité (modèle de Lee (1994) pour a_{PHY}).

Dans un deuxième temps, nous présentons les résultats obtenus sur les images HypLitt. Comme pour l'estimation d'une seule pente, nous ne disposons pas de toute la vérité terrain, mais la seule visualisation de l'image couleur nous permet d'avoir une bonne indication sur la fiabilité de notre méthode.

Les résultats observés sur les figures 4.19 et 4.20 confirment la tendance, à savoir que l'apport de la contrainte de continuité est intéressant et permet de mieux estimer le relief. Quant aux différences liées au modèle d'absorption du phytoplancton choisi, il semblerait là encore que celui de Lee (1994) donne de meilleurs résultats que celui de Brando et al. (2009) et Bricaud et al. (1995).

4.3.4 En résumé

On a donc vu qu'il était possible d'adopter une représentation linéaire de la profondeur, et d'insérer ce modèle dans l'expression de la vraisemblance. Nous avons également pu modéliser une pente linéaire par morceaux, où l'on cherche d'abord la position de la vallée ou de la crête,

avant d'estimer la profondeur dans les deux régions. Ces deux méthodes ont été testées sur les simulations et sur les données HypLitt avec succès, puisque dans les deux cas, la précision sur l'estimation de la hauteur est augmentée. Par ailleurs, il est aussi intéressant de constater qu'en zone peu profonde, cette amélioration se traduit également à travers les résultats d'estimation des autres paramètres C_{PHY} , C_{CDOM} , et C_{NAP} .

4.4 ETUDE DU MAXIMUM A POSTERIORI (MAP)

4.4.1 Introduction et méthode

L'étude de l'EMV nous montre que cette méthode est très intéressante dans une gamme de profondeurs allant d'un mètre à une quinzaine de mètres (gamme variant selon la turbidité de l'eau). En revanche, à l'instar d'autres techniques comme cela est évoqué par Cannizzaro et Carder (2006) pour l'estimation de C_{PHY} en eau peu profonde, elle s'avère un peu moins performante pour de très faibles et très grandes hauteurs d'eau (quelques dizaines de centimètres et au-delà de 15-20 m). En effet, dans le premier cas, la couche d'eau est trop mince pour que les différents constituants influent vraiment sur la réflectance, et leur estimation est donc délicate. Dans le deuxième cas, le fond n'a quasiment plus d'influence sur cette même réflectance (ici, $r \approx r_\infty$), et il est donc très difficile d'estimer la profondeur.

Dans ces cas extrêmes, l'approche aveugle a ses limites, et il est intéressant d'apporter, si possible, un *a priori* sur l'ensemble des paramètres à estimer θ afin de conserver de bonnes performances. Cette loi *a priori* $P(\theta)$ intervient dans la fonction objectif à maximiser sous la forme suivante :

$$\hat{\theta}_{MAP}(\chi_N) = \underset{\theta}{\operatorname{argmax}} P(\theta|\chi_N) \quad (4.32)$$

$$\hat{\theta}_{MAP}(\chi_N) = \underset{\theta}{\operatorname{argmax}} \frac{P(\chi_N|\theta) \times P(\theta)}{P(\chi_N)} \quad (4.33)$$

$$\hat{\theta}_{MAP}(\chi_N) = \underset{\theta}{\operatorname{argmax}} P(\chi_N|\theta) \times P(\theta). \quad (4.34)$$

(= $\hat{\theta}_{MV}(\chi_N)$ pour une distribution uniforme)

Dans le cas où $\theta = \{H, C_{PHY}, C_{CDOM}, C_{NAP}\}$, on peut dans un premier temps, considérer que les variables H , C_{PHY} , C_{CDOM} , et C_{NAP} sont indépendantes (même si C_{PHY} et C_{CDOM} peuvent être corrélées, notamment en haute mer comme affirmé par Kutser (2004), mais ces relations sont mal connues actuellement), et ainsi, que :

$$P(\theta) = P(H) \times P(C_{PHY}) \times P(C_{CDOM}) \times P(C_{NAP}). \quad (4.35)$$

Cela revient donc à rechercher la valeur de θ qui maximise le logarithme de la fonction objectif suivante :

$$\begin{aligned} \hat{\theta}_{MAP}(\chi_N) &= \underset{\theta}{\operatorname{argmax}} \ln [P(\theta|\chi_N)] \\ \hat{\theta}_{MAP}(\chi_N) &= \underset{\theta}{\operatorname{argmax}} \{ l(\chi_N|\theta) + \ln [P(H)] + \ln [P(C_{PHY})] \\ &\quad + \ln [P(C_{CDOM})] + \ln [P(C_{NAP})] \} \end{aligned} \quad (4.36)$$

où $l(\chi_N|\boldsymbol{\theta}) = \ln [P(\chi_N|\boldsymbol{\theta})]$ est la fonction de log-vraisemblance.

Dans un premier temps, on considère ici des distributions *a priori* gaussiennes pour H , C_{PHY} , C_{CDOM} , et C_{NAP} , c'est-à-dire que l'on a par exemple $H \sim \mathcal{N}(\mu_H, \sigma_H^2)$. Un test d'adéquation du χ^2 , ayant été réalisé sur un jeu de données réelles fourni par l'IFREMER (2011), a d'ailleurs permis de valider cette hypothèse sur la distribution des paramètres C_{PHY} et C_{NAP} (les seuls paramètres accessibles, C_{CDOM} n'ayant pas été mesuré dans ce jeu de données).

Finalement, trouver $\hat{\boldsymbol{\theta}}_{MAP}(\chi_N)$ revient donc à maximiser la fonction $f_1(\chi_N, \boldsymbol{\theta})$ suivante :

$$\begin{aligned} f_1(\chi_N, \boldsymbol{\theta}) = & -\frac{N}{2} \ln \left[(2\pi)^L |\hat{\boldsymbol{\Gamma}}_{MV}(\chi_N; \boldsymbol{\theta})| \right] - \frac{L \times N}{2} \\ & - \frac{1}{2} \ln (2\pi\sigma_H^2) - \frac{1}{2\sigma_H^2} (H - \mu_H)^2 \\ & - \frac{1}{2} \ln (2\pi\sigma_{C_{PHY}}^2) - \frac{1}{2\sigma_{C_{PHY}}^2} (C_{PHY} - \mu_{C_{PHY}})^2 \\ & - \frac{1}{2} \ln (2\pi\sigma_{C_{CDOM}}^2) - \frac{1}{2\sigma_{C_{CDOM}}^2} (C_{CDOM} - \mu_{C_{CDOM}})^2 \\ & - \frac{1}{2} \ln (2\pi\sigma_{C_{NAP}}^2) - \frac{1}{2\sigma_{C_{NAP}}^2} (C_{NAP} - \mu_{C_{NAP}})^2 \end{aligned} \quad (4.37)$$

où $\hat{\boldsymbol{\Gamma}}_{MV}(\chi_N; \boldsymbol{\theta})$ est donné par l'Equation 4.9.

Dans un deuxième temps, il est intéressant d'étudier d'autres distributions, peut-être plus adaptées que la distribution normale, celle-ci autorisant en effet des valeurs négatives pour les paramètres. C'est pourquoi nous avons également utilisé des lois Gamma pour représenter la variabilité de C_{PHY} et C_{NAP} . Pour chacun d'entre eux, les tests du χ^2 réalisés sur deux ensembles de données réelles nous indiquent, à chaque fois, que ce type de loi permet de mieux décrire cette variabilité que les distributions gaussiennes. Sur les images HypLitt, nous testons donc une deuxième régularisation, où la variabilité de H est toujours décrit par une loi normale, mais où celles de C_{PHY} et C_{NAP} sont modélisées par des lois Gamma. En ce qui concerne le CDOM, aucune donnée réelle permettant d'estimer les paramètres statistiques des lois n'est disponible, et par conséquent, aucune régularisation n'est effectuée sur C_{CDOM} . En résumé, on a donc $H \sim \mathcal{N}(\mu_H, \sigma_H^2)$, $C_{PHY} \sim \Gamma(\alpha_{C_{PHY}}, \beta_{C_{PHY}})$, et $C_{NAP} \sim \Gamma(\alpha_{C_{NAP}}, \beta_{C_{NAP}})$.

La nouvelle fonction à maximiser $f_2(\chi_N, \boldsymbol{\theta})$ est la suivante :

$$\begin{aligned} f_2(\chi_N, \boldsymbol{\theta}) = & -\frac{N}{2} \ln \left[(2\pi)^L |\hat{\boldsymbol{\Gamma}}_{MV}(\chi_N; \boldsymbol{\theta})| \right] - \frac{L \times N}{2} \\ & - \frac{1}{2} \ln (2\pi\sigma_H^2) - \frac{1}{2\sigma_H^2} (H - \mu_H)^2 \\ & + \ln(C_{PHY}^{\alpha_{C_{PHY}}}) - \ln(\beta_{C_{PHY}}^{\alpha_{C_{PHY}}}) - \ln[\Gamma(\alpha_{C_{PHY}})] - \frac{C_{PHY}}{\beta_{C_{PHY}}} \\ & + \ln(C_{NAP}^{\alpha_{C_{NAP}}}) - \ln(\beta_{C_{NAP}}^{\alpha_{C_{NAP}}}) - \ln[\Gamma(\alpha_{C_{NAP}})] - \frac{C_{NAP}}{\beta_{C_{NAP}}} \end{aligned} \quad (4.38)$$

où Γ est la fonction Gamma.

En guise d'exemple, on trace sur la Figure 4.21 la fonction de vraisem-

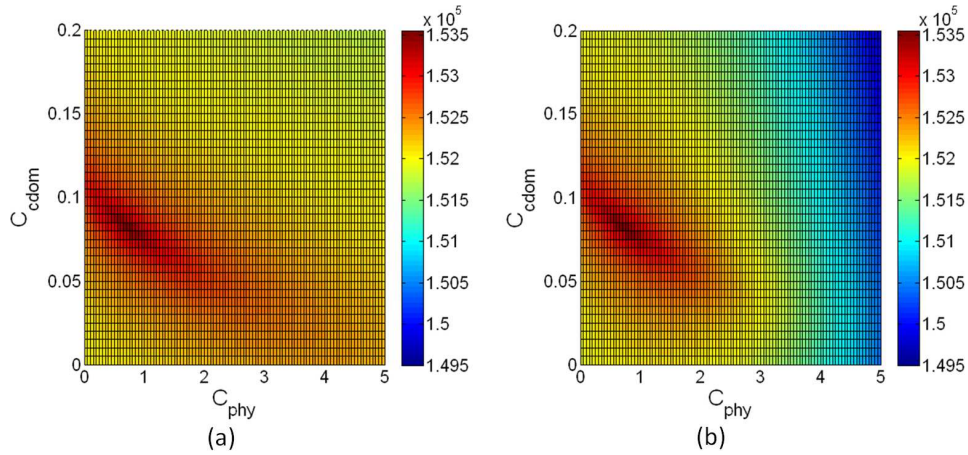


FIGURE 4.21 – (a) Fonction de log-vraisemblance, (b) fonction de log-vraisemblance régularisée en C_{PHY} et C_{CDOM} ($\mu_{C_{PHY}} = C_{PHY,VRAI} = 0.7 \mu\text{g}\cdot\text{L}^{-1}$, $\sigma_{C_{PHY}} = 10\% \times C_{PHY,VRAI}$, $\mu_{C_{CDOM}} = C_{CDOM,VRAI} = 0.08 \text{ m}^{-1}$, $\sigma_{C_{CDOM}} = 10\% \times C_{CDOM,VRAI}$).

blance $l(\chi_N|\theta)$ (à gauche) et la fonction $f_1(\chi_N, \theta)$ (à droite) régularisée en C_{PHY} et C_{CDOM} . On voit bien l'apport de la régularisation : la fonction devient plus "piquée", et rend les valeurs extrêmes moins vraisemblables. En revanche, cela ne modifie quasiment pas la zone autour du maximum. Grâce à cette méthode, il est donc possible d'éliminer certaines valeurs aberrantes obtenues lorsque les conditions d'estimation sont difficiles (faible et fortes profondeurs pour l'estimation de $\{C_{PHY}, C_{CDOM}, C_{NAP}\}$ et $\{H\}$ respectivement, faibles SNRs, etc).

Comme énoncé précédemment, le choix de la loi *a priori* est important : par exemple, dans le cas d'une distribution *a priori* gaussienne centrée sur la vraie valeur, plus l'écart-type choisi sera faible, et meilleure sera la précision.

4.4.2 Résultats

Nous avons mesuré l'apport de la régularisation de la vraisemblance en simulations et sur les données HypLitt.

Images simulées : tests avec la régularisation $f_1(\chi_N, \theta)$

En ce qui concerne les résultats sur données synthétiques, on compare sur la Figure 4.22, les erreurs calculées avec les estimateurs du MV et du MAP (ce dernier étant obtenu à partir de la fonction régularisée $f_1(\chi_N, \theta)$ de l'Equation 4.37), en fonction de la profondeur. Tout d'abord, on voit bien qu'ici, pour des profondeurs comprises entre 5 m et 10 m, il n'y a pas de différences significatives entre les deux : en effet, nous ne sommes pas ici dans des conditions d'estimation difficiles et la régularisation n'apporte pas grand chose. En revanche, pour de faibles profondeurs (ici 0.10 m), on constate qu'avec des écarts-types sur la distribution des paramètres égaux à 30% et 50% de la valeur moyenne, l'estimateur du MAP des concentrations permet d'obtenir une précision bien meilleure que celui du MV : par exemple, l'erreur sur C_{PHY} passe de près de 60% à environ 25% dans le cas où $\sigma_{C_{PHY}} = C_{PHY,VRAI} \times 50\%$. De la même façon pour de

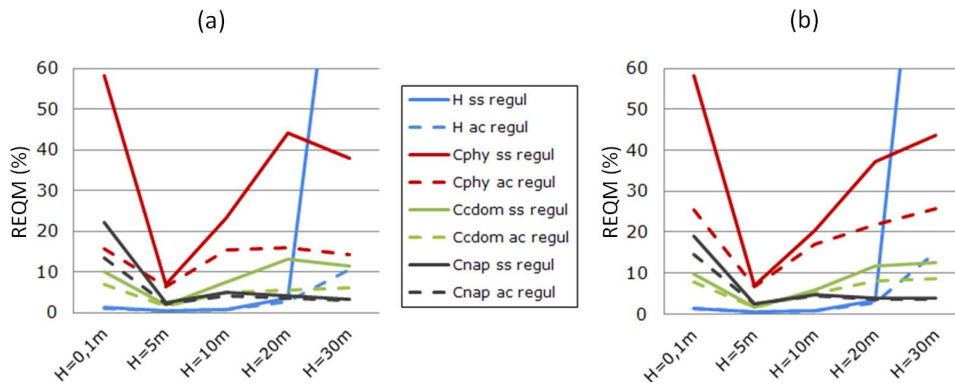


FIGURE 4.22 – REQMs relatives obtenues avec (lignes pointillées) ou sans (lignes pleines) régularisation, en fonction de la profondeur (données simulées). Pour chaque paramètre X , on a (a) $\mu_X = X_{VRAI}$, $\sigma_X = 30\% \times X_{VRAI}$, et (b) $\mu_X = X_{VRAI}$, $\sigma_X = 50\% \times X_{VRAI}$ ($C_{PHY} = 0.7 \mu\text{g}\cdot\text{L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$, $C_{NAP} = 2.8 \text{ mg}\cdot\text{L}^{-1}$).

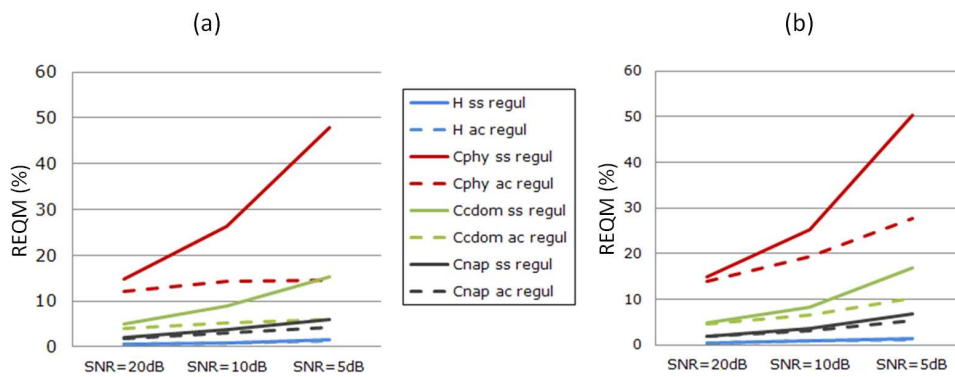


FIGURE 4.23 – REQMs relatives obtenues avec (lignes pointillées) ou sans (lignes pleines) régularisation, en fonction du SNR (données simulées). Pour chaque paramètre X , on a (a) $\mu_X = X_{VRAI}$, $\sigma_X = 30\% \times X_{VRAI}$, et (b) $\mu_X = X_{VRAI}$, $\sigma_X = 50\% \times X_{VRAI}$ ($H = 14 \text{ m}$, $C_{PHY} = 0.7 \mu\text{g}\cdot\text{L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$, $C_{NAP} = 2.8 \text{ mg}\cdot\text{L}^{-1}$).

fortes profondeurs (ici 30 m), le MAP permet d’améliorer notablement l’estimation de la hauteur H , puisque l’erreur passe de plus de 100% (non affichée ici par souci de visibilité) à environ 15% pour $\sigma_H = H_{VRAI} \times 50\%$. On vérifie également l’intérêt du MAP dans des conditions difficiles sur la Figure 4.23, où l’on montre l’évolution de l’erreur en fonction du SNR. On constate là encore, que plus l’image est bruitée, et plus grande est la différence entre MV et MAP.

Images HypLitt : tests avec les régularisations $f_1(\chi_N, \theta)$ et $f_2(\chi_N, \theta)$

Par ailleurs, nous avons comparé ces deux méthodes d’estimation sur les données HypLitt.

En ce qui concerne la régularisation $f_1(\chi_N, \theta)$, deux paramétrisations différentes pour le MAP ont été testées.

La première s’appuie sur un premier jeu de données fourni par l’IFREMER (2011) et déjà évoqué dans la section précédente. Comme indiqué sur la Figure 4.24, la zone de mesures se situe dans la baie de Quiberon, non loin du site S_3 . Nous avons extrait 35 (resp. 90) mesures de la concentration en phytoplancton (resp. de turbidité), effectuées entre 1996 et 2003, sur



FIGURE 4.24 – Localisation de la zone de mesure de la concentration en phytoplancton et de la turbidité.

		I ₁	I ₂	I ₃	I ₄	I ₅
Fond	$\mu_{C_{PHY}}$	1.5	1.5	1.6	3	2.4
	$\mu_{C_{NAP}}$	<2	<2	<2	<2	<2
Surface	$\mu_{C_{PHY}}$	1.5	2.4	2.2	2	1.6
	$\mu_{C_{NAP}}$	3	<2	<2	<2	<2
Moyenne	$\mu_{C_{PHY}}$	1.5	1.95	1.9	2.5	2
	$\mu_{C_{NAP}}$	<2.5	<2	<2	<2	<2

TABLE 4.7 – Valeurs mesurées par ActiMar sur les différents sites en S_3 ($\mu_{C_{PHY}}$ en $\mu\text{g.L}^{-1}$, $\mu_{C_{NAP}}$ en mg.L^{-1}).

une période de deux mois autour mi-septembre (date à laquelle ont été prises les images). La relation entre turbidité et concentration en matière en suspension est notamment évoquée par Doxaran et al. (2009), le rapport des deux grandeurs étant proche de l'unité. Les paramètres $\mu_{C_{PHY}}$, $\sigma_{C_{PHY}}$, $\mu_{C_{NAP}}$ et $\sigma_{C_{NAP}}$ sont ainsi estimés sur cet échantillon, et on a ainsi :

$$\begin{aligned}\hat{\mu}_{C_{PHY}} &= 1.88 \mu\text{g.L}^{-1}, & \hat{\sigma}_{C_{PHY}} &= 1.12 \mu\text{g.L}^{-1} \\ \hat{\mu}_{C_{NAP}} &= 1.47 \text{mg.L}^{-1}, & \hat{\sigma}_{C_{NAP}} &= 1.48 \text{mg.L}^{-1}.\end{aligned}$$

En ce qui concerne C_{CDOM} , aucune donnée n'est disponible, et par conséquent, nous ne régularisons pas la vraisemblance par rapport à ce paramètre.

La deuxième paramétrisation s'inspire des relevés réalisés par ActiMar lors de la campagne de mesures. Les paramètres $\mu_{C_{PHY}}$ et $\mu_{C_{NAP}}$ correspondent aux valeurs mesurées, et $\sigma_{C_{PHY}}$ et $\sigma_{C_{NAP}}$ sont pris égaux à 100% des valeurs moyennes. Les données sont présentées dans le Tableau 4.7. Dans nos tests, nous prendrons les valeurs moyennes des deux mesures effectuées en surface et au fond. En ce qui concerne le paramètre C_{NAP} où seuls des encadrements des concentrations réelles sont fournis, nous choisissons d'utiliser $\mu_{C_{NAP}} = 2 \text{mg.L}^{-1}$ en I₁, et $\mu_{C_{NAP}} = 1 \text{mg.L}^{-1}$ en I₂, I₃, I₄, et I₅. Enfin, comme pour la première paramétrisation, aucune régularisation n'est faite sur la concentration C_{CDOM} .

Lors des tests réalisés sur les images HypLitt, nous nous sommes rendus compte que les différences de résultats entre ces deux paramétrisations étaient finalement assez minimes. Par conséquent, nous ne les distinguons

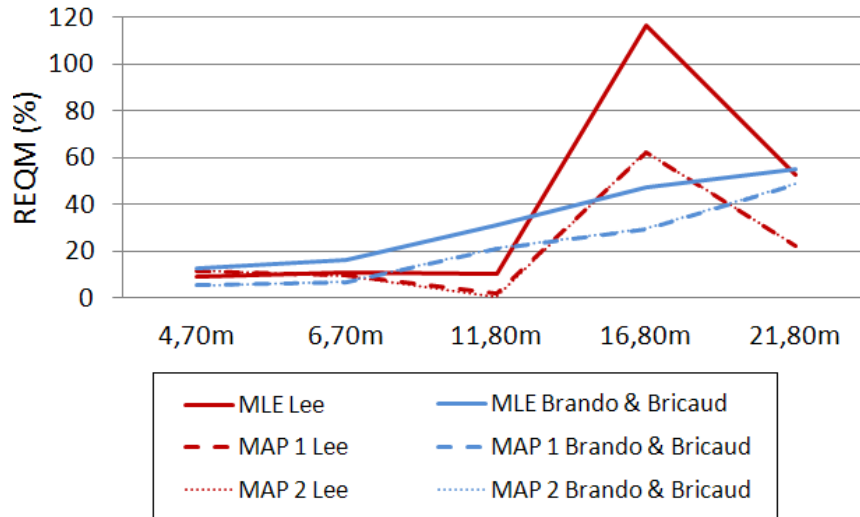


FIGURE 4.25 – REQMs relatives sur l'estimation de la hauteur d'eau, moyennées sur deux images de la même profondeur, et obtenues avec (lignes pointillées) ou sans (lignes pleines) régularisation, en fonction de la profondeur (données HypLitt).

pas sur la Figure 4.25 et les Tableaux 4.8 et 4.9.

En ce qui concerne la deuxième régularisation $f_2(\chi_N, \theta)$, nous utilisons un deuxième jeu de données fourni par l'IFREMER (2011). Cette fois-ci, nous avons extrait 88 (resp. 191) mesures de la concentration en phytoplancton (resp. de turbidité), effectuées entre 1996 et 2003, sur une période de quatre mois autour mi-septembre (de mi-juillet à mi-novembre) afin d'avoir davantage de données pour estimer les paramètres statistiques des lois. Ici, les moyennes et écarts-types empiriques de C_{PHY} et C_{NAP} sont :

$$\begin{aligned}\hat{\mu}_{C_{PHY}} &= 1.64 \mu\text{g.L}^{-1}, & \hat{\sigma}_{C_{PHY}} &= 1.00 \mu\text{g.L}^{-1} \\ \hat{\mu}_{C_{NAP}} &= 1.57 \text{mg.L}^{-1}, & \hat{\sigma}_{C_{NAP}} &= 1.77 \text{mg.L}^{-1}.\end{aligned}$$

La méthode des moments nous donne des estimateurs des paramètres α et β des lois Gamma en fonction des moments empiriques d'ordre 1 et 2, et on a ainsi :

$$\begin{aligned}\hat{\alpha}_{C_{PHY}} &= 2.73, & \hat{\beta}_{C_{PHY}} &= 0.60 \\ \hat{\alpha}_{C_{NAP}} &= 0.79, & \hat{\beta}_{C_{NAP}} &= 1.98.\end{aligned}$$

Enfin, en ce qui concerne la profondeur, dans les deux cas, μ_H est donné par les mesures réalisées *in situ*, et σ_H est choisi égal à 50% de μ_H .

Seules les images du site S_3 sont utilisées lors des tests, puisque les prélèvements d'eau n'ont pas été effectués en S_1 et S_2 .

On présente sur la Figure 4.25 l'évolution des erreurs sur l'estimation de la profondeur pour les deux modèles d'absorption du phytoplancton présentés dans la Section 2.3.1. On voit bien ici l'intérêt du MAP, notamment lorsque la hauteur d'eau augmente : par exemple, l'erreur diminue de moitié pour $H = 16.80$ m, et ce, quelque soit l'expression de a_{PHY} . En revanche, comme sur les simulations, il n'y a que peu d'écart entre

	I ₁ (H = 4.70 m)			I ₂ (H = 6.70 m)		
	MV	MAP 1	MAP 2	MV	MAP 1	MAP 2
C _{PHY}	4.82	4.72	4.76	4.49	4.37	4.35
C _{CDOM}	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵
C _{NAP}	1.53	1.43	1.37	2.82	2.66	2.72
	I ₃ (H = 11.80 m)			I ₄ (H = 16.80 m)		
	MV	MAP 1	MAP 2	MV	MAP 1	MAP 2
C _{PHY}	3.60	3.52	3.52	3.37	3.34	3.35
C _{CDOM}	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵
C _{NAP}	4.19	4.04	4.10	6.12	5.96	6.00
	I ₅ (H = 21.80 m)					
	MV	MAP 1	MAP 2			
C _{PHY}	3.18	3.12	3.12			
C _{CDOM}	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵			
C _{NAP}	4.77	4.60	4.65			

TABLE 4.8 – Concentrations estimées avec le modèle de Brando et al. (2009) et Bricaud et al. (1995) (C_{PHY} en $\mu\text{g.L}^{-1}$, C_{CDOM} en m^{-1} , et C_{NAP} en mg.L^{-1}). Les deux régularisations sont testées.

MV et MAP pour $H = 4.70$ m et $H = 6.70$ m. Par ailleurs, on constate ici que les deux méthodes du MAP proposées donnent des résultats quasi-identiques. On peut expliquer cela par le fait que la régularisation effectuée en H est la même dans les deux cas, et que par conséquent, les modifications apportées par le passage aux lois Gamma sur les autres paramètres ne sont pas suffisantes pour influencer l'estimation de la profondeur.

En ce qui concerne l'estimation des différentes concentrations, on constate sur les Tableaux 4.8 et 4.9 que, quelque soit le modèle d'absorption du phytoplancton, les deux méthodes du MAP améliorent très légèrement les résultats. Cependant, l'écart est tel qu'on ne peut pas considérer que cette amélioration soit significative. Deux raisons peuvent expliquer cela. La première provient du fait que l'on ne considère ici que des profondeurs supérieures à 4.70 m. Or, comme nous l'avons déjà expliqué auparavant, le MAP permet d'améliorer l'estimation d'un paramètre lorsque les conditions relatives à cette dernière sont difficiles. Pour les concentrations, cela correspond à une gamme de faibles profondeurs, et par conséquent, ce phénomène est moins marqué lorsque la hauteur d'eau est plus importante. D'ailleurs, cela permet également d'expliquer le fait que les deux méthodes du MAP donnent des résultats équivalents. La deuxième raison des faibles écarts de résultats entre MV et MAP est le fait que les écarts-types choisis ou estimés à partir des vérités terrain disponibles sont trop importants pour que cela ait une réelle influence par rapport aux résultats de l'EMV. Il serait difficile de les réduire car, comme on a pu le constater sur les données de l'IFREMER (2011), ces concentrations peuvent varier significativement sur la même période. Enfin, de façon générale, on voit ici l'intérêt du modèle de Lee (1994) pour évaluer C_{PHY}, puisqu'il suffit de comparer avec le Tableau 4.7 pour s'apercevoir que les valeurs estimées sont beaucoup plus réalistes. En outre, ces résultats sont meilleurs si l'on distingue les eaux optiquement

	I ₁ (H = 4.70 m)			I ₂ (H = 6.70 m)		
	MV	MAP 1	MAP 2	MV	MAP 1	MAP 2
C _{PHY}	2.74	2.66	2.66	2.06	2.02	2.01
C _{CDOM}	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵
C _{NAP}	3.86	3.59	3.56	4.23	3.95	4.02
	I ₃ (H = 11.80 m)			I ₄ (H = 16.80 m)		
	MV	MAP 1	MAP 2	MV	MAP 1	MAP 2
C _{PHY}	1.76	1.74	1.74	1.51	1.51	1.51
C _{CDOM}	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵
C _{NAP}	4.84	4.57	4.66	5.69	5.68	5.69
	I ₅ (H = 21.80 m)					
	MV	MAP 1	MAP 2			
C _{PHY}	1.49	1.49	1.49			
C _{CDOM}	<10 ⁻⁵	<10 ⁻⁵	<10 ⁻⁵			
C _{NAP}	5.81	5.80	5.80			

TABLE 4.9 – Concentrations estimées avec le modèle de Lee (1994) (C_{PHY} en $\mu\text{g.L}^{-1}$, C_{CDOM} en m^{-1} , et C_{NAP} en mg.L^{-1}). Les deux régularisations sont testées.

peu profondes, où le fond est visible ou presque (voir les images à 4.70 m et 6.70 m sur la Figure 1.5), des eaux optiquement profondes (voir les images à 11.80 m, 16.80 m et 21.80 m sur la Figure 1.5). En effet, dans le premier cas, il est pertinent de comparer les résultats d'estimation de C_{PHY} avec la moyenne des concentrations mesurées en surface et au fond, puisque les deux valeurs influent sur la réflectance subsurface. En revanche, dans le deuxième cas, le spectre observé (qui est proche de r_∞) dépend essentiellement de la valeur de C_{PHY} mesurée en surface et par conséquent, les valeurs estimées doivent plutôt être comparées à ces concentrations-là.

Cela nous permet de souligner une nouvelle fois l'importance du choix d'un modèle réaliste dans la méthode EMV. Nous l'avons déjà constaté sur le Tableau 4.4, où nous voyions que le modèle de réflectance M2 permettait d'augmenter notablement les performances de l'EMV par rapport au modèle M1. Ce choix joue un rôle déterminant sur les résultats, et l'utilisation de modèles encore plus fiables représente donc une piste d'amélioration majeure de nos méthodes.

4.5 ETUDE DES BORNES DE CRAMÉR-RAO

4.5.1 Introduction et méthode

Afin d'évaluer les performances de la méthode EMV, nous pouvons comparer les résultats avec ceux donnés par d'autres techniques. Mais cela ne nous donne aucune indication sur la précision maximale que l'on peut atteindre, et on ne sait donc pas si, en théorie, il est possible d'obtenir de meilleurs résultats. Voilà pourquoi la borne de Cramér-Rao (BCR) est un outil très intéressant puisqu'elle correspond justement à la valeur minimale théorique que peut prendre la variance d'un estimateur

non-biaisé de θ .

Dans le cas scalaire $\theta = \{\theta_1\}$ et pour un N -échantillon χ_N , elle s'exprime comme l'inverse de l'information de Fisher $I_N(\theta)$:

$$\mathcal{BCR}_N(\theta)^{-1} = I_N(\theta) = \mathbb{E} \left[\left(\frac{\partial \ln(P(\chi_N|\theta))}{\partial \theta_1} \right)^2 \right]. \quad (4.39)$$

Dans le cas multivarié où $\theta = \{\theta_1, \dots, \theta_P\}$, cette relation devient :

$$\begin{aligned} [\mathcal{BCR}_N(\theta)^{-1}]_{i,j} &= [I_N(\theta)]_{i,j} \\ &= \mathbb{E} \left[\frac{\partial \ln(P(\chi_N|\theta))}{\partial \theta_i} \frac{\partial \ln(P(\chi_N|\theta))}{\partial \theta_j} \right]. \end{aligned} \quad (4.40)$$

Présentée par Stoica et Moses (1997), la formule de Slepian-Bangs donne l'expression intégrée de $I_1(\theta)$ dans le cas multivarié. Il est alors facile de retrouver $I_N(\theta)$, puisque, pour $\chi_N = \{r_1, \dots, r_N\}$ où tous les r_i sont indépendants et identiquement distribués, on peut montrer que $I_N(\theta) = N \times I_1(\theta)$.

La formule de Slepian-Bangs relie $I_1(\theta)$ avec μ , Γ et leurs dérivées par rapport aux composantes de θ :

$$[I_1(\theta)]_{i,j} = \frac{1}{2} \text{tr} \left(\Gamma^{-1} \frac{\partial \Gamma}{\partial \theta_i} \Gamma^{-1} \frac{\partial \Gamma}{\partial \theta_j} \right) + \frac{\partial \mu^t}{\partial \theta_i} \Gamma^{-1} \frac{\partial \mu}{\partial \theta_j}. \quad (4.41)$$

Afin de pouvoir calculer les dérivées $\frac{\partial \mu}{\partial \theta_i}$ et $\frac{\partial \Gamma}{\partial \theta_i}$, il est donc nécessaire de connaître la relation entre les paramètres statistiques μ et Γ , et les paramètres physiques H , C_{PHY} , C_{CDOM} et C_{NAP} . Ici, on utilise le modèle M1 de Maritorea et al. (1994) mis sous forme matricielle dans la Section 4.2.1 pour représenter la réflectance.

On utilise les expressions de μ et Γ données par les Equations 4.7 et 4.8 respectivement. Comme dans les tests précédents, on prend $\sigma_B^2 = \sigma_S^2$.

Sous ces hypothèses, il est alors possible de calculer les BCRs (ces calculs étant détaillés en Annexe A.3).

Dans le cas scalaire (c'est-à-dire pour $\theta = \{H\}$, $\theta = \{C_{PHY}\}$, $\theta = \{C_{CDOM}\}$ et $\theta = \{C_{NAP}\}$), on obtient :

$$\begin{aligned} \mathcal{BCR}_1(H)^{-1} &= \sum_{l=1}^L \frac{8k(\lambda_l)^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right)^2 [\Gamma^{-1}]_{l,l} \\ \mathcal{BCR}_1(C_{PHY})^{-1} &= \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_{PHY}} \right)^2 H^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_{PHY}} \right)^2 [\Gamma^{-1}]_{l,l} \\ \mathcal{BCR}_1(C_{CDOM})^{-1} &= \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_{CDOM}} \right)^2 H^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_{CDOM}} \right)^2 [\Gamma^{-1}]_{l,l} \\ \mathcal{BCR}_1(C_{NAP})^{-1} &= \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_{NAP}} \right)^2 H^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_{NAP}} \right)^2 [\Gamma^{-1}]_{l,l}. \end{aligned} \quad (4.42)$$

Dans le cas multivarié, le calcul de ces bornes pour l'estimation de $\theta = \{H, C_{PHY}, C_{CDOM}, C_{NAP}\}$ nous donne :

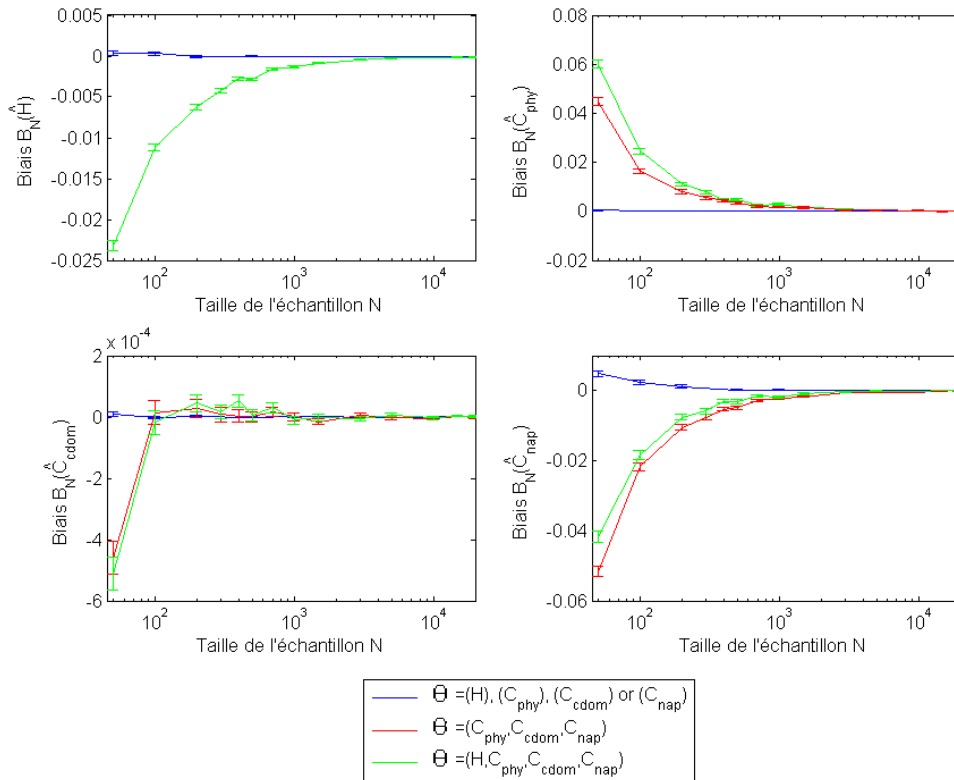


FIGURE 4.26 – Evolution du biais expérimental en fonction de la taille de l'échantillon ($M = 50000$). Ici, on a : $H = 10 \text{ m}$, $C_{PHY} = 0.7 \mu \text{ g.L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$, et $C_{NAP} = 2.8 \text{ mg.L}^{-1}$.

– Lorsque $\theta_i = H$ et $\theta_j = H$:

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8k(\lambda_l)^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right)^2 [\boldsymbol{\Gamma}^{-1}]_{l,l} \quad (4.43)$$

– Lorsque $\theta_i = H$ et $\theta_j = C_c$:

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8k(\lambda_l) \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) H e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l} \quad (4.44)$$

– Lorsque $\theta_i = C_c$ et $\theta_j = C_d$:

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial k(\lambda_l)}{\partial C_d} \right) H^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_d} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l} \quad (4.45)$$

4.5.2 Résultats

Avant de comparer nos résultats avec les BCRs, il est nécessaire de vérifier que les EMVs sont non biaisés. Etant données leurs propriétés asymptotiques, cela revient à savoir combien de pixels d'entraînement

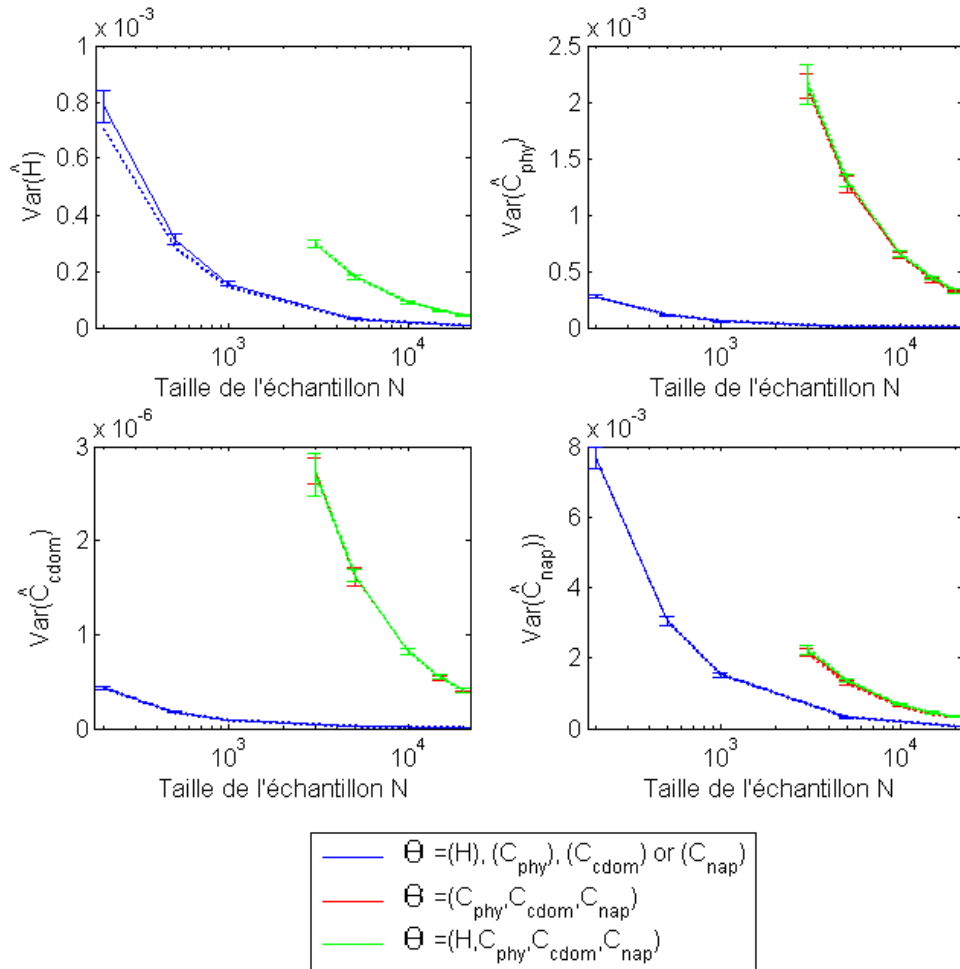


FIGURE 4.27 – Comparaison entre les BCRs (lignes pointillées) et les variances expérimentales des estimateurs (lignes pleines) en fonction de la taille de l'échantillon ($H = 10 \text{ m}$, $C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$, $C_{NAP} = 2.8 \text{ mg.L}^{-1}$).

est-il nécessaire d'avoir afin d'obtenir un estimateur sans biais.

Sur la Figure 4.26, on montre l'évolution du biais expérimental $B_N(\hat{\theta}_j)$ calculé à partir de M valeurs de $\hat{\theta}_{MV}(\chi_N)$ (chaque valeur de $\hat{\theta}_{MV}(\chi_N)$ étant calculée à partir d'un N -échantillon), tel que $B_N(\hat{\theta}_j) = \frac{1}{M} \sum_{i=1}^M (\hat{\theta}_{j,i} - \theta_j)$.

L'incertitude sur $B_N(\hat{\theta}_j)$ est donnée par son écart-type expérimental $\sigma(B_N(\hat{\theta}_j)) = \frac{1}{\sqrt{M}} \sigma(\hat{\theta}_j)$, où $\sigma(\hat{\theta}_j)$ est l'écart-type de la j^{eme} composante de $\hat{\theta}_{MV}(\chi_N)$.

Dans le cas où θ est scalaire, on peut voir qu'environ 100 à 200 pixels suffisent pour que chaque estimateur $\hat{\theta}_j$ soit non biaisé. En revanche, dans le cas multivarié ($\theta = \{C_{PHY}, C_{CDOM}, C_{NAP}\}$ ou $\theta = \{H, C_{PHY}, C_{CDOM}, C_{NAP}\}$), plusieurs milliers de pixels sont nécessaires afin d'obtenir le même résultat. Par la suite, il sera donc important de veiller à choisir un échantillon de taille suffisante.

Sur la Figure 4.27 où l'on compare les BCRs et les variances des estimateurs, on choisit donc $N = 200$ (resp. $N = 3000$) pour θ scalaire (resp. multivarié). On constate que quelque soit l'ensemble θ à estimer,

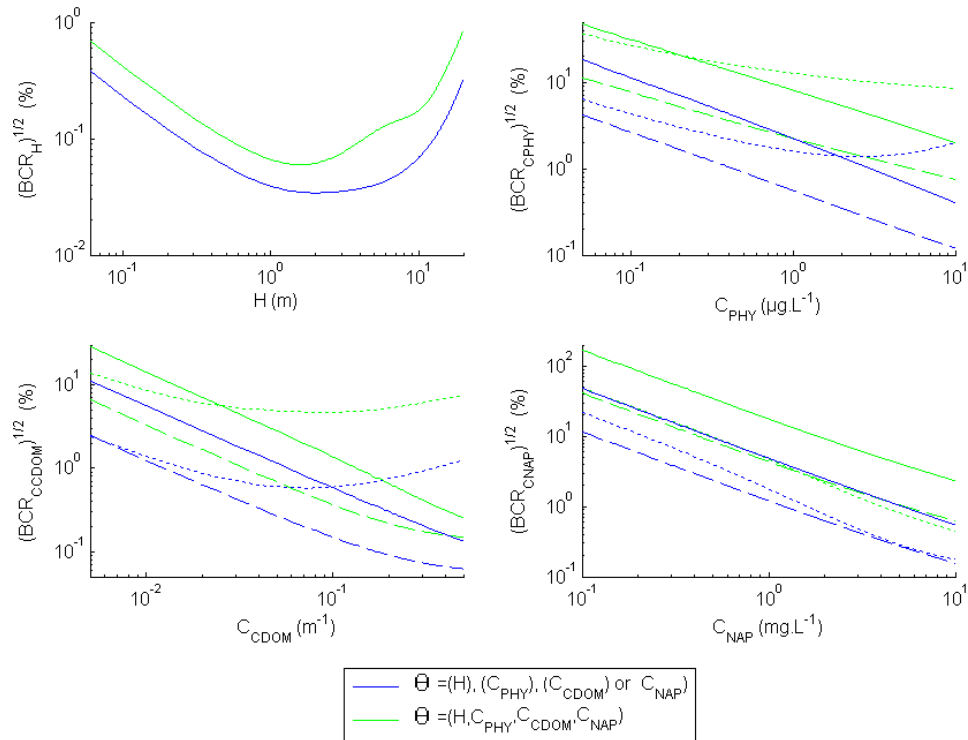


FIGURE 4.28 – Evolution des BCRs en fonction des valeurs des paramètres ($N = 3000$). Les lignes pleines correspondent à $H = 0.20$ m, les tirets à $H = 1$ m, et les pointillés à $H = 20$ m. Lorsque ces paramètres ne varient pas, on a : $C_{PHY} = 0.7 \mu\text{g.L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$, $C_{NAP} = 2.8 \text{ mg.L}^{-1}$.

les variances sont égales aux BCRs associées. Cela confirme donc la théorie disant que dans la famille exponentielle, dès lors que l'EMV $\hat{\theta}_{MV}(\chi_N)$ est non biaisé, il est à variance minimale (et sa variance est égale à la BCR).

Sachant cela, nous pouvons désormais étudier le comportement de ces bornes (et donc de la précision de l'estimation) lorsque les différents paramètres H , C_{PHY} , C_{CDOM} , et C_{NAP} varient. Sur la Figure 4.28, on montre l'évolution des BCRs en fonction des paramètres, et ce pour trois profondeurs différentes (H faible, intermédiaire et grand). En ce qui concerne l'estimation de H , on voit que la précision est maximale entre un peu moins d'un mètre et une dizaine de mètres. De même, la précision sur l'estimation des concentrations est également plus grande à 1 m qu'à 0.20 m et 20 m, et comme expliqué dans la Section 4.2.5, on retrouve le fait que la méthode EMV est plus performante pour cette gamme de profondeurs.

D'une façon générale, on voit d'ailleurs que plus l'eau est trouble, et mieux on estime les concentrations : cela peut s'expliquer par le fait que pour des valeurs faibles de C_{PHY} , C_{CDOM} , ou C_{NAP} , les spectres d'absorption et de rétro-diffusion de l'eau sont principalement influencés par ceux de l'eau pure, et il est donc difficile de retrouver l'apport de ces paramètres. Ce phénomène est plus marqué en faible (0.20 m et 1 m) qu'en forte profondeur (20 m), car la réflectance subsurface dépend également beaucoup de la réflectance du fond, ce qui diminue encore davantage l'influence des concentrations.

Sur les simulations, l'étude des bornes de Cramér-Rao nous permet donc de voir que la méthode EMV est très intéressante puisque, lorsque nous possédons suffisamment de pixels, elle donne l'unique estimateur sans biais à variance minimale. Toute autre méthode aboutira à des résultats inférieurs (au sens du critère "non biaisé et variance minimale"). Par ailleurs, au-delà de la caractérisation de la précision, on pourrait également utiliser cet outil statistique pour qualifier le réalisme d'un modèle de réflectance. En effet, on a vu que lorsque le modèle choisi suit parfaitement les données (ce qui est le cas dans les simulations), la variance est égale à la BCR. Inversement, sur les images réelles, cette borne pourrait donc correspondre à une mesure de la fiabilité du modèle sélectionné. Enfin, il sera intéressant de calculer ces bornes en changeant d'une part, de modèle de réflectance afin d'être encore plus fidèle à la réalité, et d'autre part de modèle de covariance en choisissant une forme plus générale.

4.6 CONCLUSIONS

Nous avons développé une méthode d'estimation par maximum de vraisemblance permettant de retrouver tout ou partie des paramètres caractéristiques de la colonne d'eau. Les tests réalisés sur les simulations et sur les images HypLitt démontrent l'intérêt de cette technique, en particulier par rapport à l'inversion par minimisation d'une fonction d'erreur, pour une gamme de profondeurs situées entre plusieurs dizaines de centimètres et dix ou vingt mètres selon la turbidité de l'eau. Nous voyons également que cette méthode est robuste pour de forts niveaux de bruit. Toutefois, cette méthode suppose la connaissance parfaite du spectre du matériau constituant le fond marin. Or, en pratique, celle-ci est rarement parfaite, et il est plus réaliste de considérer le fond connu à un paramètre multiplicatif près. Nous avons donc développé l'estimation au sens du maximum de vraisemblance en intégrant l'estimation de ce paramètre et de la variance spectrale, en supposant la matrice de covariance diagonale. Nous montrons sur des simulations et sur les données réelles que, malgré cette hypothèse sur la matrice de covariance, l'estimation des paramètres caractéristiques de la colonne d'eau est largement améliorée lorsque le spectre moyen du fond n'est pas parfaitement connu, en particulier pour les faibles profondeurs. Cependant, cette approche n'est plus pertinente lorsque la hauteur d'eau augmente, la forme spectrale du fond ayant moins d'influence sur la réflectance subsurface.

Par ailleurs, nous avons inséré un modèle de pente linéaire pour estimer les variations de la hauteur d'eau au sein de la zone étudiée. De plus, nous avons étendu cette représentation à la modélisation d'une pente linéaire par morceaux, où l'on repère dans un premier temps la zone de rupture de pente, avant d'estimer la profondeur avec la méthode précédente. Nous voyons sur les simulations et sur les données HypLitt que cela améliore notablement l'estimation de la hauteur et des concentrations lorsque la profondeur reste modérée. Au-delà, les variations de H par rapport à sa valeur moyenne sont faibles, et l'hypothèse de fond plat est suffisante pour conserver de bons résultats.

Afin d'améliorer l'estimation lorsque les conditions sont difficiles (faible et forte profondeurs, faible SNR), nous avons étudié l'intérêt d'apporter une connaissance *a priori* sur la distribution des paramètres inconnus. Nous avons testé deux régularisations différentes, selon que les variabilités de C_{PHY} et C_{NAP} soient modélisées par des lois gaussiennes ou des lois Gamma. Sur les simulations et sur les images Hyplitt, nous voyons que cela permet de mieux estimer les concentrations et la profondeur lorsque la hauteur d'eau est faible et forte respectivement. En revanche, si les lois Gamma permettent de mieux représenter les variabilités de C_{PHY} et C_{NAP} que les lois gaussiennes, les résultats obtenus par les deux approches sont assez proches. Dans cette section, nous avons également constaté l'importance primordiale du choix du modèle bathymétrique sur les résultats : en effet, les deux modèles d'absorption du phytoplancton comparés conduisent à des performances d'estimation notablement différentes.

Enfin, lors de simulations, nous avons calculé les bornes de Cramér-Rao associées à l'estimation des paramètres caractéristiques de la colonne d'eau. Ce puissant outil nous permet de vérifier expérimentalement que, lorsque l'estimateur du MV est sans biais, il est à variance minimale et représente donc le meilleur choix selon le critère classique "sans biais et à variance minimale". Le tracé de ces bornes en fonction de H , C_{PHY} , C_{CDOM} , et C_{NAP} nous montre l'évolution de la précision : on voit notamment que plus l'eau est trouble, et mieux on estime les concentrations. De plus, la précision théorique sur l'estimation de la profondeur est maximale entre plusieurs dizaines de centimètres et une dizaine de mètres, ce qui est conforme aux résultats expérimentaux et permet d'apporter une justification théorique à ceux-ci. Par ailleurs, l'étude de cet outil et l'atteinte des BCRs pour les erreurs expérimentales lors des simulations, permet de valider le choix de la méthode d'optimisation que nous avons fait, confirmant que celle-ci permet d'approcher le maximum de la fonction de vraisemblance de façon satisfaisante.

MÉTHODES DE DÉTECTION DE CIBLES IMMERGÉES

5.1 INTRODUCTION

EN imagerie hyperspectrale, la détection de cibles représente une problématique majeure pour la défense et la surveillance de l'environnement. On entend par cible tout objet d'intérêt (véhicules, mines, etc) dont une des caractéristiques fréquentes en télédétection est sa petite taille vis-à-vis de la scène qui l'entoure.

L'approche classique consiste à exploiter la parcimonie des cibles et considérer que leur spectre est suffisamment différent de celui du fond. Le processus de détection s'effectue en deux étapes. Dans un premier temps, une fonction de contraste est dérivée à partir des données, des hypothèses formulées sur la valeur du pixel étudié, et des connaissances *a priori* (si disponibles). Appliquée à l'ensemble de l'image, elle permet d'obtenir une carte de détection en niveau de gris, chaque niveau de gris étant lié à la probabilité de présence d'une cible. Dans un deuxième temps, cette carte est seuillée afin d'obtenir un masque de détection binaire, indiquant la présence ou l'absence de cible dans chaque pixel.

Le terme de détection supervisée est utilisé lorsque le spectre de la cible est connu *a priori*. Dans la littérature, les détecteurs supervisés de référence sont le filtre adapté (MF pour Matched Filter, AMF pour Adaptive Matched Filter) et le filtre ACE (Adaptive Cosine/Coherence Estimator), présentés par Manolakis et Shaw (2002). Leur construction se base sur des hypothèses de présence (H_1) ou d'absence (H_0) de la cible, et sur l'écriture du test du rapport de vraisemblance associé (test de Neyman-Pearson) :

$$D(\mathbf{r}) = \frac{P(\mathbf{r}; \boldsymbol{\theta}_1 | H_1)}{P(\mathbf{r}; \boldsymbol{\theta}_0 | H_0)} \underset{H_0}{\overset{H_1}{\geq}} \eta \quad (5.1)$$

où $\boldsymbol{\theta}_0$ et $\boldsymbol{\theta}_1$ sont les ensembles de paramètres caractérisant les lois de probabilité associées aux hypothèses H_0 et H_1 respectivement. La fonction D est la fonction de contraste évoquée précédemment, et η le seuil de prise de décision permettant d'obtenir un masque binaire.

Pour le filtre adapté, on suppose que chaque pixel spectral est une réalisation d'un vecteur aléatoire gaussien multivarié \mathbf{r} pouvant suivre l'une ou

l'autre des deux hypothèses suivantes :

$$H_0 : \text{la cible est absente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_{0,B}, \boldsymbol{\Gamma}_0)$$

$$H_1 : \text{la cible est présente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_{0,T}, \boldsymbol{\Gamma}_0)$$

où $\boldsymbol{\mu}_{0,B}$ est le spectre moyen du fond, $\boldsymbol{\mu}_{0,T}$ le spectre moyen de la cible, et $\boldsymbol{\Gamma}_0$ la matrice de covariance du fond supposée égale sous les deux hypothèses (Manolakis et Shaw 2002). L'écriture du logarithme du rapport de vraisemblance mène ainsi à l'expression linéaire du filtre MF :

$$d_{MF}(\mathbf{r}) = (\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B})^t \boldsymbol{\Gamma}_0^{-1} (\mathbf{r} - \boldsymbol{\mu}_{0,B}). \quad (5.2)$$

En pratique, il est bien souvent nécessaire d'estimer la matrice de covariance du fond. Le test du rapport de vraisemblance est alors remplacé par le GLRT, qui ne possède plus les mêmes propriétés d'optimalité, mais qui est cependant largement utilisé. On utilise alors l'AMF, la version adaptative du filtre MF, où ce paramètre statistique est remplacé par son EMV $\hat{\boldsymbol{\Gamma}}$ (Robey et al. (1992)).

$$d_{AMF}(\mathbf{r}) = \frac{\left[(\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B})^t \hat{\boldsymbol{\Gamma}}_0^{-1} (\mathbf{r} - \boldsymbol{\mu}_{0,B}) \right]^2}{(\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B})^t \hat{\boldsymbol{\Gamma}}_0^{-1} (\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B})}. \quad (5.3)$$

Le filtre ACE a été développé avec une approche subpixel ; il est basé sur les hypothèses suivantes :

$$H_0 : \text{la cible est absente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_{0,B}, \boldsymbol{\Gamma}_0)$$

$$H_1 : \text{la cible est présente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_{0,T}, \sigma^2 \boldsymbol{\Gamma}_0).$$

Ici, les matrices de covariance des classes fond et cible ont la même structure, et le facteur de proportionnalité inconnu σ^2 est relié au taux de remplissage de la cible dans le pixel. Le rapport de vraisemblance généralisé conduit au détecteur ACE (Conte et al. 1995) :

$$d_{ACE}(\mathbf{r}) = \frac{\left[(\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B})^t \hat{\boldsymbol{\Gamma}}_0^{-1} (\mathbf{r} - \boldsymbol{\mu}_{0,B}) \right]^2}{\left[(\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B})^t \hat{\boldsymbol{\Gamma}}_0^{-1} (\boldsymbol{\mu}_{0,T} - \boldsymbol{\mu}_{0,B}) \right] \left[(\mathbf{r} - \boldsymbol{\mu}_{0,B})^t \hat{\boldsymbol{\Gamma}}_0^{-1} (\mathbf{r} - \boldsymbol{\mu}_{0,B}) \right]}. \quad (5.4)$$

L'AMF s'avère être très intéressant lorsque la cible est résolue, et que le fond est globalement homogène. Il possède une sensibilité au spectre cible plus importante que l'ACE par exemple. En revanche, ce dernier est également moins sensible aux outliers, et permet donc d'obtenir de meilleures performances de détection lorsque le fond est hétérogène (ce qui est souvent le cas en situation réelle). Sa construction fait qu'il est aussi adapté aux problèmes de détection subpixel (lorsqu'un pixel est composé d'un mélange de fond et de cible).

Les trois filtres présentés possèdent la caractéristique d'être à taux de fausse alarme constant (CFAR) ; par ailleurs, un résumé de l'ensemble de leurs propriétés est décrit par Manolakis (2004).

De nombreux autres algorithmes de détection ont ainsi été développés ces dernières années pour faire face à différents cas pratiques. Cependant,

aucun d'entre eux n'est adapté à la détection d'objets immergés. En effet, lors de la traversée des rayons lumineux dans la colonne d'eau, les spectres du fond et de la cible observés en surface peuvent être considérablement déformés, notamment dans le visible et le proche-infrarouge. Cette déformation dépend non seulement de la profondeur, mais aussi du degré de turbidité de l'eau. Afin de conserver de bonnes performances, il est donc nécessaire de prendre en compte et de modéliser cette distortion avant de développer les filtres de détection.

Pour cela, les deux modèles présentés dans la Section 4.2.1 sont utilisés. Comme pour l'estimation, le modèle M_1 est étudié en simulations, tandis que le modèle M_2 est testé sur les données réelles. Tous les développements théoriques des filtres bathymétriques sont effectués avec M_1 , mais ils peuvent être aisément réécrits avec M_2 .

Nous adoptons ici la même approche statistique que pour la méthode d'estimation présentée dans la Section 4.2.2. Chaque pixel spectral \mathbf{r} est ainsi considéré comme une réalisation d'un vecteur aléatoire gaussien multivarié, paramétré par le modèle M_1 ou M_2 .

Deux ensembles de données sont utilisés dans le processus de détection : un échantillon d'apprentissage, constitué uniquement de pixels de fond, et un échantillon de pixels tests (cible ou fond).

Comme pour les filtres de détection classiques, l'écriture d'un test du rapport de vraisemblance conduit à différentes fonctions selon les hypothèses et les connaissances *a priori* dont on dispose. Dans les sections suivantes, on transpose les filtres MF, AMF, et ACE dans le contexte de la détection de cibles immergées. De plus, nous proposons un nouveau filtre issu du test de vraisemblance généralisé, permettant de détecter sans avoir besoin de posséder de connaissances *a priori* sur la colonne d'eau.

5.2 QUELQUES FILTRES CLASSIQUES ADAPTÉS À LA DÉTECTION DE CIBLES IMMERGÉES

Pour chaque pixel test, on considère les hypothèses suivantes :

$$H_0 : \text{la cible est absente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_B, \boldsymbol{\Gamma})$$

$$H_1 : \text{la cible est présente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_T, \boldsymbol{\Gamma}).$$

Ici, on a $\boldsymbol{\mu}_B = \mathbf{r}_\infty + \mathbf{K}(\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty)$ et $\boldsymbol{\mu}_T = \mathbf{r}_\infty + \mathbf{K}(\boldsymbol{\mu}_{0,T} - \mathbf{r}_\infty)$ où $\boldsymbol{\mu}_{0,B}$ et $\boldsymbol{\mu}_{0,T}$ sont les spectres moyens du fond et de la cible respectivement. Tous les paramètres sont supposés connus.

Comme pour le filtre MF classique de l'Equation 5.2, le logarithme du rapport de vraisemblance conduit à l'expression du filtre adapté bathymétrique :

$$d_{BMF}(\mathbf{r}) = (\boldsymbol{\mu}_T - \boldsymbol{\mu}_B)^t \boldsymbol{\Gamma}^{-1} (\mathbf{r} - \boldsymbol{\mu}_B). \quad (5.5)$$

On peut insérer le modèle dans les hypothèses des autres détecteurs classiques, et ainsi, la méthode décrite par Robey et al. (1992) nous conduit à

l'expression du filtre adapté adaptatif bathymétrique (BAMF) :

$$d_{BAMF}(\mathbf{r}) = \frac{\left[(\boldsymbol{\mu}_T - \boldsymbol{\mu}_B)^t \hat{\boldsymbol{\Gamma}}^{-1} (\mathbf{r} - \boldsymbol{\mu}_B) \right]^2}{(\boldsymbol{\mu}_T - \boldsymbol{\mu}_B)^t \hat{\boldsymbol{\Gamma}}^{-1} (\boldsymbol{\mu}_T - \boldsymbol{\mu}_B)}. \quad (5.6)$$

En ce qui concerne le détecteur angulaire adaptatif bathymétrique (BACE), on peut aussi l'exprimer comme le carré du cosinus de l'angle entre entre le pixel test et le pixel cible, les données ayant été blanchies au préalable (Manolakis et Shaw 2002) :

$$\begin{aligned} d_{BACE}(\mathbf{r}) &= \cos^2 \left(\hat{\boldsymbol{\Gamma}}^{-\frac{1}{2}} (\mathbf{r} - \boldsymbol{\mu}_B), \hat{\boldsymbol{\Gamma}}^{-\frac{1}{2}} (\boldsymbol{\mu}_T - \boldsymbol{\mu}_B) \right) \\ &= \frac{\left[(\boldsymbol{\mu}_T - \boldsymbol{\mu}_B)^t \hat{\boldsymbol{\Gamma}}^{-1} (\mathbf{r} - \boldsymbol{\mu}_B) \right]^2}{\left[(\boldsymbol{\mu}_T - \boldsymbol{\mu}_B)^t \hat{\boldsymbol{\Gamma}}^{-1} (\boldsymbol{\mu}_T - \boldsymbol{\mu}_B) \right] \left[(\mathbf{r} - \boldsymbol{\mu}_B)^t \hat{\boldsymbol{\Gamma}}^{-1} (\mathbf{r} - \boldsymbol{\mu}_B) \right]}. \end{aligned} \quad (5.7)$$

Ces trois filtres bathymétriques possèdent les mêmes propriétés de taux de fausse alarme constant que leur homologue non-bathymétrique.

5.3 LE GBF (GLRT-BASED BATHYMETRIC FILTER), UNE SOLUTION LORSQUE LES PARAMÈTRES DE LA COLONNE D'EAU SONT INCONNUS

En pratique, les paramètres physiques tels que la profondeur ou la concentration en phytoplancton sont inconnus. Cependant, ils sont indispensables pour corriger les spectres du fond et de la cible avant le processus de détection (pour passer de $\boldsymbol{\mu}_{0,T}$ à $\boldsymbol{\mu}_T$ dans les Equations 4.1 et 4.2). Le but est donc de développer un filtre issu du rapport de vraisemblance généralisé, et fonctionnant de façon autonome en estimant l'ensemble de paramètres $\boldsymbol{\theta}$ par MV sur l'ensemble des données (pixel test \mathbf{r} et échantillon d'apprentissage $\chi_N = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$).

Par exemple, si la matrice de covariance et la profondeur sont inconnues, les hypothèses de départ sont les suivantes :

$$H_0 : \text{la cible est absente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_B, \boldsymbol{\Gamma})$$

$$H_1 : \text{la cible est présente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_T, \boldsymbol{\Gamma})$$

où $\boldsymbol{\mu}_B$ et $\boldsymbol{\mu}_T$ sont inconnus du fait de la méconnaissance de H . Hormis la matrice $\boldsymbol{\Gamma}$, tous les autres paramètres sont supposés connus.

Le test du rapport de vraisemblance généralisé (GLRT) calculé sur l'ensemble des données est donc :

$$D_{GBF}(\mathbf{r}) = \frac{\max_{\boldsymbol{\theta}} P(\mathbf{r}, \chi_N; \boldsymbol{\theta} | H_1)_{H_1}}{\max_{\boldsymbol{\theta}} P(\mathbf{r}, \chi_N; \boldsymbol{\theta} | H_0)_{H_0}} \stackrel{\geq}{\geq} \eta, \text{ avec } \boldsymbol{\theta} = \{\boldsymbol{\Gamma}, H\}. \quad (5.8)$$

Le développement de ce détecteur est détaillé en Annexe A.1, et aboutit à l'expression :

$$D_{GBF}(\mathbf{r}) = \frac{1 + (\mathbf{r} - \boldsymbol{\mu}_B(\hat{H}))^t \mathbf{S}(\hat{H})^{-1} (\mathbf{r} - \boldsymbol{\mu}_B(\hat{H}))}{1 + (\mathbf{r} - \boldsymbol{\mu}_T(\hat{H}))^t \mathbf{S}(\hat{H})^{-1} (\mathbf{r} - \boldsymbol{\mu}_T(\hat{H}))}. \quad (5.9)$$

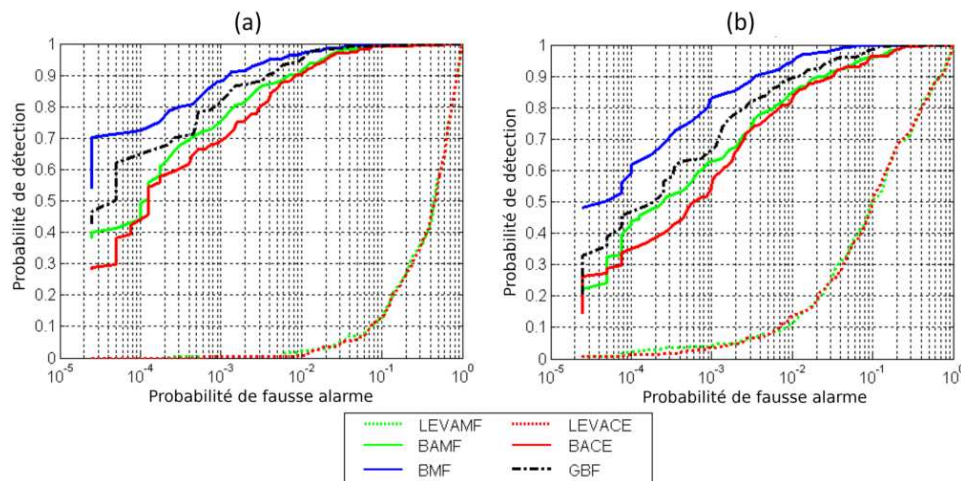


FIGURE 5.1 – Courbes ROC illustrant la détection du métal galvanisé en : (a) eau pure ($H = 55$ m, $SNR = 5.6$ dB); (b) eau turbide ($H = 14$ m, $C_{PHY} = 0.7$ $\mu\text{g.L}^{-1}$, $C_{CDOM} = 0.08$ m^{-1} et $C_{NAP} = 2.8$ mg.L^{-1} , $SNR = 9.9$ dB).

De la même façon, il est possible de développer le GBF dans le cas où la profondeur et les différentes concentrations sont inconnues :

$$D_{GBF}(\mathbf{r}) = \frac{\max_{\theta} P(\mathbf{r}, \chi_N; \theta | H_1)_{H_1}}{\max_{\theta} P(\mathbf{r}, \chi_N; \theta | H_0)_{H_0}} \geq \eta, \quad \text{avec } \theta = \{\Gamma, H, C_{PHY}, C_{CDOM}, C_{NAP}\}. \quad (5.10)$$

A partir de l'expression 5.9, nous n'avons pu établir la propriété de taux de fausse alarme constant, le paramètre \hat{H} intervenant dans les expressions de $\mu_B(\hat{H})$, $\mu_T(\hat{H})$, et $S(\hat{H})$ étant estimé numériquement.

5.4 RÉSULTATS

En pratique, le processus de détection peut s'effectuer de la manière suivante : en supposant que les pixels de cible soient suffisamment peu nombreux pour ne pas affecter l'estimation, on estime les paramètres inconnus dans le voisinage du pixel test, avant d'appliquer le filtrage.

Afin d'évaluer les performances des filtres bathymétriques, nous les comparons avec la méthode suivante : nous inversons les données avec les valeurs des paramètres estimées par la technique d'inversion présentée dans la Section 3.4, puis nous appliquons les détecteurs classiques (AMF, ACE) aux images ainsi transformées.

L'échantillon d'apprentissage est un cube de 21×21 pixels, de manière à ce que les EMVs sous chaque hypothèse H_0 et H_1 soient égaux (voir Annexe A.1).

On présente les résultats de détection sous la forme de courbes ROC (Receiver Operating Characteristic) : celles-ci sont des courbes paramétrées par le seuil η du test du rapport de vraisemblance 5.1. Pour chaque valeur de η , on calcule les probabilités de détection (P_D) et de fausse alarme

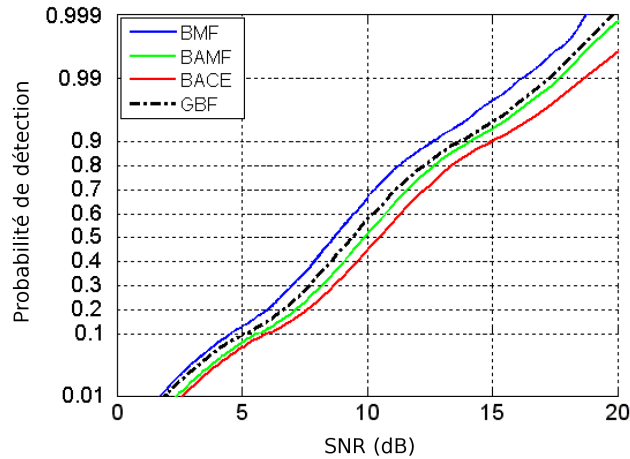


FIGURE 5.2 – Probabilité de détection du métal galvanisé en fonction du SNR ($H = 14$ m, $C_{PHY} = 0.7 \mu\text{g}\cdot\text{L}^{-1}$, $C_{CDOM} = 0.08 \text{ m}^{-1}$ et $C_{NAP} = 2.8 \text{ mg}\cdot\text{L}^{-1}$). La P_{FA} est fixée à 10^{-4} .

(P_{FA}) :

$$P_D = \frac{\text{Nombre de pixels de cibles détectés}}{\text{Nombre total de pixels de cible}} \quad (5.11)$$

$$P_{FA} = \frac{\text{Nombre de pixels de fond détectés}}{\text{Nombre total de pixels de fond}} \quad (5.12)$$

La probabilité de détection est donc la probabilité de bien détecter un pixel de cible, tandis que la probabilité de fausse alarme est la probabilité de détecter un pixel de fond comme étant un pixel de cible.

Sur la Figure 5.1.a, on compare les performances de chaque filtre pour la détection du métal galvanisé (cible présente sur la première ligne sur la Figure 4.5) à une profondeur de 55 m et un SNR de 5.6 dB. D'un point de vue général, on peut voir clairement que les filtres bathymétriques sont supérieurs aux détecteurs classiques appliqués après inversion. Pour une P_{FA} de 10^{-3} , les filtres bathymétriques détectent environ 80% des pixels de cible, tandis que le LEVAMF ou le LEVACE n'en détectent aucun. Bien entendu, ces performances dépendent notamment des différences entre les spectres de la cible et du fond, et de la turbidité de l'eau. Pour une eau plus trouble (voir Figure 5.1.b), la probabilité de détection est d'environ 0.7 pour une P_{FA} de 10^{-3} à 14 m et pour un SNR de 9.9 dB.

Ces courbes présentant des résultats pour un SNR fixe, il serait intéressant d'étudier aussi la robustesse de ces détecteurs ; en effet, on pourrait craindre que le GBF soit sensible aux forts niveaux de bruit, car la précision de l'estimation diminue avec le SNR. Sur la Figure 5.2, on a donc tracé l'évolution de la probabilité de détection du métal galvanisé en fonction du SNR pour une profondeur fixée à $H = 14$ m. On peut voir que le BMF et le GBF donnent de meilleurs résultats que le BAMF ou le BACE, et ce même sur images très bruitées. Le GBF est plus robuste que le BAMF puisqu'en effet, si les performances des deux filtres semblent équivalentes pour un SNR élevé, celles du GBF décroissent moins vite que celle du BAMF lorsque le SNR diminue. Par ailleurs, ces résultats peuvent être combinés avec ceux du Tableau 4.2 : on voit que même si la REQM

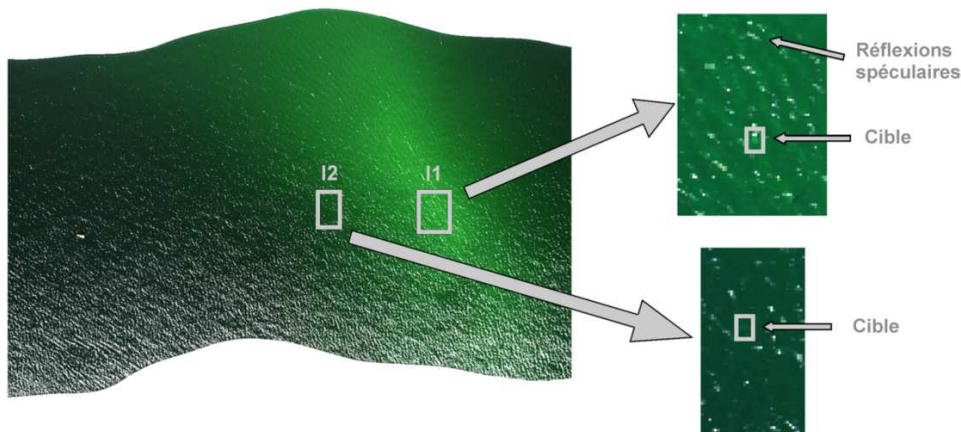


FIGURE 5.3 – Image HypLitt H3-0500m-Lo1-270deg-20100914-13h15mn. Une bâche blanche et une bâche noire ont été placées respectivement en I_1 et I_2 . La profondeur en I_1 est de 4.70 m tandis qu'elle est de 6.70 m en I_2 .

sur l'estimation de C_{PHY} et C_{CDOM} est respectivement de 17% et 11% pour un SNR de 5 dB, le GBF reste équivalent aux autres filtres. Nous en concluons que ce nouveau détecteur est robuste aux erreurs d'estimation sur les paramètres inconnus, et peut être utilisé sans aucune connaissance *a priori* sur la colonne d'eau.

Nous avons également implémenté nos algorithmes sur les données HypLitt présentées sur la Figure 5.3. Ceux-ci ont été développés avec le modèle M2; de plus, une série de résultats a été obtenue pour chaque modèle d'absorption du phytoplancton. Pour des profondeurs inférieures à 4.70 m (sites S_1 et S_2), les cibles sont très bien détectées, et il n'y a pas de différences entre les différentes méthodes.

En ce qui concerne le site S_3 , les courbes ROC présentées sur la Figure 5.4 ont été tracées en utilisant le modèle de Brando et al. (2009) et Bricaud et al. (1995). La tendance globale confirme la supériorité des filtres bathymétriques. Pour de telles profondeurs, les bâches blanches sont bien détectées par chaque filtre, puisque pour une P_{FA} de 10^{-3} , 100% des pixels de cible sont retrouvés. En ce qui concerne les bâches noires qui sont plus difficiles à détecter, le BAMF et le BACE sont sensiblement plus performants que le LEVAMF et le LEVACE respectivement, en I_1 et I_2 . Dans ce cas, on peut également remarquer que le GBF est le filtre le plus efficace : en effet, en I_1 et I_2 et pour une P_{FA} de 10^{-3} , la probabilité de détection est d'environ 0.8, tandis qu'elle est inférieure à 0.2 pour les autres filtres. Cela confirme donc le fait que le GBF est plus performant que ses concurrents lorsque les conditions sont difficiles, et ce, malgré les erreurs d'estimation sur la profondeur et les concentrations des constituants de l'eau.

Par ailleurs, afin d'étudier encore davantage l'influence du choix du modèle sur nos méthodes et nos résultats, on effectue les mêmes tests en utilisant cette fois-ci le modèle d'absorption du phytoplancton de Lee (1994). Les courbes ROC obtenues sont présentées sur la Figure 5.5. On voit bien cette modélisation permet d'améliorer globalement les résultats de tous les filtres. Cela est d'autant plus vrai pour ceux effectuant une pré-étape d'estimation avant la détection (i.e. le GBF, le LEVAMF, et le LEVACE). Par exemple, le GBF détecte désormais plus de 80% des pixels

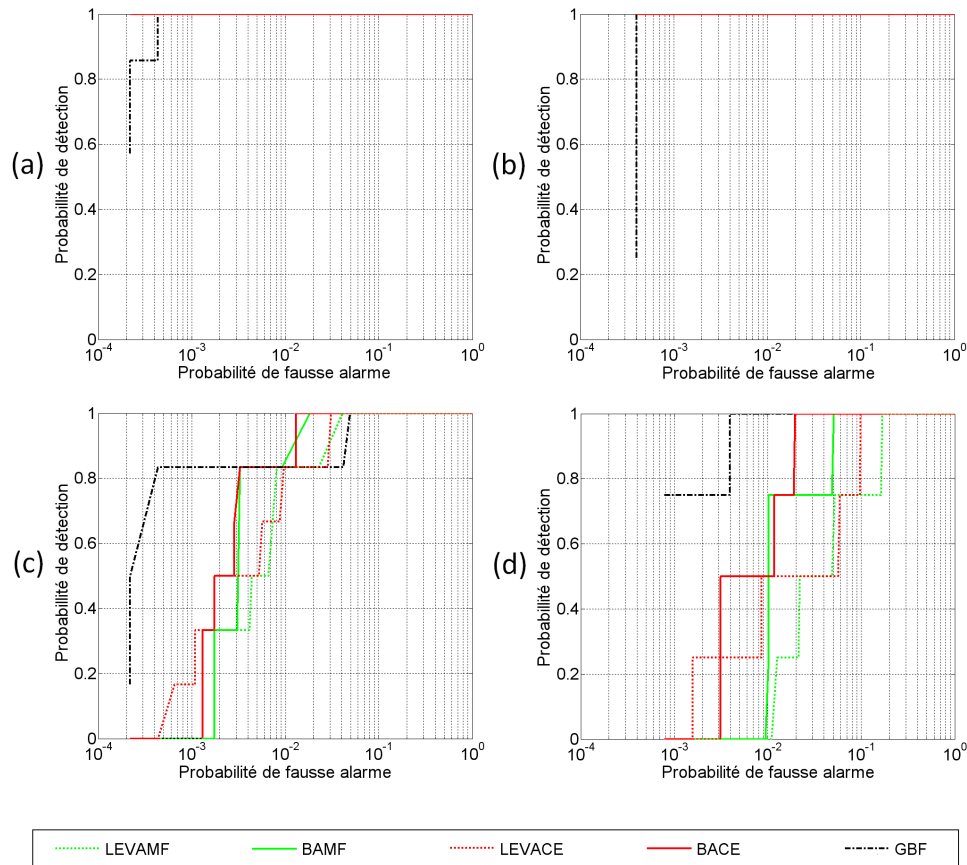


FIGURE 5.4 – Courbes ROC illustrant la détection de : (a) bache blanche en I_1 : $H = 4.70$ m, $C_{PHY} = 1.5 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2.5 \text{ mg.L}^{-1}$; (b) bache blanche en I_2 : $H = 6.70$ m, $C_{PHY} = 1.95 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2 \text{ mg.L}^{-1}$; (c) bache noire en I_1 : $H = 4.70$ m, $C_{PHY} = 1.5 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2.5 \text{ mg.L}^{-1}$; (d) bache noire en I_2 : $H = 6.70$ m, $C_{PHY} = 1.95 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2 \text{ mg.L}^{-1}$. Ici, le modèle d'absorption du phytoplancton est celui de Brando et al. (2009) et Bricaud et al. (1995).

de cible sur l'ensemble des images étudiées. La seule légère baisse de résultats lors du passage au modèle de Lee (1994) concerne la détection de la bache noire à 6.70 m par le BAMF et le BACE. On peut expliquer cela par le fait qu'il est nécessaire de connaître la valeur exacte des paramètres H , C_{PHY} , C_{CDOM} , et C_{NAP} pour appliquer ces filtres; or, comme nous l'avons déjà présenté dans la Section 4.4.2, et plus particulièrement dans le Tableau 4.7, nous n'avons qu'une connaissance approximative de C_{NAP} , et aucune connaissance sur C_{CDOM} . Les valeurs utilisées dans le BAMF et le BACE ne sont donc peut-être pas les valeurs réelles de ces concentrations, ce qui induit cette petite baisse des performances.

Par ailleurs, nous avons mesuré les temps de calcul pour les deux méthodes. Avec un processeur de 3 GHz et pour les méthodes de détection basées sur l'estimation par inversion, il est de 12 ms et 13 ms par pixel en I_1 et I_2 respectivement. En comparaison, nos algorithmes ont besoin de 4 ms et 2.4 ms par pixel en I_1 et I_2 respectivement. Bien entendu, il serait possible d'améliorer encore davantage ces derniers chiffres en utilisant un schéma d'optimisation plus rapide comme les techniques de gradient. Finalement, ces derniers résultats obtenus sur les images HypLitt

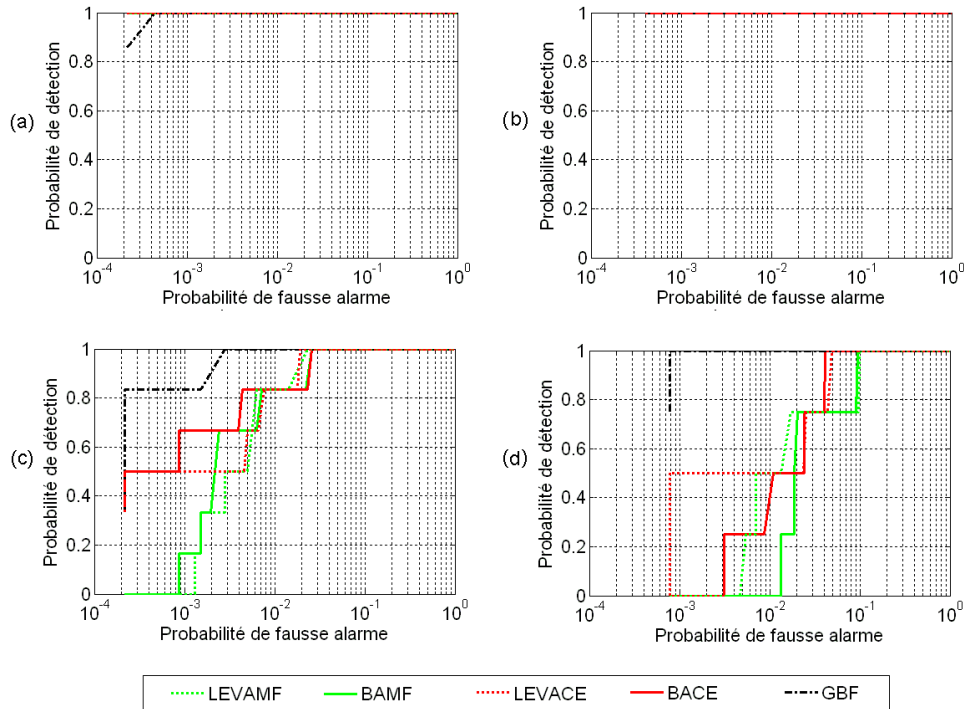


FIGURE 5.5 – Courbes ROC illustrant la détection de : (a) bêche blanche en I_1 : $H = 4.70$ m, $C_{PHY} = 1.5 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2.5 \text{ mg.L}^{-1}$; (b) bêche blanche en I_2 : $H = 6.70$ m, $C_{PHY} = 1.95 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2 \text{ mg.L}^{-1}$; (c) bêche noire en I_1 : $H = 4.70$ m, $C_{PHY} = 1.5 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2.5 \text{ mg.L}^{-1}$; (d) bêche noire en I_2 : $H = 6.70$ m, $C_{PHY} = 1.95 \mu\text{g.L}^{-1}$, et $C_{NAP} < 2 \text{ mg.L}^{-1}$. Ici, le modèle d'absorption du phytoplancton est celui de Lee (1994).

confirment l'intérêt de ces filtres bathymétriques pour la détection de cibles immergées.

5.5 CONCLUSIONS

Dans cette partie, nous transposons les filtres MF, AMF et ACE dans le contexte de la détection de cibles immergées. Ces détecteurs supposant que les caractéristiques de la colonne d'eau sont connues, nous développons le GBF, un nouveau filtre issu du test de vraisemblance généralisé, qui estime au préalable les paramètres inconnus. Les résultats obtenus sur les simulations et sur les images HypLitt sont supérieurs à ceux obtenus par les méthodes classiques, consistant à inverser la colonne d'eau avec les paramètres estimés, avant d'appliquer les détecteurs usuels. On constate notamment sur les données simulées que le GBF permet d'obtenir des performances équivalentes aux autres filtres bathymétriques, et que par conséquent, l'estimation par maximum de vraisemblance est suffisamment performante pour que cela n'influe pas sur les performances de détection. Sur les données réelles, lorsque la connaissance des paramètres caractéristiques de la colonne d'eau est imparfaite, le GBF est le détecteur le plus performant. Enfin, nous avons étudié l'influence du modèle d'absorption du phytoplancton sur les résultats de détection : la modélisation de Lee (1994) est encore celle qui permet d'obtenir les meilleurs taux de détection, confirmant ainsi la supériorité face à la représentation

de Brando et al. (2009) et Bricaud et al. (1995) déjà constatée dans le Chapitre 4.

CONCLUSIONS ET PERSPECTIVES

6

6.1 CONCLUSIONS

CETTE thèse aborde deux problématiques : la première est relative à la caractérisation de la colonne d'eau et l'estimation des paramètres inconnus associés. La deuxième traite de la détection de cibles immergées en zones côtières. Pour cela, après le Chapitre d'introduction 1, nous présentons dans les Chapitres 2 et 3 un état de l'art sur les modèles physiques utilisés, ainsi que sur les méthodes d'estimation des paramètres de la colonne d'eau existantes. Dans le Chapitre 4, nous présentons les différentes méthodes d'estimation statistiques proposées. Enfin, le Chapitre 5 traite de la problématique de la détection de cibles immergées.

Dans le Chapitre 2, nous abordons la question cruciale du choix du modèle bathymétrique, reliant la réflectance subsurface à la réflectance du fond et aux paramètres "optiquement actifs" caractéristiques de la colonne d'eau (profondeur, concentrations en phytoplancton, en matière organique dissoute, et en particules non algales). Ce choix est très important car le succès de nos méthodes repose en grande partie sur la fidélité de ce modèle à la réalité. Par conséquent, il est donc primordial d'effectuer l'état de l'art le plus complet possible. En environnement côtier, la réflectance subsurface peut être vue comme la somme pondérée de la réflectance du fond et de la réflectance d'une colonne d'eau infinie. Ces coefficients de pondération dépendent de la profondeur et de la turbidité de l'eau, notamment à travers les spectres d'absorption et de rétro-diffusion. Ces derniers peuvent s'écrire eux-mêmes comme la somme des absorptions ou des rétro-diffusions de chaque constituant de l'eau optiquement actif, c'est-à-dire les molécules d'eau, le phytoplancton, le CDOM, et les particules non algales. Afin d'évaluer l'impact d'un modèle plus précis sur les résultats d'estimation, nous avons sélectionné plusieurs représentations de la réflectance, ainsi que plusieurs façons de modéliser l'absorption du phytoplancton.

Dans le Chapitre 3, nous présentons un ensemble de méthodes existantes estimant tout ou partie des paramètres caractéristiques de la colonne d'eau. Nous choisissons ensuite l'une d'entre elles afin d'évaluer par comparaison les performances de nos algorithmes. Deux approches sont utilisées dans les méthodes d'estimation globales. La première consiste à minimiser une fonction d'erreur calculée à partir de l'observation et

du spectre modélisé. Les paramètres estimés sont ceux pour lesquels la mesure et le modèle sont les plus proches au sens de l'erreur quadratique moyenne. Dans la deuxième approche, on construit une librairie spectrale tabulée en fonction de la profondeur et des concentrations. L'observation est comparée à cet ensemble de spectres, et les paramètres estimés sont alors ceux pour lesquels la distance entre la mesure et le modèle est minimale (méthode des moindres carrés). La taille de cette librairie doit être bien choisie afin de représenter une majorité de situations réelles tout en gardant des temps de calcul raisonnables.

Afin d'évaluer les performances de nos algorithmes, nous choisissons de travailler avec la méthode d'inversion par minimisation d'une fonction d'erreur, adaptée à la modélisation choisie.

Dans le Chapitre 4, nous présentons les différentes méthodes d'estimation statistiques que nous avons développées au cours de cette thèse. Celles-ci se basent sur l'hypothèse de gaussianité des données et sur la paramétrisation de la réflectance à l'aide des modèles évoqués dans le Chapitre 2. Dans un premier temps, nous considérons un échantillon d'apprentissage de réalisations indépendantes et identiquement distribuées, et nous calculons la vraisemblance associée. Cette dernière dépend des paramètres à estimer, et les estimations numériques sont alors les valeurs les plus probables, c'est-à-dire celles qui maximisent cette fonction. Sur les simulations, les résultats obtenus sont très intéressants, notamment dans une gamme de profondeurs comprises entre 1 m et 10 m à 20 m selon la turbidité de l'eau. Cette méthode montre également une bonne robustesse au bruit. Grâce aux tests sur les données HypLitt, on voit que l'EMV est globalement supérieure à l'inversion par minimisation d'une fonction d'erreur pour l'estimation de la profondeur. Cette dernière étude permet de souligner l'importance du choix d'un modèle réaliste, particulièrement pour la méthode EMV, puisque les résultats obtenus avec le modèle M2 sont notablement meilleurs que ceux obtenus avec M1.

Nous avons envisagé la cas réaliste où le spectre du fond est connu à un facteur multiplicatif près, en intégrant ce paramètre dans le modèle des données, et en supposant la matrice de covariance diagonale et de variance inconnue. Nous montrons que, malgré l'hypothèse simplificatrice sur la forme de cette matrice, l'estimation du vecteur de paramètres inconnu est notablement améliorée pour des hauteurs d'eau faibles, là où la méthode générale donne de moins bons résultats (lorsque le spectre de fond supposé est différent du spectre réel). Pour des profondeurs plus importantes, cette modélisation n'est plus efficace puisque la forme spectrale du fond n'a alors que peu d'importance, la réflectance subsurface étant quasiment égale à la réflectance d'une colonne d'eau infinie.

Les performances obtenues par l'EMV sont donc satisfaisantes. Cependant, l'implémentation de cette méthode suppose que l'on puisse disposer d'un échantillon suffisamment grand, ce qui est facilement envisageable en zones côtières lorsque l'on travaille avec des résolutions spatiales comme celles des données HypLitt. Il est donc intéressant d'affiner le processus d'estimation en modélisant les variations de profondeur, qui est *a priori* le paramètre qui peut évoluer le plus rapidement au sein de l'image. Nous avons alors inséré un modèle de pente linéaire dans la

vraisemblance : le spectre moyen du fond et sa matrice de covariance sont exprimés en fonction de H_{ij} , la hauteur d'eau dans le pixel (i,j) . Les résultats montrent que cela permet d'améliorer significativement l'estimation de tous les paramètres (y compris les concentrations), notamment lorsque les profondeurs considérées sont peu importantes (jusqu'à quelques mètres). Lorsque les variations du relief sont faibles par rapport à H , la modélisation par un fond plat suffit pour décrire convenablement la scène. Cette technique a été testée et validée sur les données simulées et sur les images HypLitt.

Ces résultats étant satisfaisants, nous avons étendu cette méthode à l'estimation d'une pente linéaire par morceaux. Celle-ci se décompose en deux étapes : dans la première, nous recherchons la zone de rupture de pente (crête ou vallée) en étudiant le profil spectral des pixels dans chaque direction. Puis, une fois que les deux régions sont repérées, nous estimons la profondeur dans chacune d'entre elles, et ce, de façon indépendante ou non (en contraignant ou non les deux plans à se rejoindre). Là encore, les résultats obtenus sur les données simulées et les images HypLitt sont très intéressants et montrent que cette méthode permet d'affiner l'estimation de la profondeur en zones côtières.

Par ailleurs, toujours dans un souci d'améliorer le processus d'estimation, nous avons voulu travailler sur les gammes de profondeurs où l'EMV est un peu moins performante, c'est-à-dire les très faibles ou les très grandes hauteurs d'eau. Dans le premier cas, l'atténuation est principalement causée par l'eau pure, et par conséquent, il est délicat de retrouver l'effet des constituants. Dans le deuxième cas, le fond n'a quasiment plus d'influence sur la réflectance subsurface, ce qui rend difficile l'estimation de la profondeur. Dans ces situations, il est donc intéressant, dans la mesure du possible, d'insérer un *a priori* sur la répartition des paramètres à estimer. Dans un premier temps, nous avons régularisé l'expression de la vraisemblance en introduisant des distributions gaussiennes centrées sur les vraies valeurs de la profondeur et des concentrations. Dans un deuxième temps, nous avons utilisé des lois Gamma pour modéliser la variabilité des paramètres C_{PHY} et C_{NAP} . Les résultats obtenus sur les simulations confirment l'intérêt de cette approche, puisque cela permet d'éviter les valeurs aberrantes et d'améliorer significativement la précision de l'estimation lorsque les conditions sont difficiles (faibles et fortes profondeurs, faible SNR). Sur les images HypLitt, l'amélioration est surtout notable sur l'estimation de H , puisque par exemple, l'erreur obtenue avec le modèle d'absorption du phytoplancton de Lee (1994) est diminuée de moitié pour des hauteurs d'eau supérieures ou égales à 12 m. Pour l'estimation des concentrations, la baisse est moins significative car d'une part, il est moins difficile de retrouver ces paramètres dans cette gamme de profondeurs et par conséquent, il y a moins d'écart entre MV et MAP, et d'autre part, les écarts-types choisis ou estimés à partir des vérités terrain disponibles sont trop grands pour que cela ait une réelle influence sur les résultats. En ce qui concerne les différences entre les deux régularisations proposées, bien que les tests d'adéquation du χ^2 prouvent que les lois Gamma permettent de mieux représenter la variabilité de C_{PHY} et C_{NAP} que les lois gaussiennes, les résultats obtenus dans les deux cas sont très proches. Il serait toutefois intéressant de comparer ces deux méthodes en

eau peu profonde, là où l'estimation des concentrations est plus difficile. En conclusion, cette approche, lorsqu'elle est possible, permet d'améliorer les performances d'estimation lorsque les conditions sont délicates. Enfin, pour compléter nos travaux sur l'estimation des paramètres de l'eau, nous avons calculé les bornes de Cramér-Rao associées. Ce puissant outil statistique donne une expression théorique de la variance minimale d'un estimateur non biaisé, et son étude permet de savoir s'il est possible ou non d'améliorer la précision de nos méthodes. Sous certaines hypothèses, notamment sur la forme de la matrice de covariance, on peut ainsi calculer ces bornes et les comparer aux variances expérimentales. Une étude préliminaire sur l'évolution du biais est nécessaire afin de savoir quelle est la taille minimale de l'échantillon permettant d'avoir des estimateurs non biaisés. Une fois cela effectué, la comparaison entre les BCRs et les variances nous montre que l'EMV permet d'obtenir des estimateurs efficaces des paramètres lorsque le nombre de pixels est suffisant, et qu'au sens du critère "non biaisé et à variance minimale", toute autre méthode donnera des résultats au mieux égaux, sinon inférieurs. Ceci nous permet de prévoir la précision optimale que l'on peut attendre pour l'estimation des paramètres. Enfin, le tracé de ces BCRs en fonction de H , C_{PHY} , C_{CDOM} , et C_{NAP} est intéressant et nous renseigne sur l'évolution de la précision de ces différents paramètres pour différentes colonnes d'eau. Cet outil nous permet par ailleurs de valider la méthode d'optimisation numérique de la vraisemblance que nous avons choisie.

Dans le Chapitre 5, nous abordons la problématique de la détection de cibles immergées en environnement côtier. Dans un premier temps, nous insérons les modèles présentés dans les chapitres précédents dans les filtres classiquement utilisés (MF, AMF, ACE), ce qui conduit à l'écriture de filtres dits bathymétriques (BMF, BAMF, BACE). L'utilisation de tels détecteurs nécessitant de connaître la nature de la colonne d'eau, ce qui est délicat en pratique, nous avons également développé le GBF, un nouveau filtre bathymétrique issu du test de vraisemblance généralisé calculé à partir du pixel test et d'un ensemble de pixels d'apprentissage. Les paramètres inconnus sont estimés au préalable par MV, et sont ensuite réinsérés dans l'expression du détecteur. Comme pour les méthodes d'estimation, nous avons testé ces algorithmes sur les images simulées et les données HypLitt. Dans les deux cas, on voit que d'une part, les filtres bathymétriques sont supérieurs aux détecteurs classiques obtenus en combinant la méthode d'inversion par minimisation d'une fonction d'erreur et le filtrage par AMF ou ACE. D'autre part, le GBF est intéressant par rapport au BAMF et au BACE, lorsque les conditions de détection sont difficiles : en particulier, il s'avère être robuste à des niveaux de bruit importants, ce qui prouve que l'effet des erreurs d'estimation induites par un faible SNR ont peu d'influence sur les performances de détection. Par ailleurs, dans la lignée du chapitre précédent, nous évaluons l'influence du modèle d'absorption du phytoplancton sur les performances de détection. D'une façon générale, nous voyons là encore que celui de Lee (1994) permet d'obtenir les meilleurs résultats, et s'affirme ainsi comme la modélisation la plus fiable.

Pour finir, nous avons comparé les temps de calcul, et là encore, nous

voyons que nos algorithmes sont plus performants, puisqu'ils sont au moins trois fois plus rapides que les filtres classiques (2.4 ms à 4 ms par pixel contre 12 ms à 13 ms). Cette différence provient essentiellement de la méthode d'estimation, l'inversion par minimisation d'une fonction d'erreur étant plus lente que l'EMV proposée.

Nous pouvons conclure que bien que les méthodes statistiques soient peu utilisées dans le cadre de la bathymétrie et de l'analyse de la composition de l'eau, elles présentent un intérêt certain, tant que la résolution spatiale reste suffisante pour réaliser l'estimation sur un échantillon.

6.2 PERSPECTIVES

Plusieurs pistes restent à explorer, notamment sur l'estimation des paramètres.

L'une des plus importantes concerne le choix des modèles physiques. On a vu que cet aspect était primordial et impactait considérablement les résultats d'estimation. La complexification du modèle bathymétrique de réflectance (par exemple, l'utilisation du modèle de Albert et Mobley (2003)), l'affinage de l'influence de chaque constituant, ou un regroupement différent des effets des composants (par exemple, regrouper la rétro-diffusion du phytoplancton et des particules non algales en un seul terme) peuvent représenter des perspectives intéressantes.

En ce qui concerne l'estimation d'une profondeur variable, on peut envisager d'introduire une modélisation non linéaire du relief, et pour cela, l'utilisation de champs de Markov pourrait être intéressante.

La méthode du MAP appliquée ici pourrait également faire l'objet de modifications. Il serait intéressant d'étudier par exemple l'inter-dépendance des paramètres C_{PHY} , C_{CDOM} , et C_{NAP} , et d'insérer des lois conjointes dans le calcul de la fonction objectif. Par ailleurs, il serait intéressant de pouvoir comparer les méthodes du MAP avec les distributions gaussiennes et Gamma sur des données réelles en eau peu profonde : en effet, c'est dans cette gamme de hauteurs que l'estimation des concentrations est la plus délicate et donc que les plus grandes améliorations sont attendues. C'est également là *a priori* que l'on pourrait décider laquelle des deux lois gaussienne ou Gamma permet d'obtenir les meilleurs résultats.

Concernant les BCRs, on pourrait réaliser la même étude en choisissant des modèles plus complexes pour la réflectance et la matrice de covariance.

Une perspective plus générale est relative au choix du modèle gaussien pour modéliser la variabilité des données hyperspectrales. Bien que celui-ci permette d'obtenir des résultats satisfaisants, on pourrait envisager de développer toutes les méthodes présentées dans cette thèse avec le modèle elliptique, qui est connu pour mieux représenter les données hyperspectrales, notamment au niveau des queues de distribution.

Enfin, jusqu'ici, l'influence de la colonne d'eau s'est traduite par une transformation affine de la réflectance du fond. Des modélisations plus fines prenant d'éventuelles interactions non linéaires entre pixels voisins pourraient être prises en compte dans le futur pour développer de nouveaux filtres bathymétriques.

ANNEXES

A

A.1 UN NOUVEAU FILTRE BATHYMÉTRIQUE BASÉ SUR LE GLRT (GBF)

Si certains paramètres (H , C_{PHY} , C_{CDOM} et/ou C_{NAP}) sont inconnus, nous ne pouvons pas utiliser le BMF, le BAMF ou le BACE, car nous en avons besoin afin de corriger les spectres du fond et de la cible. Nous avons donc développé un nouveau filtre permettant de surmonter cette difficulté en estimant au préalable les paramètres inconnus. Nous montrons le développement de ce détecteur dans le cas où seules la matrice de covariance et la profondeur sont inconnues, mais il est facile de faire de même lorsque nous n'avons de plus aucune connaissance *a priori* sur les concentrations.

Les hypothèses de départ sont donc les suivantes :

$$H_0 : \text{la cible est absente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_B, \boldsymbol{\Gamma})$$

$$H_1 : \text{la cible est présente} \Rightarrow \mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_T, \boldsymbol{\Gamma}).$$

où $\boldsymbol{\mu}_B$ et $\boldsymbol{\mu}_T$ sont inconnus du fait de la méconnaissance de H . Hormis la matrice $\boldsymbol{\Gamma}$, tous les autres paramètres sont supposés connus.

Nous pouvons ainsi écrire le GLRT suivant :

$$D_{GBF}(\mathbf{r}) = \frac{\max_{\boldsymbol{\theta}} \text{P}(\mathbf{r}, \chi_N; \boldsymbol{\theta} | H_1)}_{H_1}}{\max_{\boldsymbol{\theta}} \text{P}(\mathbf{r}, \chi_N; \boldsymbol{\theta} | H_0)}_{H_0}} \stackrel{\geq}{\eta}, \text{ avec } \boldsymbol{\theta} = \{\boldsymbol{\Gamma}, H\}. \quad (\text{A.1})$$

Suivant la méthodologie développée par Kelly (1986), nous estimons d'abord la matrice de covariance sous chaque hypothèse. Pour H_0 et H_1 , nous obtenons respectivement :

$$\hat{\boldsymbol{\Gamma}}_{H_0}(\mathbf{r}, \chi_N; H) = \frac{1}{N+1} [(\mathbf{r} - \boldsymbol{\mu}_B(H))(\mathbf{r} - \boldsymbol{\mu}_B(H))^t + \mathbf{S}(H)] \quad (\text{A.2})$$

$$\hat{\boldsymbol{\Gamma}}_{H_1}(\mathbf{r}, \chi_N; H) = \frac{1}{N+1} [(\mathbf{r} - \boldsymbol{\mu}_T(H))(\mathbf{r} - \boldsymbol{\mu}_T(H))^t + \mathbf{S}(H)] \quad (\text{A.3})$$

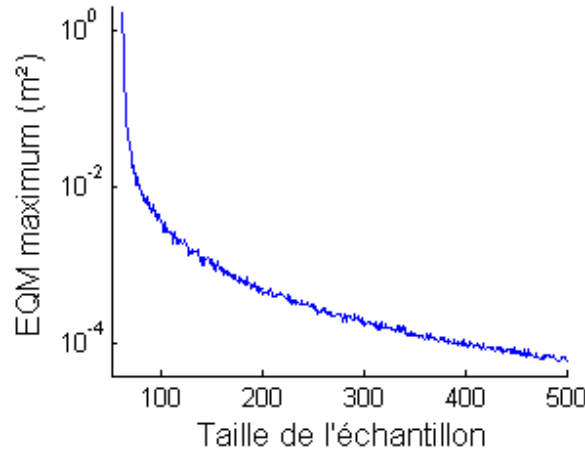


FIGURE A.1 – EQM maximum de la différence entre la profondeur estimée sous les hypothèses H_0 ou H_1 , et la profondeur estimée \hat{H} calculée uniquement sur l'échantillon d'apprentissage, en fonction de la taille de l'échantillon N ($H = 10$ m, eau turbide).

où $\mathbf{S}(H) = \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}_B(H))(\mathbf{r}_i - \boldsymbol{\mu}_B(H))^t$ est calculé à partir de l'échantillon d'apprentissage χ_N . Le GLR devient alors :

$$D_{GBF}(\mathbf{r}) = \frac{\max_H |\hat{\mathbf{T}}_{H_1}(\mathbf{r}, \chi_N; H)|^{-1}}{\max_H |\hat{\mathbf{T}}_{H_0}(\mathbf{r}, \chi_N; H)|^{-1}}. \quad (\text{A.4})$$

On peut montrer que, si la taille de l'échantillon N est assez grand, il est équivalent d'optimiser la fonction de vraisemblance sous les hypothèses H_0 ou H_1 , ou même uniquement sur les données d'apprentissage. On peut vérifier cela sur la Figure A.1, sur laquelle est tracée en fonction de N , l'erreur quadratique moyenne maximale de la différence entre la profondeur estimée sous l'hypothèse H_0 ou H_1 , et la profondeur \hat{H} estimée uniquement sur l'échantillon χ_N . On peut voir que lorsque N est environ sept fois plus grand que le nombre de bandes, l'EQM est inférieure à 10^{-4} m^2 , soit une erreur de moins de 1 cm. Nous pouvons ainsi considérer que les estimateurs du MV \hat{H}_{H_1} et \hat{H}_{H_0} sont presque égaux à \hat{H} .

En utilisant le lemme pour le calcul du déterminant utilisé par Kelly (1986), nous obtenons finalement le GLR suivant :

$$D_{GBF}(\mathbf{r}) = \frac{1 + (\mathbf{r} - \boldsymbol{\mu}_B(\hat{H}))^t \mathbf{S}(\hat{H})^{-1} (\mathbf{r} - \boldsymbol{\mu}_B(\hat{H}))}{1 + (\mathbf{r} - \boldsymbol{\mu}_T(\hat{H}))^t \mathbf{S}(\hat{H})^{-1} (\mathbf{r} - \boldsymbol{\mu}_T(\hat{H}))}. \quad (\text{A.5})$$

A.2 MÉTHODE D'ESTIMATION LORSQUE LE FOND EST INCONNU

La fonction de log-vraisemblance est donnée à partir de l'Equation 4.5 par :

$$l(\chi_N|\theta) = -\frac{N}{2} \ln \left[(2\pi)^L |\Gamma| \right] - \frac{1}{2} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu})^t \Gamma^{-1} (\mathbf{r}_i - \boldsymbol{\mu}). \quad (\text{A.6})$$

Ici, le spectre moyen et la matrice de covariance sont donnés par :

$$\boldsymbol{\mu} = \mathbf{r}_\infty + \mathbf{K}(f\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty) \quad (\text{A.7})$$

$$\Gamma = \sigma^2(\mathbf{K}^2 + \mathbf{Id}_L). \quad (\text{A.8})$$

où \mathbf{r}_∞ et \mathbf{K} dépendent de $\boldsymbol{\theta} = \{H, C_{PHY}, C_{CDOM}, C_{NAP}\}$.

Dans un premier temps, on recherche l'EMV du facteur multiplicatif f . Pour cela, on a besoin de connaître la dérivée de $\boldsymbol{\mu}$ par rapport à f . Celle-ci est simplement donnée par :

$$\frac{\partial \boldsymbol{\mu}}{\partial f} = \mathbf{K}\boldsymbol{\mu}_{0,B}. \quad (\text{A.9})$$

Le spectre moyen étant le seul paramètre statistique dépendant de f , la dérivée de la log-vraisemblance est donc égale à :

$$\frac{\partial l(\chi_N|\theta)}{\partial f} = -\frac{1}{2} \sum_{i=1}^N \frac{\partial}{\partial f} \left[\Gamma^{-1/2} (\mathbf{r}_i - \boldsymbol{\mu}) \right]^2 \quad (\text{A.10})$$

$$\frac{\partial l(\chi_N|\theta)}{\partial f} = -\frac{1}{2} \sum_{i=1}^N 2 \left[\Gamma^{-1/2} \left(-\frac{\partial \boldsymbol{\mu}}{\partial f} \right) \right]^t \Gamma^{-1/2} (\mathbf{r}_i - \boldsymbol{\mu}) \quad (\text{A.11})$$

$$\frac{\partial l(\chi_N|\theta)}{\partial f} = \sum_{i=1}^N \left(\mathbf{K}\boldsymbol{\mu}_{0,B} \right)^t \Gamma^{-1} (\mathbf{r}_i - \boldsymbol{\mu}) \quad (\text{A.12})$$

$$\frac{\partial l(\chi_N|\theta)}{\partial f} = N \left[\boldsymbol{\mu}_{0,B}^t \mathbf{K} \Gamma^{-1} (\bar{\mathbf{r}}(\chi_N) + (\mathbf{K} - \mathbf{Id}_L) \mathbf{r}_\infty) - f \boldsymbol{\mu}_{0,B}^t \mathbf{K} \Gamma^{-1} \mathbf{K} \boldsymbol{\mu}_{0,B} \right].$$

En égalant $\frac{\partial l(\chi_N|\theta)}{\partial f}$ à zéro, on obtient ainsi l'expression analytique de l'EMV de f :

$$\hat{f}_{MV}(\chi_N; \boldsymbol{\theta}) = \frac{\boldsymbol{\mu}_{0,B}^t \mathbf{K} \Gamma^{-1} (\bar{\mathbf{r}}(\chi_N) + (\mathbf{K} - \mathbf{Id}_L) \mathbf{r}_\infty)}{\boldsymbol{\mu}_{0,B}^t \mathbf{K} \Gamma^{-1} \mathbf{K} \boldsymbol{\mu}_{0,B}} \quad (\text{A.13})$$

$$\hat{f}_{MV}(\chi_N; \boldsymbol{\theta}) = \frac{\boldsymbol{\mu}_{0,B}^t \mathbf{K} \mathbf{G}^{-1} (\bar{\mathbf{r}}(\chi_N) + (\mathbf{K} - \mathbf{Id}_L) \mathbf{r}_\infty)}{\boldsymbol{\mu}_{0,B}^t \mathbf{K} \mathbf{G}^{-1} \mathbf{K} \boldsymbol{\mu}_{0,B}} \quad (\text{A.14})$$

où $\mathbf{G} = \text{diag} \left[1 + e^{-4k(\lambda_i)H} \right]_{l \in \llbracket 1; L \rrbracket}$.

On remplace f par son EMV dans l'expression de la log-vraisemblance, qui devient alors :

$$l(\chi_N|\boldsymbol{\theta}) = -\frac{N}{2} \ln \left[(2\pi)^L |\Gamma(\boldsymbol{\theta})| \right] - \frac{1}{2} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \Gamma(\boldsymbol{\theta})^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})) \quad (\text{A.15})$$

où $\boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}) = \mathbf{r}_\infty + \mathbf{K}(\widehat{f}_{MV}(\chi_N; \boldsymbol{\theta})\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty)$.

Dans un deuxième temps, on recherche l'expression de l'EMV de σ^2 . Pour cela, on a besoin de calculer la dérivée de $\Gamma(\boldsymbol{\theta})$ par rapport à σ^2 , et plus particulièrement, celle de $\Gamma(\boldsymbol{\theta})^{-1}$ par rapport à σ^2 .

$$\frac{\partial \Gamma(\boldsymbol{\theta})^{-1}}{\partial \sigma^2} = \frac{\partial}{\partial \sigma^2} \left\{ \text{diag} \left[\frac{1}{\sigma^2 (1 + e^{-4k(\lambda_l)H})} \right]_{l \in \llbracket 1; L \rrbracket} \right\} \quad (\text{A.16})$$

$$\frac{\partial \Gamma(\boldsymbol{\theta})^{-1}}{\partial \sigma^2} = \text{diag} \left[-\frac{1}{\sigma^4 (1 + e^{-4k(\lambda_l)H})} \right]_{l \in \llbracket 1; L \rrbracket} \quad (\text{A.17})$$

$$\frac{\partial \Gamma(\boldsymbol{\theta})^{-1}}{\partial \sigma^2} = -\sigma^{-4} \mathbf{G}^{-1}. \quad (\text{A.18})$$

En ce qui concerne $\boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})$, on voit facilement qu'il ne dépend pas de σ^2 , et par conséquent, on a :

$$\begin{aligned} \frac{\partial l(\chi_N | \boldsymbol{\theta})}{\partial \sigma^2} &= -\frac{N}{2} \underbrace{\frac{\partial}{\partial \sigma^2} [\ln |\Gamma(\boldsymbol{\theta})|]}_{\textcircled{\text{A}}} \\ &\quad - \underbrace{\frac{1}{2} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \frac{\partial \Gamma(\boldsymbol{\theta})^{-1}}{\partial \sigma^2} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))}_{\textcircled{\text{B}}}. \end{aligned} \quad (\text{A.19})$$

Dans un premier temps, on calcule le terme $\textcircled{\text{A}}$:

$$\textcircled{\text{A}} = \frac{\partial}{\partial \sigma^2} \left\{ \ln \left[\prod_{i=1}^L \sigma^2 (1 + e^{-4k(\lambda_i)H}) \right] \right\} \quad (\text{A.20})$$

$$\textcircled{\text{A}} = \frac{\partial}{\partial \sigma^2} \left[L \ln \sigma^2 + \sum_{i=1}^L \ln (1 + e^{-4k(\lambda_i)H}) \right] \quad (\text{A.21})$$

$$\textcircled{\text{A}} = \frac{L}{\sigma^2}. \quad (\text{A.22})$$

Le terme $\textcircled{\text{B}}$ est lui donné par :

$$\textcircled{\text{B}} = -\frac{1}{\sigma^4} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})). \quad (\text{A.23})$$

La dérivée de la log-vraisemblance par rapport à σ^2 est donc égale à :

$$\frac{\partial l(\chi_N | \boldsymbol{\theta})}{\partial \sigma^2} = -\frac{N \times L}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})). \quad (\text{A.24})$$

En égalant $\frac{\partial l(\chi_N | \boldsymbol{\theta})}{\partial \sigma^2}$ à zéro, on obtient ainsi l'expression analytique de l'EMV de σ^2 :

$$\widehat{\sigma}_{MV}^2(\chi_N; \boldsymbol{\theta}) = \frac{1}{N \times L} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})). \quad (\text{A.25})$$

La log-vraisemblance devient :

$$l(\chi_N|\boldsymbol{\theta}) = -\frac{N}{2} \ln \left[(2\pi)^L |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})| \right] - \frac{1}{2} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})) \quad (\text{A.26})$$

$$l(\chi_N|\boldsymbol{\theta}) = -\frac{N}{2} \ln \left[(2\pi)^L |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})| \right] - \frac{1}{2} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \frac{\mathbf{G}^{-1}}{\hat{\sigma}_{MV}^2(\chi_N; \boldsymbol{\theta})} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})) \quad (\text{A.27})$$

$$l(\chi_N|\boldsymbol{\theta}) = -\frac{N}{2} \ln \left[(2\pi)^L |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})| \right] - \frac{N \times L \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))}{2 \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))} \quad (\text{A.28})$$

$$l(\chi_N|\boldsymbol{\theta}) = -\frac{N \times L}{2} [1 + \ln(2\pi)] - \frac{N}{2} \ln |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})|. \quad (\text{A.29})$$

Bien entendu, il est possible de développer les mêmes expressions avec le modèle M2, et dans ce cas-là, il s'agit d'optimiser la fonction :

$$l(\chi_N|\boldsymbol{\theta}) = -\frac{N \times L}{2} [1 + \ln(2\pi)] - \frac{N}{2} \ln |\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta})| \quad (\text{A.30})$$

avec

$$\boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}) = \mathbf{r}_\infty + \mathbf{K}_B \hat{f}_{MV}(\chi_N; \boldsymbol{\theta}) \boldsymbol{\mu}_{0,B} - \mathbf{K}_C \mathbf{r}_\infty \quad (\text{A.31})$$

$$\boldsymbol{\Gamma}(\chi_N; \boldsymbol{\theta}) = \hat{\sigma}_{MV}^2(\chi_N; \boldsymbol{\theta}) (\mathbf{K}_B^2 + \mathbf{Id}_L) \quad (\text{A.32})$$

$$\hat{f}_{MV}(\chi_N; \boldsymbol{\theta}) = \frac{\boldsymbol{\mu}_{0,B}^t \mathbf{K}_B \mathbf{G}^{-1} (\bar{\mathbf{r}}(\chi_N) + (\mathbf{K}_C - \mathbf{Id}_L) \mathbf{r}_\infty)}{\boldsymbol{\mu}_{0,B}^t \mathbf{K}_B \mathbf{G}^{-1} \mathbf{K}_B \boldsymbol{\mu}_{0,B}^t} \quad (\text{A.33})$$

$$\hat{\sigma}_{MV}^2(\chi_N; \boldsymbol{\theta}) = \frac{1}{N \times L} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta}))^t \mathbf{G}^{-1} (\mathbf{r}_i - \boldsymbol{\mu}(\chi_N; \boldsymbol{\theta})) \quad (\text{A.34})$$

où $\mathbf{G} = \text{diag} \left[1 + e^{-2(k_d(\lambda_l) + k_u^b(\lambda_l))H} \right]_{l \in \llbracket 1; L \rrbracket}$.

A.3 CALCUL DES BORNES DE CRAMÉR-RAO

A.3.1 Introduction

Dans le cas scalaire (c'est-à-dire $\boldsymbol{\theta} = \{\theta_1\}$), la BCR pour l'estimation de θ_1 est donnée par l'inverse de l'information de Fisher $\mathbf{I}_1(\boldsymbol{\theta})$:

$$\mathcal{BCR}_1(\boldsymbol{\theta})^{-1} = \mathbf{I}_1(\boldsymbol{\theta}) = \mathbb{E} \left[\left(\frac{\partial \ln(\mathbb{P}(\chi_N|\boldsymbol{\theta}))}{\partial \theta_1} \right)^2 \right]. \quad (\text{A.35})$$

Dans le cas multivarié (c'est-à-dire $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_P\}$), elle est égale à :

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = [\mathbf{I}_1(\boldsymbol{\theta})]_{i,j} = \mathbb{E} \left[\frac{\partial \ln(\mathbb{P}(\chi_N|\boldsymbol{\theta}))}{\partial \theta_i} \frac{\partial \ln(\mathbb{P}(\chi_N|\boldsymbol{\theta}))}{\partial \theta_j} \right]. \quad (\text{A.36})$$

Dans le cas gaussien, $\mathbf{I}_1(\boldsymbol{\theta})$ est donnée par la formule de Slepian-Bangs :

$$[\mathbf{I}_1(\boldsymbol{\theta})]_{i,j} = \frac{1}{2} \text{tr} \left(\underbrace{\Gamma^{-1} \frac{\partial \Gamma}{\partial \theta_i} \Gamma^{-1} \frac{\partial \Gamma}{\partial \theta_j}}_{\text{(A)}} + \underbrace{\frac{\partial \boldsymbol{\mu}^t}{\partial \theta_i} \Gamma^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \theta_j}}_{\text{(B)}} \right). \quad (\text{A.37})$$

Dans le cas scalaire, on a bien entendu $i = j = 1$.

Enfin, les paramètres statistiques $\boldsymbol{\mu}$ et Γ sont donnés ici par :

$$\boldsymbol{\mu} = \mathbf{r}_\infty + \mathbf{K}(\boldsymbol{\mu}_{0,B} - \mathbf{r}_\infty) \quad (\text{A.38})$$

$$\Gamma = \sigma_B^2 \mathbf{K}^2 + \sigma_S^2 \mathbf{Id}_L = \text{diag} \left[\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2 \right]_{l \in \llbracket 1;L \rrbracket}. \quad (\text{A.39})$$

Par la suite, nous avons besoin de calculer les dérivées des coefficients d'absorption et de rétro-diffusion (exprimés par les Equations 2.39 et 2.44) en fonction des trois concentrations C_{PHY} , C_{CDOM} , et C_{NAP} . En supposant ici que l'absorption du phytoplancton est donnée par la formulation 2.32, nous obtenons :

$$\frac{\partial a(\lambda)}{\partial C_{PHY}} = a_{PHY}^*(\lambda) \quad (\text{A.40})$$

$$\frac{\partial a(\lambda)}{\partial C_{CDOM}} = e^{-S_{CDOM}(\lambda - \lambda_0)} \quad (\text{A.41})$$

$$\frac{\partial a(\lambda)}{\partial C_{NAP}} = a_{NAP}^*(\lambda_0) e^{-S_{NAP}(\lambda - \lambda_0)} \quad (\text{A.42})$$

$$\frac{\partial b_b(\lambda)}{\partial C_{PHY}} = b_{b,PHY}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{PHY}} \quad (\text{A.43})$$

$$\frac{\partial b_b(\lambda)}{\partial C_{CDOM}} = 0 \quad (\text{A.44})$$

$$\frac{\partial b_b(\lambda)}{\partial C_{NAP}} = b_{b,NAP}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{NAP}}. \quad (\text{A.45})$$

Les dérivées du coefficient d'atténuation (exprimé par l'Equation 2.21) sont alors égales à :

$$\frac{\partial k(\lambda)}{\partial C_{PHY}} = \frac{1}{\mu_d} \left(a_{PHY}^*(\lambda) + b_{b,PHY}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{PHY}} \right) \quad (\text{A.46})$$

$$\frac{\partial k(\lambda)}{\partial C_{CDOM}} = \frac{1}{\mu_d} \left(e^{-S_{CDOM}(\lambda-\lambda_0)} \right) \quad (\text{A.47})$$

$$\frac{\partial k(\lambda)}{\partial C_{NAP}} = \frac{1}{\mu_d} \left(a_{NAP}^*(\lambda_0) e^{-S_{NAP}(\lambda-\lambda_0)} + b_{b,NAP}^*(\lambda_2) \left(\frac{\lambda_2}{\lambda} \right)^{Y_{NAP}} \right) \quad (\text{A.48})$$

A.3.2 Cas scalaire

La BCR pour l'estimation du paramètre inconnu est donc un scalaire donné par la formule A.35.

– **Cas $\theta = \{H\}$**

En ce qui concerne le calcul de \textcircled{A} , la dérivée de Γ est donnée par :

$$\frac{\partial \Gamma}{\partial H} = \text{diag} \left[-4k(\lambda_l) \sigma_B^2 e^{-4k(\lambda_l)H} \right]_{l \in \llbracket 1;L \rrbracket}. \quad (\text{A.49})$$

Quant à l'inverse Γ^{-1} , il est simplement donné par :

$$\Gamma^{-1} = \text{diag} \left[\frac{1}{\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2} \right]_{l \in \llbracket 1;L \rrbracket}. \quad (\text{A.50})$$

On a donc :

$$\textcircled{A} = \text{tr} \left[\text{diag} \left(\frac{8k(\lambda_l)^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} \right)_{l \in \llbracket 1;L \rrbracket} \right] \quad (\text{A.51})$$

$$\textcircled{A} = \sum_{l=1}^L \frac{8k(\lambda_l)^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2}. \quad (\text{A.52})$$

En ce qui concerne le calcul de \textcircled{B} , pour chaque longueur d'onde λ_l , on a :

$$\frac{\partial \mu(\lambda_l)}{\partial H} = -2k(\lambda_l) e^{-2k(\lambda_l)H} (\mu_{0,B}(\lambda_l) - r_\infty(\lambda_l)). \quad (\text{A.53})$$

\textcircled{B} est donc donné par :

$$\textcircled{B} = \sum_{l=1}^L \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right)^2 \left[\Gamma^{-1} \right]_{ll} \quad (\text{A.54})$$

$$\textcircled{B} = \sum_{l=1}^L \frac{4k(\lambda_l)^2 e^{-4k(\lambda_l)H} (\mu_{0,B}(\lambda_l) - r_\infty(\lambda_l))^2}{\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2}. \quad (\text{A.55})$$

Finalement, l'expression de la BCR est donnée par :

$$\begin{aligned} BCR_1(H)^{-1} &= \sum_{l=1}^L \frac{8k(\lambda_l)^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} \\ &\quad + \frac{4k(\lambda_l)^2 e^{-4k(\lambda_l)H} (\mu_{0,B}(\lambda_l) - r_\infty(\lambda_l))^2}{\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2}. \end{aligned} \quad (\text{A.56})$$

Dans le cas où l'on a $\sigma_B^2 = \sigma_S^2 = \sigma^2$, on obtient enfin :

$$\begin{aligned} \mathcal{BCR}_1(H)^{-1} &= \sum_{l=1}^L \frac{8k(\lambda_l)^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} \\ &+ \frac{4k(\lambda_l)^2 e^{-4k(\lambda_l)H} (\mu_{0,B}(\lambda_l) - r_\infty(\lambda_l))^2}{\sigma^2(1 + e^{-4k(\lambda_l)H})}. \end{aligned} \quad (\text{A.57})$$

– **Cas $\theta = \{C_c\}$, où C_c est l'une des trois concentrations C_{PHY} , C_{CDOM} , et C_{NAP}**

Ici, en ce qui concerne la calcul de \textcircled{A} , la dérivée de Γ est donnée par :

$$\frac{\partial \Gamma}{\partial C_c} = \text{diag} \left[-4 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) H \sigma_B^2 e^{-4k(\lambda_l)H} \right]_{l \in \llbracket 1;L \rrbracket}. \quad (\text{A.58})$$

Ici, en utilisant les dérivées de $k(\lambda_l)$ par rapport à chaque C_c présentées dans les Equations A.46, A.47, et A.48, \textcircled{A} est donc donné par :

$$\textcircled{A} = \text{tr} \left[\text{diag} \left(\frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right)^2 H^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} \right)_{l \in \llbracket 1;L \rrbracket} \right] \quad (\text{A.59})$$

$$\textcircled{A} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right)^2 H^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2}. \quad (\text{A.60})$$

En ce qui concerne le calcul de \textcircled{B} , pour chaque longueur d'onde λ_l , on a :

$$\begin{aligned} \frac{\partial \mu(\lambda_l)}{\partial C_c} &= \frac{\partial r_\infty(\lambda_l)}{\partial C_c} (1 - e^{-2k(\lambda_l)H}) \\ &- 2 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) H e^{-2k(\lambda_l)H} (\mu_{0,B}(\lambda_l) - r_\infty(\lambda_l)). \end{aligned} \quad (\text{A.61})$$

Pour dériver la réflectance d'une colonne d'eau infinie $r_\infty(\lambda_l)$ en fonction de C_c , nous avons besoin de dériver le rapport $\frac{b_b(\lambda_l)}{a(\lambda_l) + b_b(\lambda_l)}$ (par souci de lisibilité, la dépendance en λ_l est ici exprimée ponctuellement à travers l'indice l) :

$$\frac{\partial \frac{b_{b,l}}{a_l + b_{b,l}}}{\partial C_c} = \frac{a_l \frac{\partial b_{b,l}}{\partial C_c} - b_{b,l} \frac{\partial a_l}{\partial C_c}}{(a_l + b_{b,l})^2}. \quad (\text{A.62})$$

La dérivée de $r_{\infty,l}$ s'exprime alors telle que :

$$\frac{\partial r_{\infty,l}}{\partial C_c} = 0.084 \left(\frac{\partial \frac{b_{b,l}}{a_l + b_{b,l}}}{\partial C_c} \right) + 0.34 \left(\frac{\partial \frac{b_{b,l}}{a_l + b_{b,l}}}{\partial C_c} \right) \frac{b_{b,l}}{a_l + b_{b,l}} \quad (\text{A.63})$$

$$\frac{\partial r_{\infty,l}}{\partial C_c} = 0.084 \frac{a_l \frac{\partial b_{b,l}}{\partial C_c} - b_{b,l} \frac{\partial a_l}{\partial C_c}}{(a_l + b_{b,l})^2} + 0.34 \frac{a_l b_{b,l} \frac{\partial b_{b,l}}{\partial C_c} - b_{b,l}^2 \frac{\partial a_l}{\partial C_c}}{(a_l + b_{b,l})^3}. \quad (\text{A.64})$$

Pour calculer cette dérivée pour chaque concentration, il suffit d'utiliser les Equations 2.39, 2.44, A.40, A.41, A.42, A.43, A.44, et A.45, et de remplacer les termes correspondants dans la formulation A.64 ci-dessus. Ces termes étant très longs, nous ne les explicitons pas ici. \textcircled{B} est ainsi égal à :

$$\textcircled{B} = \sum_{l=1}^L \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right)^2 [\Gamma^{-1}]_{l,l} \quad (\text{A.65})$$

où $\frac{\partial \mu(\lambda_l)}{\partial C_c}$ et $[\Gamma^{-1}]_{l,l}$ sont donnés par les Equations A.61 et A.50 respectivement.

Finalement, l'expression (simplifiée) de la BCR est donnée par :

$$\text{BCR}_1(C_c)^{-1} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right)^2 H^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right)^2 [\Gamma^{-1}]_{l,l}. \quad (\text{A.66})$$

Dans le cas où l'on a $\sigma_B^2 = \sigma_S^2 = \sigma^2$, on obtient enfin :

$$\text{BCR}_1(C_c)^{-1} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right)^2 H^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right)^2 [\Gamma^{-1}]_{l,l}. \quad (\text{A.67})$$

A.3.3 Cas multivarié

Dans le cas multivarié, on calcule la matrice des BCRs dont chaque élément est donné par la formule A.36. Ici, les développements sont effectués pour $\boldsymbol{\theta} = \{H, C_{PHY}, C_{CDOM}, C_{NAP}\}$. Cependant, la matrice des BCRs de tout sous-ensemble de $\boldsymbol{\theta}$ sera simplement obtenue en supprimant la ou les lignes et la ou les colonnes correspondant aux paramètres connus.

En ce qui concerne les éléments diagonaux $[\text{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,i}$, ils sont égaux aux bornes précédemment calculées dans le cas scalaire.

En ce qui concerne les éléments hors-diagonaux $[\text{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j}$, lorsque $\theta_i = H$ et $\theta_j = C_c$, la formule de Slepian-Bangs devient :

$$[\mathbf{I}_1(\boldsymbol{\theta})]_{i,j} = \frac{1}{2} \text{tr} \left(\Gamma^{-1} \frac{\partial \Gamma}{\partial H} \Gamma^{-1} \frac{\partial \Gamma}{\partial C_c} \right) + \frac{\partial \mu}{\partial H} \Gamma^{-1} \frac{\partial \mu}{\partial C_c}. \quad (\text{A.68})$$

Tous les termes intervenant dans cette équation sont donnés précédemment dans les Equations A.50, A.49, A.58, A.53, et A.61, et on a donc :

$$\textcircled{A} = \sum_{l=1}^L \frac{8k(\lambda_l) \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) H \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} \quad (\text{A.69})$$

$$\textcircled{B} = \sum_{l=1}^L \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) [\Gamma^{-1}]_{l,l}. \quad (\text{A.70})$$

où $\frac{\partial k(\lambda_l)}{\partial C_c}$ est donné par les Equations A.46, A.47, et A.48.

D'où l'expression (simplifiée) des éléments de la matrice des BCRs lorsque

$\theta_i = H$ et $\theta_j = C_c$:

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8k(\lambda_l) \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) H \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l}. \quad (\text{A.71})$$

Lorsque $\sigma_B^2 = \sigma_S^2 = \sigma^2$, on a :

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8k(\lambda_l) \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) H e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial H} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l}. \quad (\text{A.72})$$

Par ailleurs, lorsque $\theta_i = C_c$ et $\theta_j = C_d$, la formule de Slepian-Bangs devient :

$$[\mathbf{I}_1(\boldsymbol{\theta})]_{i,j} = \frac{1}{2} \text{tr} \left(\boldsymbol{\Gamma}^{-1} \frac{\partial \boldsymbol{\Gamma}}{\partial C_c} \boldsymbol{\Gamma}^{-1} \frac{\partial \boldsymbol{\Gamma}}{\partial C_d} \right) + \frac{\partial \boldsymbol{\mu}^t}{\partial C_c} \boldsymbol{\Gamma}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial C_d} \quad (\text{A.73})$$

et on a :

$$\textcircled{\text{A}} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial k(\lambda_l)}{\partial C_d} \right) H^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} \quad (\text{A.74})$$

$$\textcircled{\text{B}} = \sum_{l=1}^L \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_d} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l}. \quad (\text{A.75})$$

L'expression des éléments de la matrice des BCRs lorsque $\theta_i = C_c$ et $\theta_j = C_d$ est donc :

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial k(\lambda_l)}{\partial C_d} \right) H^2 \sigma_B^4 e^{-8k(\lambda_l)H}}{(\sigma_B^2 e^{-4k(\lambda_l)H} + \sigma_S^2)^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_d} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l} \quad (\text{A.76})$$

qui devient, lorsque $\sigma_B^2 = \sigma_S^2 = \sigma^2$:

$$[\mathcal{BCR}_1(\boldsymbol{\theta})^{-1}]_{i,j} = \sum_{l=1}^L \frac{8 \left(\frac{\partial k(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial k(\lambda_l)}{\partial C_d} \right) H^2 e^{-8k(\lambda_l)H}}{(1 + e^{-4k(\lambda_l)H})^2} + \left(\frac{\partial \mu(\lambda_l)}{\partial C_c} \right) \left(\frac{\partial \mu(\lambda_l)}{\partial C_d} \right) [\boldsymbol{\Gamma}^{-1}]_{l,l}. \quad (\text{A.77})$$

LISTE DES PRODUCTIONS SCIENTIFIQUES

Revue internationale avec comité de lecture :

- S. Jay, M. Guillaume, J. Blanc-Talon. "Underwater target detection with hyperspectral data : solutions for both known and unknown water quality", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5 :1213-1221, 2012.

Conférences internationales avec actes et comité de lecture :

- S. Jay, M. Guillaume. "Estimation of water column parameters with a maximum likelihood approach", *Proc. of the 3rd IEEE Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, pages 1-4, 2011.
- S. Jay, M. Guillaume. "Underwater target detection with hyperspectral remote-sensing imagery", *Proc. of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 2820-2823, 2010.

Colloque national sans comité de lecture :

- M. Guillaume, S. Jay, O. Eches, M. Lennon, G. Sicot, N. Vigneau, J. Blanc-Talon, V. Serfaty, M. Lennon. "Evaluation des capacités de la télédétection HYPerspectrale et développement de méthodes innovantes de traitement d'images pour des applications Défense en zone LITtorale (HYPLITT)", *2^{ème} colloque scientifique de la Société Française de Télédétection Hyperspectrale*, 2012.

BIBLIOGRAPHIE

- S.M. Adler-Golden, P.K. Acharya, A. Berk, M.W. Matthew, et D. Gorodetzky. Remote bathymetry of the littoral zone from AVIRIS, LASH and QuickBird imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 43(2) :337–347, 2005. (Cité page 1.)
- Y.H. Ahn, A. Bricaud, et A. Morel. Light backscattering efficiency and related properties of some phytoplankters. *Deep-Sea Research*, 39 :1835–1855, 1992. (Cité page 16.)
- A. Albert et C.D. Mobley. An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters. *Optics Express*, 11(22) :2873–2890, 2003. (Cité pages 7, 10, 13, 14, 16, 17, 18, 19, 21 et 83.)
- L. Bertels, T. Vanderstraete, S. Van Coillie, E. Knaeps, S. Sterckx, R. Goossens, et B. Deronde. Mapping of coral reefs using hyperspectral CASI data ; a case study : Fordata, tanimbar, indonesia. *International journal of remote sensing*, 29(8) :2359–2391, 2008. (Cité pages 8 et 12.)
- V.E. Brando, J.M. Anstee, M. Wettle, A.G. Dekker, S.R. Phinn, et C. Roelfsema. A physics based retrieval and quality assessment of bathymetry from suboptimal hyperspectral data. *Remote Sensing of Environment*, 113 : 755–770, 2009. (Cité pages xv, 1, 13, 16, 17, 18, 19, 20, 21, 23, 28, 30, 39, 47, 48, 49, 53, 54, 61, 75, 76 et 78.)
- A. Bricaud, M. Babin, A. Morel, et H. Claustre. Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton : analysis and parameterization. *Journal of geophysical research*, 100(C7) : 13321–13332, 1995. (Cité pages xv, 17, 19, 23, 47, 48, 49, 53, 54, 61, 75, 76 et 78.)
- H. Buiteveld, J.H.M. Hakvoort, et M. Donze. The optical properties of pure water. Dans *Ocean Optics XII*, pages 174–183, 1994. (Cité pages 16 et 19.)
- G.S. Campbell et J.M. Norman. *An introduction to environmental biophysics*. Springer, 1998. ISBN 0-387-94937-2. (Cité page 15.)
- J.P. Cannizzaro et K.L. Carder. Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters. *Remote Sensing of Environment*, 101 :13–24, 2006. (Cité pages 8, 10, 13, 16, 17, 18, 19, 20, 21, 26 et 55.)
- K.L. Carder, F.R. Chen, J.P. Cannizzaro, J.W. Campbell, et B.G. Mitchell. Performance of the MODIS semi-analytical ocean color algorithm for chlorophyll-a. *Advances in Space Research*, 33 :1152–1159, 2004. (Cité pages 16, 17, 18 et 19.)

- Y. Cherruault et G. Mora. *Optimisation globale, Théorie des courbes alpha-denses*. Economica, 2006. (Cité page 37.)
- E. Christophe. *Compression des images hyperspectrales et son impact sur la qualité des données*. PhD thesis, Ecole Nationale Supérieure de l'Aéronautique et de l'Espace, 2006. (Cité page 2.)
- R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, et S.J. Sutley. Usgs digital spectral library splib06a : U.s. geological survey, data series 231. <http://speclab.cr.usgs.gov/spectral-lib.html>, 2007. (Cité page 38.)
- M. Cococcioni, B. Lazzerini, et F. Marcelloni. Estimating the concentration of optically active constituents of sea water by takagi-sugeno models with quadratic rule consequents. *Pattern Recognition*, 40 :2846–2860, 2007. (Cité page 9.)
- E. Conte, M Lops, et G. Ricci. Asymptotically optimum radar detection in compound-gaussian clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 31 :617–625, 1995. (Cité page 70.)
- G. Dall'Olmo, A.A. Gitelson, D.C. Rundquist, B. Leavitt, T. Barrow, et J.C. Holz. Assessing the potential of seawiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. *Remote Sensing of Environment*, 96 :176–187, 2005. (Cité page 26.)
- A.G. Dekker, S.R. Phinn, J. Anstee, P. Bissett, V.E. Brando, B. Casey, P. Fearn, J. Hedley, W. Klonowski, Z.P. Lee, M. Lynch, M. Lyons, C. Mobley, et C. Roelfsema. Intercomparison of shallow water bathymetry, hydro-optics, and benthos mapping techniques in australian and caribbean coastal environments. *Limnology and Oceanography*, 9 :396–425, 2011. (Cité pages 27 et 28.)
- A.G. Dekker, R.J. Vos, et S.W.M. Peters. Comparison of remote sensing data, model results and in situ data for total suspended matter TSM in the southern Frisian lakes. *The science of the total environment*, 268 :197–214, 2001. (Cité pages 9 et 26.)
- D. Doxaran, J.M. Froidefond, Castaing P., et Babin M. Dynamics of the turbidity maximum zone in a macrotidal estuary (the gironde, france) : Observations from field and modis satellite data. *Estuarine, Coastal and Shelf Science*, 81 :312–332, 2009. (Cité page 59.)
- P.R.C. Fearn, W. Klonowski, R.C. Babcock, P. England, et J. Phillips. Shallow water substrate mapping using hyperspectral remote sensing. *Continental Shelf Research*, 31 :1249–1259, 2011. (Cité pages 1, 8, 13, 14, 16, 17, 18, 20 et 21.)
- P. Forget, S. Ouillon, et F. Lahet. Modélisation et inversion des spectres de réflectance d'eaux turbides côtières. *Physical measurements and signatures in remote sensing*, 2 :495–500, 1997. (Cité pages 9, 16, 18, 19 et 26.)

- P. Gege. The water color simulator WASI : an integrating software tool for analysis and simulation of optical in situ spectra. *Computers and Geosciences*, 30 :523–532, 2004. (Cité pages 8, 14, 16, 17, 18 et 19.)
- C. Gerardino-Neira, J. Goodman, M. Velez-Reyes, et W. Rivera. Sensitivity analysis of a hyperspectral inversion model for remote sensing of shallow coastal ecosystems. 2008. (Cité pages 8 et 13.)
- C. Giardino, V.E. Brando, A.G. Dekker, N. Strömbeck, et G. Candiani. Assessment of water quality in lake garda (italy) using hyperion. *Remote Sensing of Environment*, 109 :183–195, 2007. (Cité pages 8, 9, 10, 16, 17, 18, 19 et 21.)
- A.A. Gitelson, J.F. Schalles, et C.M. Hladik. Remote chlorophyll-a retrieval in turbid, productive estuaries : Chesapeake bay case study. *Remote Sensing of Environment*, 109 :464–472, 2007. (Cité pages 1 et 26.)
- H.R. Gordon. Can the lambert-beer law be applied to the diffuse attenuation coefficient of ocean water ? *Limnology and Oceanography*, 34(8) : 1389–1409, 1989. (Cité pages 13 et 14.)
- H.R. Gordon, O.B. Brown, R.H. Evans, J.W. Brown, R.C. Smith, K.S. Baker, et D.K. Clark. A semianalytic radiance model of ocean color. *Journal of Geophysical Research*, 93 :10909–10924, 1988. (Cité pages 9 et 10.)
- H.R. Gordon, O.B. Brown, et M.M. Jacobs. Computed relations between the inherent and apparent optical properties of a flat homogeneous ocean. *Applied Optics*, 14 :417–427, 1975. (Cité page 9.)
- H. Hakvoort, J. De Haan, R. Jordans, R. Vos, S. Peters, et M. Rijkeboer. Towards airborne remote sensing of water quality in the netherlands - validation and error analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57 :171–183, 2002. (Cité pages 9, 17, 18, 21 et 27.)
- L. Han. Estimating chlorophyll-a concentration using first-derivative spectra in coastal water. *International Journal of Remote Sensing*, 26(23) :5235–5244, 2005. (Cité page 27.)
- J. Hedley, C. Roelfsema, et S.R. Phinn. Efficient radiative transfer model inversion for remote sensing applications. *Remote Sensing of Environment*, 113 :2527–2532, 2009. (Cité page 29.)
- T. Heege, P. Hausknecht, et H. Kobryn. Hyperspectral seafloor mapping and direct bathymetry calculation using hymap data from the ningaloo reef and rotnest island areas in western australia. Dans *Proc. of the 5th EARSeL Workshop in Imaging Spectroscopy*, 2007. (Cité pages 1 et 26.)
- IFREMER. Site ifremer consacré à l'environnement littoral. <http://www.ifremer.fr/envlit/>, 2011. (Cité pages 56, 58, 60 et 61.)
- J.H. Jerome, R.P. Bukata, et J.R. Miller. Remote sensing reflectance and its relationship to optical properties of natural waters. *International Journal of remote sensing*, 17 :3135–3155, 1996. (Cité page 9.)

- K. Kallio, T. Kutser, T. Hannonen, S. Koponen, J. Pulliainen, J. Vepsäläinen, et T. Pyhälähti. Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons. *The Science of the Total Environment*, 268 :59–77, 2001. (Cité pages 25 et 26.)
- P.A. Keller. Comparison of two inversion techniques of a semi-analytical model for the determination of lake water constituents using imaging spectrometry data. *The Science of the Total Environment*, 268 :189–196, 2001. (Cité pages 9, 17, 18, 19 et 27.)
- E.J. Kelly. An adaptive detection algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, AES-22(1) :115–127, 1986. (Cité pages 87 et 88.)
- J.T.O. Kirk. Monte carlo study of the nature of the underwater light held in, and the relationships between optical properties of, turbid yellow waters. *Australian Journal of Marine and Freshwater Research*, 32 :517–532, 1981. (Cité page 13.)
- J.T.O. Kirk. Dependence of relationship between inherent and apparent optical properties of water on solar altitude. *Limnology and Oceanography*, 29 :350–356, 1984. (Cité pages 9 et 13.)
- S. Kirkpatrick, C.D. Gelatt, et M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598) :671–680, 1983. (Cité page 37.)
- W.M. Klonowski, P.R.C.S. Fearn, et M.J. Lynch. Retrieving key benthic cover types and bathymetry from hyperspectral imagery. *Journal of Applied Remote Sensing*, 1, 2007. (Cité page 28.)
- F.A. Kruse, A.B. Lefkoff, J.W. Boardman, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon, et A.F.H. Goetz. The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44 :145–163, 1993. (Cité page 28.)
- T. Kutser. *Estimation of water quality in turbid inland and coastal waters by passive optical remote sensing*. Tartu University Press, 1997. (Cité page 18.)
- T. Kutser. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnology and Oceanography*, 49 :2179–2189, 2004. (Cité pages 8, 9, 16, 18, 19, 26 et 55.)
- T. Kutser, A. Herlevi, K. Kallio, et H. Arst. A hyperspectral model for interpretation of passive optical remote sensing data from turbid lakes. *The Science of the Total Environment*, pages 47–58, 2001. (Cité page 21.)
- T. Kutser, D.C. Pierson, K.Y. Kallio, A. Reinart, et S. Sobek. Mapping lake cdom by satellite remote sensing. *Remote Sensing of Environment*, 94 : 535–540, 2005. (Cité pages 1, 17 et 26.)
- C. Le, Y. Li, Y. Zha, D. Sun, C. Huang, et H. Lu. A four-band semi-analytical model for estimating chlorophyll a in highly turbid lakes : The case of taihu lake, china. *Remote Sensing of Environment*, 113 :1175–1182, 2009. (Cité pages 1 et 26.)

- M. Lee. *Benthic mapping of coastal waters using data fusion of hyperspectral imagery and airborne laser bathymetry*. PhD thesis, University of Florida, 2003. (Cité page 12.)
- Z.P. Lee. *Visible-infrared remote-sensing model and applications for ocean waters*. PhD thesis, University of South Florida, 1994. (Cité pages xv, 17, 19, 23, 47, 48, 49, 53, 54, 61, 62, 75, 76, 77, 81 et 82.)
- Z.P. Lee. Applying narrowband remote-sensing reflectance models to wideband data. *Applied Optics*, 48(17) :3177–3183, 2009. (Cité pages 8, 10, 17, 18, 19, 20 et 21.)
- Z.P. Lee et K.L. Carder. Effect of spectral band numbers on the retrieval of water column and bottom properties from ocean color data. *Applied Optics*, 41(12) :2191–2201, 2002. (Cité pages 8, 13, 17 et 19.)
- Z.P. Lee et K.L. Carder. Absorption spectrum of phytoplankton pigments derived from hyperspectral remote-sensing reflectance. *Remote Sensing of Environment*, 89 :361–368, 2003. (Cité page 18.)
- Z.P. Lee, K.L. Carder, et R.A. Arnone. Deriving inherent optical properties from water color : a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, 41(27) :5755–5772, 2002. (Cité pages 8 et 10.)
- Z.P. Lee, K.L. Carder, et K. Du. Effects of molecular and particle scatterings on the model parameter for remote-sensing reflectance. *Applied Optics*, 43(25) :4957–4964, 2004. (Cité pages 9 et 10.)
- Z.P. Lee, K.L. Carder, S.K. Hawes, R.G. Steward, T.G. Peacock, et C.O. Davis. Model for the interpretation of hyperspectral remote-sensing reflectance. *Applied Optics*, 33(24) :5721–5732, 1994. (Cité page 8.)
- Z.P. Lee, K.L. Carder, C.D. Mobley, R.G. Steward, et J.S. Patch. Hyperspectral remote sensing for shallow waters. ii. deriving bottom depths and water properties by optimization. *Applied Optics*, 38 :3831–3843, 1999. (Cité pages 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 27, 28, 30, 32 et 35.)
- Z.P. Lee, K.L. Carder, C.D. Mobley, R.G. Steward, et J.S. Patch. Hyperspectral remote sensing for shallow waters. i. a semi-analytical model. *Applied Optics*, 37(27) :6329–6338, 1998a. (Cité pages 8, 9, 12, 14, 16, 17, 18, 19, 20, 21 et 26.)
- Z.P. Lee, M.R. Zhang, K.L. Carder, et L.O. Hall. A neural network approach to deriving optical properties and depths of shallow waters. *Ocean Optics*, XIV, 1998b. (Cité pages 19 et 26.)
- K. Levenberg. A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 2 :164–168, 1944. (Cité page 28.)
- E.M. Louchard, R.P. Reid, F.C. Stephens, C.O. Davis, R.A. Leathers, et T.V. Downes. Optical remote sensing of benthic habitats and bathymetry in coastal environments at lee stocking island, bahamas : A comparative

- spectral classification approach. *Limnology and Oceanography*, 48 :511–521, 2003. (Cité pages 1 et 26.)
- D.R. Lyzenga. Passive remote sensing techniques for mapping water depth and bottom features. *Applied Optics*, 17(3) :379–383, 1978. (Cité page 5.)
- D. Manolakis. Hyperspectral signal models and implications to material detection algorithms. Dans *Proc. of the IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, volume 3, pages 117–120, 2004. (Cité page 70.)
- D. Manolakis et G. Shaw. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine*, pages 29–43, 2002. (Cité pages 69, 70 et 72.)
- S. Maritorena, A. Morel, et B. Gentili. Diffuse reflectance of oceanic shallow waters - influence of water depth and bottom albedo. *Limnology and Oceanography*, 39 :1689–1703, 1994. (Cité pages 12, 13, 22, 32 et 63.)
- D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11 :431–441, 1963. (Cité page 28.)
- M.W. Matthews, S. Bernard, et K. Winter. Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in zeekoevlei, a small hypertrophic lake, using MERIS. *Remote sensing of environment*, 114 :2070–2087, 2010. (Cité pages 10, 15, 16 et 26.)
- D. McKee, A. Cunningham, et S. Craig. Estimation of absorption and backscattering coefficients from in situ radiometric measurements : theory and validation in case II waters. *Applied Optics*, 42 :2804–2810, 2003. (Cité pages 9, 10 et 14.)
- A. Minghelli-Roman, L. Mathieu-Blanc, L. Loubersac, et F. Cauneau. Bathymetric estimation using meris images in coastal sea waters. *IEEE Geoscience and Remote Sensing Letters*, 4 :274–277, 2007. (Cité pages 12 et 25.)
- C.D. Mobley. *Light and water : radiative transfer in natural waters*. Academic Press, San Diego, 1994. (Cité page 10.)
- C.D. Mobley, B. Gentili, H.R. Gordon, Z. Jin, G.W. Kattawar, A. Morel, P. Reinersman, K. Stamnes, et R.H. Stavn. Comparison of numerical models for computing underwater light fields. *Applied Optics*, 32 :7484–7504, 1993. (Cité page 10.)
- C.D. Mobley, L.K. Sundman, C.O. Davis, J.H. Bowles, T.V. Downes, R.A. Leathers, M.J. Montes, W.P. Bisseth, D.D. Kohler, R.P. Reid, E.M. Louchard, et A. Gleason. Interpretation of hyperspectral remote-sensing imagery by spectrum matching and look-up tables. *Applied Optics*, 44 : 3576–3592, 2005. (Cité pages 1, 25, 26 et 29.)
- A. Morel. Optical properties of pure water and pure sea water. *Optical Aspects of Oceanography*, pages 1–24, 1974. (Cité pages 19, 20 et 21.)

- A. Morel et B. Gentili. Diffuse reflectance of oceanic waters. ii. bidirectional aspects. *Applied Optics*, 32 :6864–6879, 1993. (Cité page 9.)
- A. Morel, B. Gentili, H. Claustre, M. Babin, A. Bricaud, J. Ras, et F. Tièche. Optical properties of the "clearest" natural waters. *Limnology and oceanography*, 52 :217–229, 2007. (Cité pages 9, 14 et 19.)
- A. Morel et L. Prieur. Analysis of variations in ocean color. *Limnology and oceanography*, 22 :709–722, 1977. (Cité pages 9 et 13.)
- J.E. O'Reilly, S. Maritorena, B.G. Mitchell, D.A. Siegel, K.L. Carder, S.A. Garver, M. Kahru, et C. McClain. Ocean color chlorophyll algorithms for seaWiFS. *Journal of Geophysical Research*, 103 :24937–24953, 1998. (Cité page 26.)
- C. Petus, G. Chust, F. Gohin, D. Doxaran, J.M. Froidefond, et Y. Sagarmina. Estimating turbidity and total suspended matter in the adour river plume (south bay of biscay) using MODIS 250-m imagery. *Continental Shelf Research*, 30 :379–392, 2010. (Cité page 26.)
- W.D. Philpot. Radiative transfer in stratified waters : a single-scattering approximation for irradiance. *Applied Optics*, 26 :4123–4132, 1987. (Cité pages 12 et 13.)
- W.D. Philpot. Bathymetric mapping with passive multispectral imagery. *Applied Optics*, 28 :1569–1578, 1989. (Cité pages 12, 13 et 25.)
- D.C. Pierson et Strömbeck. Estimation of radiance reflectance and the concentrations of optically active substances in lake malaren, sweden, based on direct and inverse solutions of a simple model. *The Science of the Total Environment*, 268 :171–188, 2001. (Cité pages 17 et 26.)
- R.M. Pope et E.S. Fry. Absorption spectrum (380-700nm) of pure water. ii. integrating cavity measurements. *Applied Optics*, 36 :8710–8723, 1997. (Cité pages 16 et 19.)
- L. Prieur et S. Sathyendranath. An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials. *Limnology and Oceanography*, 26 :671–689, 1981. (Cité page 17.)
- K. Randolph, J. Wilson, L. Tedesco, L Li, D. Lani Pascual, et E. Soyeux. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll a and phycocyanin. *Remote Sensing of Environment*, 112 :4009–4019, 2008. (Cité pages 9 et 26.)
- L.L. Richardson et E.F. LeDrew. *Remote Sensing of Aquatic Coastal Ecosystem Processes : Science and Management Applications*. Springer, Netherlands, 2006. (Cité page 19.)
- R. Richter. *Atmospheric / topographic correction for airborne imagery : ATCOR-4 User Guide*. DLR IB 565-02/11, Wessling, Germany, 2011. (Cité pages 5 et 7.)

- F.C. Robey, D.R. Fuhrmann, E.J. Kelly, et R. Nitzberg. A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1) :208–216, 1992. (Cité pages 70 et 71.)
- S. Sathyendranath et T. Platt. Analytic model of ocean color. *Applied Optics*, 36 :2620–2629, 1997. (Cité page 14.)
- S. Sathyendranath, L. Prieur, et A. Morel. A three-component model of ocean colour and its application to remote sensing of phytoplankton pigments in coastal water. *International Journal of Remote Sensing*, 10 : 1373–1394, 1989. (Cité pages 9, 18, 19 et 27.)
- J.N. Schwarz, P. Kowalczyk, P. Kaczmarek, G.F. Cota, B.G. Mitchell, M. Kahru, F.P. Chavez, A. Cunningham, D. McKee, P. Gege, M.K. Riken, D.A. Phinney, et R. Raine. Two models for absorption by coloured dissolved organic matter (CDOM). *Oceanologia*, 44 :209–241, 2002. (Cité page 18.)
- S.G.H. Simis, S.W.M. Peters, et H.J. Gons. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and Oceanography*, 50 :237–245, 2005. (Cité pages 9 et 26.)
- R.C. Smith et K.S. Baker. Optical properties of the clearest natural waters (200–800 nm). *Applied Optics*, 20 :177–184, 1981. (Cité pages 16 et 19.)
- P. Stoica et R. Moses. *Introduction to spectral analysis*. Prentice Hall, 1997. (Cité page 63.)
- N. Strömbeck et D.C. Pierson. The effects of variability in the inherent optical properties on estimations of chlorophyll a by remote sensing in Swedish freshwaters. *The science of the total environment*, 268 :123–137, 2001. (Cité pages 9 et 26.)
- J. Sullivan, M. Twardowski, J. Zaneveld, C. Moore, A. Barnard, P. Donaghay, et B. Rhoades. Hyperspectral temperature and salt dependencies of absorption by water and heavy water in the 450–750 nm spectral range. *Applied Optics*, 45 :5294–5309, 2006. (Cité page 16.)
- S. Thiemann et H. Kaufmann. Lake water quality monitoring using hyperspectral airborne data - a semiempirical multisensor and multitemporal approach for the mecklenburg lake district, germany. *Remote Sensing of Environment*, 81 :228–237, 2002. (Cité page 25.)
- E. Vahtmäe, T. Kutser, G. Martin, et J. Kotta. Feasibility of hyperspectral remote sensing for mapping benthic macroalgal cover in turbid coastal waters - a baltic sea case study. *Remote Sensing of Environment*, 101 : 342–351, 2008. (Cité pages 9, 12, 13, 16, 17, 18, 19 et 21.)
- H.J. Van Der Woerd et R. Pasterkamp. HYDROPT : A fast and flexible method to retrieve chlorophyll-a from multispectral satellite observations of optically complex coastal waters. *Remote Sensing of Environment*, 112 : 1795–1807, 2007. (Cité pages 10 et 26.)

- V. Volpe, S. Silvestri, et M. Marani. Remote sensing retrieval of suspended sediment concentration in shallow waters. *Remote Sensing of Environment*, 115 :44–54, 2011. (Cité pages 8, 13 et 20.)
- P. Xiu, Y. Liu, G. Li, Q. Xu, H. Zong, Z. Rong, X. Yin, et F. Chai. Deriving depths of deep chlorophyll maximum and water inherent optical properties : a regional model. *Continental Shelf Research*, 29(19) :2270–2279, 2009. (Cité pages 8, 13, 16, 17, 18, 19, 20 et 21.)
- P.V. Zimba et A. Gitelson. Remote estimation of chlorophyll concentration in hyper-eutrophic aquatic systems : Model tuning and accuracy optimization. *Aquaculture*, 256 :272–286, 2006. (Cité page 26.)

Titre Estimation et détection en imagerie hyperspectrale : application aux environnements côtiers

Résumé Cette thèse aborde des problématiques d'estimation et de détection supervisée en imagerie hyperspectrale, appliquées ici aux environnements côtiers.

Des modèles bathymétriques de réflectance sont utilisés afin de représenter l'influence de la colonne d'eau sur la lumière incidente. Différents paramètres sont dits optiquement actifs et agissent sur le spectre de réflectance (phytoplancton, matière organique dissoute colorée...).

Nous proposons d'adopter une nouvelle approche statistique pour estimer ces paramètres, traditionnellement retrouvés par inversion des modèles physiques. Différentes méthodes telles que l'estimation du maximum de vraisemblance et du maximum *a posteriori*, ainsi que le calcul des bornes de Cramér-Rao, sont implémentées avec succès sur les données synthétiques et réelles.

Par ailleurs, nous adaptons les filtres supervisés couramment utilisés au contexte de la détection de cibles immergées. Dans le cas où les paramètres caractéristiques de la colonne d'eau sont inconnus, nous développons un nouveau filtre issu du test du rapport de vraisemblance généralisé permettant la détection sans aucune connaissance *a priori* sur ces paramètres.

Mots-clés Image hyperspectrale, modèle bathymétrique, estimation du maximum de vraisemblance, estimation du maximum *a posteriori*, borne de Cramér-Rao, détection supervisée, test du rapport de vraisemblance.

Title Estimation and detection in hyperspectral imagery : application in coastal environments

Abstract This thesis deals with estimation and supervised detection issues in hyperspectral imagery, applied in coastal environments.

Bathymetric models of reflectance are used for modeling the water column influence on the incident light. Various parameters are optically active and are responsible for distorting the reflectance spectrum (phytoplankton, colored dissolved organic matter...).

We adopt a new statistical approach for estimating these parameters, which are usually retrieved by inverting physical models. Various methods such as maximum likelihood estimation, maximum *a posteriori* estimation, and Cramér-Rao bound calculation, are successfully implemented on simulated and real data.

Moreover, we adapt the frequently used supervised detectors to the underwater target detection context. If some parameters describing the water column influence are unknown, we propose a new filter, based on the generalized likelihood ratio test, and that enables the detection without any *a priori* knowledge on these parameters.

Keywords Hyperspectral image, bathymetric model, maximum likelihood estimation, maximum *a posteriori* estimation, Cramér-Rao bound, supervised detection, likelihood ratio test.