



Université
de Toulouse

THÈSE

**En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par :

Institut National des Sciences Appliquées de Toulouse (INSA Toulouse)

Discipline ou spécialité :

Automatique

Présentée et soutenue par :

Rémy Lopez

le : 15 juillet 2013

Titre :

Développement d'une nouvelle algorithmie de localisation adaptée à
l'ensemble des mobiles suivis par le système ARGOS

JURY

Roland Chapuis, Polytech Clermont-Ferrand, Rapporteur

Jean-Charles Noyer, Université du Littoral, Rapporteur

Claude Jauffret, Université du Sud Toulon-Var, Examineur

Jean-Pierre Malardé, Collecte Localisation Satellites, Examineur

Philippe Gaspar, Collecte Localisation Satellites, Directeur de thèse

Patrick Danès, LAAS-CNRS, Directeur de thèse

Invité : Michel Devy, LAAS-CNRS

Ecole doctorale :

Systèmes (EDSYS)

Unité de recherche :

Collecte Localisation Satellites / LAAS-CNRS

Directeur(s) de Thèse :

Philippe Gaspar (Collecte Localisation Satellites), Patrick Danès (LAAS-CNRS)

Rapporteurs :

Roland Chapuis (Polytech Clermont-Ferrand), Jean-Charles Noyer (Université du Littoral)

REMERCIEMENTS

Un *big up* à,

Jean-Pierre Malardé, Patrick Danès, Philippe Gaspar et François Royer
pour leurs conseils et leur soutien

Le LAAS-CNRS et son équipe Robotique Action et Perception,
pour leur accueil, le café chaud et Atlanta

Collecte Localisation Satellites et la division SSIS,
pour leur accueil, le café chaud et Kyoto

Les nombreux artisans de toutes les directions de CLS, de nos filiales et
bureaux à travers le monde,
qui ont contribué à ce projet

Daniel Costa, Christophe Guinet, Neil et Liz Baker, Lucy Hawkes, Moses
Selebatso, Norv Dallin, Kellie Pendoley, Simon Benhamou, Luca Centu-
rioni, David Douglas et beaucoup d'autres,
pour leurs données ARGOS-GPS et leurs commentaires

— Résumé —

Développement d'une nouvelle algorithmie de localisation adaptée à l'ensemble des mobiles suivis par le système ARGOS

Depuis 1978, le système ARGOS assure à l'échelle mondiale la collecte de données et la localisation de plateformes pour des applications liées au suivi d'animaux, à l'océanographie et à la sécurité maritime. La localisation exploite le décalage Doppler affectant la fréquence de porteuse des messages émis par les plateformes et réceptionnés par des satellites dédiés. Au cours des vingt dernières années, les puissances d'émission des plateformes se sont réduites pour des conditions d'utilisation toujours plus extrêmes, augmentant le nombre de localisations de moindre qualité. Paradoxalement, les utilisateurs ont cherché à identifier des comportements à des échelles de plus en plus petites. L'objectif de ce projet est de développer un algorithme de localisation plus performant dans le contexte actuel afin de remplacer le traitement temps réel historique basé sur un ajustement par moindres carrés. Un service hors ligne, permettant de déterminer des localisations encore plus précises, est proposé dans un second temps.

Le problème est reformulé comme l'estimation de l'état d'un système dynamique stochastique, tenant compte d'un ensemble de modèles de déplacement admissibles pour les plateformes. La détermination exacte de la loi a posteriori de l'état présente alors une complexité exponentiellement croissante avec le temps. Le filtre « Interacting Multiple Model » (IMM) est devenu l'outil standard pour approximer en temps réel la loi a posteriori avec un coût de calcul constant. Pour des applications hors ligne, de nombreuses solutions sous-optimales de lissage multi-modèle ont aussi été proposées. La première contribution méthodologique de ce travail présente l'extension du cadre initial de l'IMM à un ensemble de modèles hétérogènes, c.-à-d. dont les vecteurs d'état sont de tailles et de sémantiques différentes. En outre, nous proposons une nouvelle méthode pour le lissage multi-modèle qui offre une complexité réduite et de meilleures performances que les solutions existantes. L'algorithme de localisation ARGOS a été réécrit en y incorporant le filtre IMM en tant que traitement temps réel et le lisseur multi-modèle comme service hors ligne. Une étude, menée sur un panel de 200 plateformes munies d'un récepteur GPS utilisé comme vérité terrain, montre que ces stratégies améliorent significativement la précision de localisation quand peu de messages sont reçus. En outre, elles délivrent en moyenne 30% de localisations supplémentaires et permettent de caractériser systématiquement l'erreur de positionnement.

Mots clés : système ARGOS, localisation Doppler, filtrage multi-modèle, lissage multi-modèle, suivi de mobile.

— Abstract —

Improving ARGOS Doppler location using multiple-model filtering and smoothing

The ARGOS service was launched in 1978 to serve environmental applications including oceanography, wildlife tracking and maritime safety. The system allows for worldwide positioning and data collection of Platform Terminal Transmitters (PTTs). The positioning is achieved by exploiting the Doppler shift in the carrier frequency of the messages transmitted by the PTTs and recorded by dedicated satellite-borne receivers. Over the last twenty years, the transmission power has decreased significantly and platforms have been used in increasingly harsh environments. This led to deliver a greater number of low quality locations while users sought to identify finer platform behavior. This work first focuses on the implementation of a more efficient location processing to replace the historical real time processing relying on a Least Squares adjustment. Secondly, an offline service to infer locations with even higher accuracy is proposed.

The location problem is formulated as the estimation of the state vector of a dynamical system, accounting for a set of admissible movement models of the platform. The exact determination of the state posterior pdf displays a complexity growing exponentially with time. The Interacting Multiple Model (IMM) algorithm has become a standard online approach to derive an approximated solution with a constant computational complexity. For offline applications, many sub-optimal multiple model schemes have also been proposed. Our methodological contributions first focused on extending the framework of the IMM filter so as to handle a bank of models with state vectors of heterogeneous size and meaning. Second, we investigated a new sub-optimal solution for multiple-model smoothing which proves to be less computationally expensive and displays markedly better performance than equivalent algorithms. The ARGOS location processing was rewritten to include the IMM filter as real time processing and the IMM smoother as offline service. We eventually analyzed their performances using a large dataset obtained from over 200 mobiles carrying both an ARGOS transmitter and a GPS receiver used as ground truth. The results show that the new approaches significantly improve the positioning accuracy, especially when few messages are received. Moreover, the algorithms deliver 30% more positions and give a systematic estimation of the location error.

Keywords : ARGOS System, Doppler location, multiple-model filtering and smoothing, target tracking.

TABLE DES MATIÈRES

Liste des symboles	1
Avant-propos	3
1. ARGOS, système satellite de localisation et de collecte	3
2. Des conditions de localisation de plus en plus difficiles	6
3. Une nouvelle approche pour la localisation	7
1. Moindres carrés : l'algorithme historique de localisation ARGOS	9
1.1. Principes de la localisation ARGOS	10
1.1.1. Systèmes de coordonnées pour la localisation	10
1.1.2. Fonction d'observation Doppler	12
1.1.3. Solutions géométriques de l'équation d'observation	13
1.2. Fonctionnement de l'algorithme historique	17
1.2.1. Ajustement par moindres carrés	17
1.2.2. Performances de localisation a priori	19
1.2.3. Particularités de l'implémentation	22
1.2.4. Distribution des positions aux utilisateurs	24
1.3. Analyse de performance	26
1.3.1. Panel de test et traitement	26
1.3.2. Statistiques de l'erreur empirique de localisation	28
1.3.3. Exemple de trajectoire	30
1.3.4. Discussion	30
Conclusions	32
2. Outils fondamentaux de filtrage et de lissage pour la localisation	37
2.1. Modélisation à espace d'état	38
2.1.1. Formulation et résolution	38
2.1.2. Avantages méthodologiques pour ARGOS	39
2.1.3. Modèles de comportement pour la localisation	40
2.2. Filtrage multi-modèle	42
2.2.1. Formulation des systèmes à sauts markoviens	42
2.2.2. Résolution par filtrage exact	43
2.2.3. Panorama des méthodes de résolution approchées	45
2.2.4. Construction du filtre IMM	48
2.3. Lissage multi-modèle	51
2.3.1. Hypothèses de base et stratégies de lissage	52
2.3.2. Solution exacte par récursion arrière	52
2.3.3. Solution exacte par double filtrage	54
2.3.4. Panorama des méthodes de résolution sous-optimales	57

Conclusions	59
3. Nouveaux outils pour le filtrage et le lissage multi-modèles	61
3.1. Lissage IMM-RTS	62
3.1.1. Vue d'ensemble	62
3.1.2. Phase 1 : lissage sur l'état continu	63
3.1.3. Phase 2 : interaction des modes avec M^2 combinaisons	65
3.1.4. Phase 2 (bis) : interaction des modes avec M combinaisons	67
3.1.5. Algorithme complet	67
3.1.6. Extension au lissage à retard fixe	71
3.2. Filtrage multi-modèle hétérogène	72
3.2.1. Formulation d'un système à sauts markoviens hétérogène	72
3.2.2. Stratégie d'interaction par composante	73
3.2.3. Algorithme de filtrage complet	76
3.3. Lissage multi-modèle hétérogène	77
3.3.1. Lisseur IMMHO-RTS	77
3.3.2. Lisseur IMMHO à double filtre	84
3.4. Application sur des données simulées	84
3.4.1. Lissage sur intervalle fixe	84
3.4.2. Lissage à retard fixe	88
3.4.3. Filtrage et lissage hétérogènes	89
Conclusions	94
4. Mise en œuvre de nouveaux algorithmes de localisation pour ARGOS	95
4.1. Choix techniques	96
4.1.1. Algorithmes et ensemble de modèles	96
4.1.2. Particularités d'implémentation	97
4.1.3. Aspects logiciels	98
4.2. Service temps-réel avec un filtre multi-modèle	98
4.2.1. Un plus grand nombre de positions	98
4.2.2. De meilleures performances dans les cas difficiles	100
4.2.3. Caractérisation de l'erreur	101
4.3. Service à valeur ajoutée sur les données archives	103
4.3.1. Un effet régularisant sur l'erreur	103
4.3.2. Caractérisation de l'erreur	104
Conclusions	105
Synthèse et perspectives	119
A. Compléments sur l'analyse d'erreur a priori	123
A.1. Biais sur l'altitude	123
A.2. Biais sur la fréquence d'émission	125
A.3. Dérive sur la fréquence d'émission	125
A.4. Déplacement pendant le passage	126

B. Outils de post-traitement des trajectoires ARGOS	131
B.1. Filtrage par décimation	131
B.2. Lissage des trajectoires	132
B.3. Modélisation d'état	133
C. Filtres de Kalman étendu et unscented	135
C.1. Filtre étendu	135
C.2. Filtre Unscented	136
Table des figures	139
Liste des tableaux	143
Bibliographie	145

LISTE DES SYMBOLES

λ	longitude
ϕ	latitude
h	altitude
f_r	fréquence de réception
f_t	fréquence de transmission
x^T	transposé de x (vecteur ou matrice)
$\mathbb{I}_n$	matrice identité de taille n par n
$\ x\ $	$\triangleq \sqrt{x^T x}$, norme Euclidienne du vecteur x
$\ x\ _A^2$	$\triangleq x^T A x$, avec la matrice A définie positive
$tr(A)$	trace de la matrice A
$P(m)$	probabilité de l'évènement m
$p(x)$	densité de probabilité d'une variable continue x
$\mathcal{N}(\hat{x}, X)$	loi gaussienne de moyenne $\hat{x}$ et de covariance X
$\mathcal{N}(:, \hat{x}, X)$	densité de la loi gaussienne $\mathcal{N}(\hat{x}, X)$
$E[x]$	espérance de la variable aléatoire x
$cov(x) = E[(x - E[x])(x - E[x])^T]$	covariance de la variable aléatoire x
$\hat{x}$	estimée de x
$\arg \max_x f(x)$	argument x qui maximise $f(x)$
k	index temporel
$v_{k:k'}$	l'ensemble des valeurs de v de l'instant k à k'

AVANT-PROPOS

1. ARGOS, système satellite de localisation et de collecte

Depuis 1978, le système ARGOS assure la collecte de données et la localisation de *plateformes*¹ réparties à la surface du globe [Fancy et al., 1988, Harris et al., 1990, Seegar et al., 1996]. Les principaux domaines d'application concernent le suivi d'animaux, l'océanographie et la sécurité maritime [CLS, 2011]. ARGOS est le fruit d'une collaboration entre le CNES, la NASA et la NOAA².

Le système s'appuie sur une constellation de satellites qui réceptionnent les messages émis par un total de 20 000 plateformes. Les messages sont retransmis au sol via des stations de réception puis acheminés vers des centres de traitement. Ces derniers se chargent de distribuer aux utilisateurs du système la position des plateformes et le contenu des messages. Depuis 1986, ARGOS est exploité par l'entreprise Collecte Location Satellites (CLS), filiale du CNES et de l'Ifremer³.

Balises

Sous sa forme la plus simple, un terminal ARGOS est l'attelage d'un émetteur, d'un oscillateur et d'une source d'énergie. Il peut être incorporé dans une bouée, fixé sur un bateau ou le dos d'un animal (cf. Fig. 1). Si les émetteurs animaux pouvaient peser plusieurs centaines de grammes trente ans auparavant, ils atteignent maintenant moins de 10 g. Le terminal ARGOS transmet des messages qui, outre un identifiant unique, peuvent contenir des données de capteurs comme la température ambiante, le rythme cardiaque d'un oiseau ou le niveau d'un cours d'eau. Il existe également des balises hybrides ARGOS-GPS qui transmettent régulièrement leur position GPS. Les messages sont émis dans la bande de fréquence 401,650 MHz $\pm$ 30 kHz avec une longueur utile de 256 bits maximum sur le canal bas débit du système ou de 4096 bits maximum sur le canal haut débit. Le système offre également la possibilité, via une voie descendante, de transmettre des informations aux balises équipées de récepteurs.

Le comportement d'une balise est caractérisé par sa période de répétition, *i.e.* l'intervalle de temps entre deux messages, ainsi que sa puissance d'émission. Pour des balises à forte puissance de types bateaux ou bouées, la puissance d'émission est de l'ordre de 1 à 2 W pour une période de répétition de 60 s. Pour les balises animaux où la durée de vie de la batterie est une contrainte prépondérante, les puissances d'émission sont plutôt de l'ordre de 250 à 500 mW avec une période de répétition de 120 s. À titre informatif, le système a été spécifié à l'origine pour des puissances de 2 W.

1. appelées aussi « balises » ou « terminaux » ARGOS.

2. CNES : Centre National d'Études Spatiales ; NASA : l'Agence spatiale américaine ; NOAA : l'Agence américaine d'étude de l'atmosphère et de l'océan.

3. IFREMER : Institut Français de Recherche pour l'Exploitation de la mer.

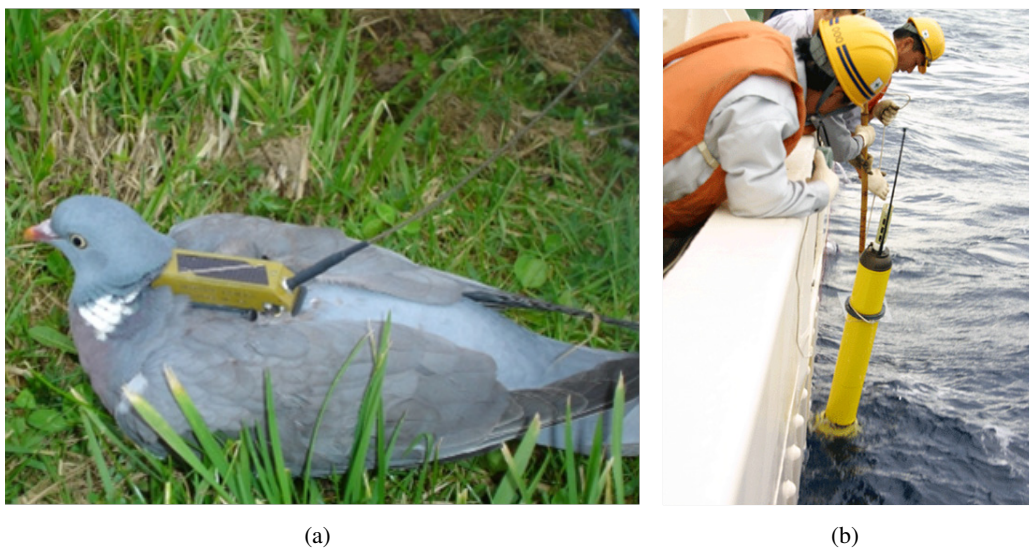


FIGURE 1.: Exemples de terminaux ARGOS. (a) Pigeon ramier équipé d'une balise de 17 grammes, et (b) flotteur sub-surface réalisant des profils de température et salinité des océans.

Satellites

La constellation compte actuellement sept satellites à vocation météorologique ou altimétrique (MA, MB, NK, NL, NN, NP et SR), lancés par la NOAA, EUMETSAT et l'ISRO⁴, où le récepteur ARGOS est embarqué comme instrument passager. Les satellites gravitent à 850 km d'altitude et se déplacent à une vitesse de 7 km/s sur un plan d'orbite polaire (inclinaison de 98 degrés par rapport au plan de l'équateur). Une révolution d'orbite dure approximativement 100 minutes. Les orbites sont héliosynchrones, c'est-à-dire qu'un satellite est en visibilité toujours à la même heure de la journée pour un point donné sur terre. La configuration de la constellation en avril 2013 est présentée dans la figure 2.

Les mesures de fréquences ne peuvent être réalisées que si la plateforme est dans le cercle de visibilité d'un satellite de la constellation (environ 2500 km de rayon). L'intervalle de temps durant lequel une plateforme se situe dans ce cercle est appelé *passage satellite*. Sa durée, qui s'étend au maximum sur 15 minutes, est d'autant plus réduite que la balise est éloignée de la trace au sol du satellite. Les plans d'orbite dérivant parfois de plusieurs degrés par mois, la forme de la constellation ainsi que les heures des passages évoluent au cours du temps.

Lorsque qu'un message est reçu par un satellite, la fréquence de sa porteuse est mesurée et enregistrée avec le contenu du message. Du fait du mouvement relatif entre la balise et les satellites de la constellation, on observe une différence, induite par l'effet Doppler-Fizeau, entre la fréquence d'émission de la balise et la fréquence de réception à bord du satellite. Le calcul de positions ARGOS tire parti de ce phénomène physique.

4. EUMETSAT : l'Organisation européenne pour l'exploitation des satellites météorologiques ; ISRO : l'Agence spatiale Indienne

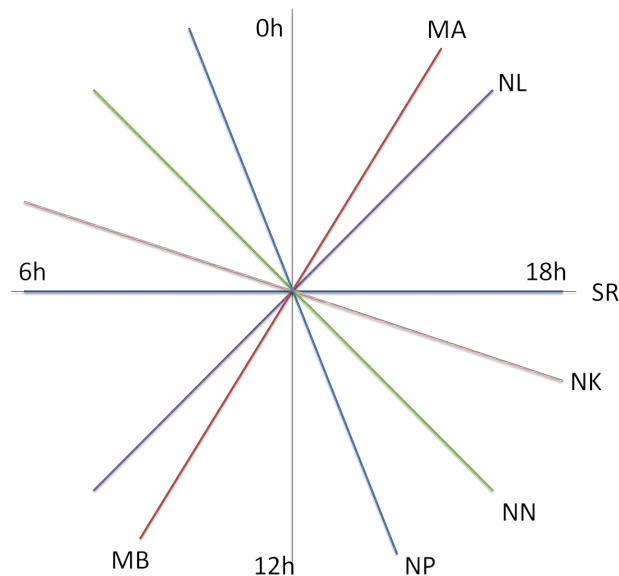


FIGURE 2.: Plans d'orbite de la constellation ARGOS en avril 2013 affichés selon les deux heures locales du nœud ascendant (points de l'orbite où le satellite traverse le plan de l'écliptique du sud vers le nord). Ce sont aussi les heures locales de passage des satellites au dessus d'un même point de la terre.

Segment sol

La télémesure du satellite, contenant les messages et les fréquences de réception, est déversée dans des stations de réception globales et locales. Lorsque le satellite est en visibilité d'une station locale, la télémesure est transmise au moment où les messages sont reçus et enregistrés à bord du satellite, permettant de délivrer plus rapidement les données au centre de calcul. En revanche, en visibilité d'une station globale, l'ensemble de la télémesure de la dernière révolution d'orbite est déchargé avant de libérer la mémoire embarquée à bord du satellite. Au sein des centres de calcul (Toulouse et Washington en redondance), la télémesure est traitée de manière à extraire le contenu des messages ainsi que les informations de datation et de fréquence reçues. Les données sont mises à disposition des clients par mail, CD-ROM, ainsi qu'au travers d'interfaces web ou machine-machine. Sur une seule journée de traitement, plusieurs centaines de milliers de localisations sont calculés et plusieurs millions de messages sont collectés.

Dans la chaîne de traitement ARGOS, on distingue plusieurs modules essentiels à la localisation : ORB, DAT et LOC. Le module d'orbitographie ORB extrapole les éphémérides satellites de la journée (vitesse et position) avec un pas de temps de 3 minutes. Par ailleurs, il recalcule par interpolation les éphémérides de la veille afin d'en avoir une estimation plus précise. Le module de datation DAT fait le lien entre le temps universel et le temps bord de chaque satellite pour établir la date de réception des messages. Enfin, le module LOC forme les jeux de données — constitués des mesures de fréquences datées et des éphémérides satellites — et applique le calcul de localisation.

Systèmes complémentaires et concurrents

Sur les activités de collecte où le temps de mise à disposition des données est un critère prépondérant, ARGOS est directement concurrencé par Iridium [Iridium Communications Inc., 2012], un système global de télécommunication qui exploite une constellation de 66 satellites. La durée de mise à disposition des données de collecte et de localisation ARGOS est de l'ordre de quelques heures alors qu'Iridium atteint quelques minutes. Le taille de la constellation, permettant la transmission en temps réel des données, impose néanmoins un coût d'utilisation plus élevé.

Pour les activités de localisation, ces questions de concurrence sont foncièrement moins tranchées. Les récepteurs GPS présentent une meilleure précision de positionnement qu'ARGOS (quelques dizaines de mètres) et, comme nous l'avons déjà évoqué, sont parfois utilisés comme capteur d'une balise. Alors que la localisation ARGOS est déterminée dans le centre de calcul, la position GPS l'est dans la plateforme embarquant le récepteur et doit être alors systématiquement collectée. L'utilisation d'un récepteur GPS dans une balise ARGOS occasionne par ailleurs un coût de fabrication plus élevé et une consommation d'énergie accrue du fait des calculs du récepteur et de la nécessité de construire les messages ARGOS contenant les positions.

Pour les balises où la puissance et le coût ne sont pas une limitation, à l'instar du suivi de bateaux, l'utilisation de la localisation Doppler permet de s'assurer de la véracité des positions GPS collectées par ARGOS. En effet, les positions GPS sont utilisées pour vérifier précisément que les navires ne franchissent pas des frontières maritimes mais il arrive qu'elles soient falsifiées préalablement à leur collecte. ARGOS est plus difficile à tromper car les calculs de position ne s'effectuent pas à bord des bateaux.

Outre le GPS, les réseaux GSM proposent un moyen de localiser des balises sans recourir à un système de collecte comme ARGOS mais sans une couverture mondiale.

2. Des conditions de localisation de plus en plus difficiles

L'algorithme de localisation ARGOS était basé depuis l'ouverture du système sur un ajustement par moindres carrés. Le calcul s'opérait initialement avec au moins quatre messages reçus pendant un passage satellite au dessus de la balise. Il s'accompagnait d'une caractérisation statistique de l'erreur de localisation. Dans ces conditions, l'incertitude de localisation s'échelonne entre plusieurs centaines de mètres et quelques kilomètres en moyenne. Fait notable, une particularité géométrique du problème de localisation implique l'existence de deux positions possibles sur terre qui satisfont l'ajustement moindres carrés. Avec une probabilité d'erreur non nulle, le traitement est en outre amené à choisir parmi ce couple la véritable position de la balise.

L'augmentation continue du nombre de balises émettant à faible puissance et dans des conditions environnementales difficiles ont poussé à compléter l'algorithme pour calculer des positions avec seulement deux ou trois messages mais sans estimation de l'erreur de localisation. Dans ces circonstances, la précision de localisation se situe entre plusieurs kilomètres et quelques dizaines de kilomètres en moyenne. La généralisation de balises à faible puissance, notamment pour le suivi d'animaux, s'est pourtant accompagnée d'un besoin accru de précision. En effet, ARGOS a d'abord servi à mettre en évidence les trajets migratoires ou encore des bassins de reproduction.

À l'heure actuelle, les utilisateurs cherchent à identifier des comportements plus fins et la communauté a déployé des trésors d'imagination pour raffiner les trajectoires ARGOS : décimation des positions aberrantes, lissage des trajectoires ou encore modélisation a priori du déplacement des animaux.

3. Une nouvelle approche pour la localisation

Dès 2008, CLS envisage une refonte profonde de son algorithme de localisation. Il est alors proposé de reformuler le problème dans le contexte des modèles d'état [Bar-Shalom and Li, 1993]. La localisation de la balise est vue comme l'état d'un système dynamique gouverné par un modèle d'évolution. L'état est relié aux mesures de fréquences acquises pendant les passages satellites par un modèle d'observation. De nombreuses méthodes permettent de fusionner de façon optimale une prédiction de l'état fournie par le modèle d'évolution avec les mesures de fréquences afin d'estimer la position.

En 2009, une maquette simplifiée du procédé, basée sur un filtre de Kalman "Unscented" [Julier and Uhlmann, 1997, Wan and Van Der Merwe, 2000], a été développée lors d'un stage de fin d'étude [Lopez, 2009] et a permis de valider l'approche sur quelques cas. Dans les mois suivants, un projet a été proposé sous la forme d'une thèse CIFRE afin de développer une version opérationnelle du code de calcul pour estimer en temps réel (*i.e.* au fil de l'arrivée des messages) les positions de l'ensemble des mobiles du système. Ce code se base sur la maquette de [Lopez, 2009]. Il s'agit aussi de prouver sur un plus large panel de balises que la méthode réduit significativement l'erreur lorsque la plateforme est suivie en conditions difficiles. Comme les mobiles du système sont de natures différentes et qu'ils peuvent changer de comportement au cours du temps, le cadre initial a été élargi aux techniques de filtrage *multi-modèle* [Li and Jilkov, 2005] qui offrent l'avantage de considérer plusieurs dynamiques de déplacement simultanément. Le second objectif est de proposer un service hors ligne aux utilisateurs qui maximise les performances de localisation avec des techniques de lissage multi-modèle. CLS stocke depuis 2008 les données de fréquences reçues, ce qui offre l'opportunité de recalculer les trajectoires en temps différé.

Le présent document relate d'abord le principe de localisation ARGOS et son application dans l'algorithme historique fondé sur un ajustement par moindres carrés (chapitre 1). C'est l'occasion de mettre en lumière les phénomènes qui modulent la précision de localisation. Celle-ci sera en outre évaluée sur un panel de 200 balises ARGOS-GPS qui inclut des animaux terrestres, des oiseaux, des animaux marins, des bouées et des bateaux avec un total de plus de 125 000 localisations. Dans un second temps (chapitre 2), on justifie l'emploi d'une modélisation d'état dans un contexte où plusieurs dynamiques de déplacement sont possibles. Du fait de la complexité exponentiellement croissante avec le temps du problème de suivi multi-modèle, on s'attache à inventorier les outils fondamentaux de résolution pour le filtrage et le lissage. Nous proposons dans un troisième temps de nouvelles stratégies, l'une pour le filtrage et l'autre pour le lissage, qui permettent de pallier les limitations des méthodes actuelles (chapitre 3). Enfin, nous analysons les performances des traitements temps réel et temps différé développés à partir de ces outils sur le panel de balises comparativement à l'ajustement par moindres carrés (chapitre 4).

MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS



Lion de mer et éléphant de mer.

Sommaire

1.1. Principes de la localisation ARGOS	10
1.1.1. Systèmes de coordonnées pour la localisation	10
1.1.2. Fonction d'observation Doppler	12
1.1.3. Solutions géométriques de l'équation d'observation	13
1.2. Fonctionnement de l'algorithme historique	17
1.2.1. Ajustement par moindres carrés	17
1.2.2. Performances de localisation a priori	19
1.2.3. Particularités de l'implémentation	22
1.2.4. Distribution des positions aux utilisateurs	24
1.3. Analyse de performance	26
1.3.1. Panel de test et traitement	26
1.3.2. Statistiques de l'erreur empirique de localisation	28
1.3.3. Exemple de trajectoire	30
1.3.4. Discussion	30
Conclusions	32

CE premier chapitre présente le calcul de position ARGOS actif depuis la mise en service du système et basé sur un ajustement par moindres carrés. Notre contribution tient ici à l'analyse des performances du traitement de localisation historique sur un jeu de données exhaustif incluant un large choix d'applications et de modes de fonctionnement.

1.1. Principes de la localisation ARGOS

1.1.1. Systèmes de coordonnées pour la localisation

ARGOS utilise le système géodésique mondial WGS84 (World Geodetic System 1984) qui définit un référentiel d'origine O , situé au centre de masse de la Terre, avec les axes $\vec{X}$, $\vec{Y}$ et $\vec{Z}$ pour se repérer dans l'espace à l'aide des coordonnées cartésiennes. L'axe $\vec{X}$ est dirigé depuis O vers l'intersection de l'équateur avec le méridien de Greenwich. L'axe $\vec{Y}$ est obtenu par une rotation de $\vec{X}$ de 90° le long de l'équateur vers l'Est de sorte que le plan de l'équateur soit confondu avec $\vec{XOY}$. Enfin, $\vec{Z}$ définit l'axe de rotation de la Terre et se déduit du produit vectoriel $\vec{X} \wedge \vec{Y}$. Du fait de sa rotation propre et de l'attraction subie par la lune et le soleil, la Terre est bien représentée par une sphère légèrement aplatie au niveau des pôles. Elle forme ainsi un ellipsoïde de révolution de centre O dont les longueurs des axes aux pôles et à l'équateur sont notées respectivement RP et RE . Le WGS84 fixe aussi la taille du demi-grand axe $RE = 6378,137$ km et l'aplatissement $f = 1/298,257223563$ de l'ellipsoïde [National Geospatial-Intelligence Agency, 1997]. La valeur du demi-petit axe se déduit de la relation $RP = RE(1 - f)$.

Bien qu'il soit possible d'utiliser les coordonnées cartésiennes (x, y, z) pour situer un point M sur terre, on utilise plus communément les coordonnées géographiques en longitude λ , latitude ϕ et altitude h par rapport à la surface de l'ellipsoïde. La Fig. 1.1 illustre les liens entre ces deux systèmes de coordonnées. Les coordonnées cartésiennes se déduisent des coordonnées géographiques au moyen des équations paramétriques d'un ellipsoïde :

$$\begin{aligned} x &= GE \cos \Theta \cos \lambda \\ y &= GE \cos \Theta \sin \lambda \\ z &= GP \sin \Theta \end{aligned} \tag{1.1}$$

où $GE = RE + h$, $GP = RP + h$ et Θ , la latitude paramétrique (ou réduite), est reliée à la latitude géographique ϕ par $\tan \Theta = \frac{GP}{GE} \tan \phi$. L'opération inverse, c'est-à-dire le calcul des coordonnées géographiques à partir du triplet (x, y, z) , n'est pas explicite :

$$\begin{aligned} \tan \lambda &= \frac{y}{x} \\ \tan \Theta &= \frac{z + RE f \sin \Theta}{\sqrt{x^2 + y^2}} \\ h &= \frac{z}{\sin \Theta} - RP. \end{aligned} \tag{1.2}$$

Si la longitude et l'altitude s'obtiennent simplement, l'évaluation de la latitude réduite impose l'utilisation de méthodes itératives ou directes [Gerdan and Deakin, 1999] plus complexes que le système d'équations (1.1).

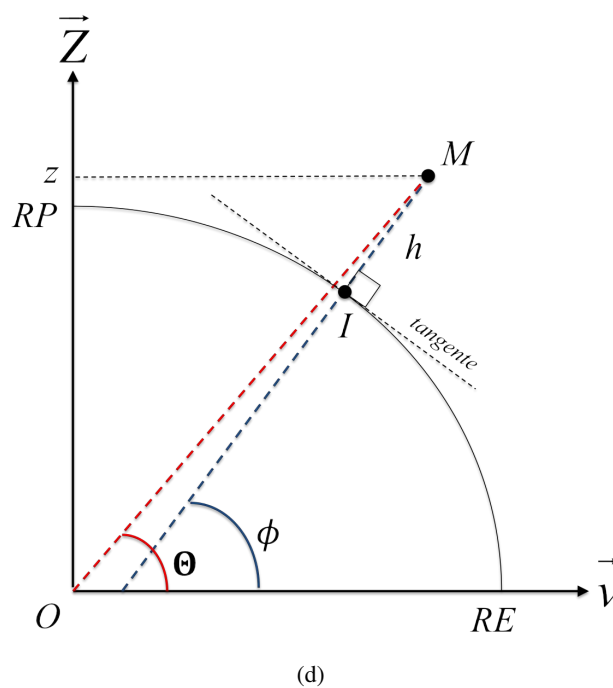
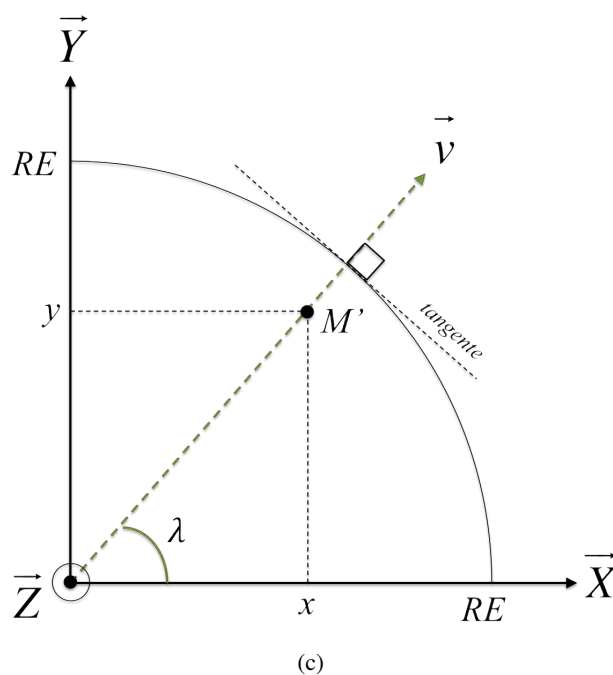


FIGURE 1.1.: Pour un point M de l'espace, on affiche sa projection M' sur le plan équatorial $\vec{X}\vec{O}\vec{Y}$ (c) et sur le plan $\vec{v}\vec{O}\vec{Z}$ (d). Les coordonnées cartésiennes sont données par le triplet (x, y, z) et les coordonnées géographiques par (λ, ϕ, h) . L'altitude h est la longueur du segment MI . La latitude ϕ est l'angle entre la normale à l'ellipsoïde passant par M et le plan équatorial. La latitude paramétrique Θ est l'angle entre ce même plan et le segment OM .

1.1.2. Fonction d'observation Doppler

Le décalage en fréquence qui s'applique à la porteuse des messages reçus est exploité pour déterminer la position de la plateforme émettrice. Soient $\vec{B} = (x, y, z)^T$ le vecteur des coordonnées cartésiennes de la plateforme et $\vec{S} = (x_s, y_s, z_s)^T$ le vecteur des coordonnées cartésiennes du satellite. On définit $\vec{V}_B$ et $\vec{V}_S$ les vecteurs vitesses de la balise et du satellite respectivement. La relation entre la fréquence de transmission f_t de la plateforme et la fréquence de réception f_r à bord du satellite s'écrit

$$\begin{aligned} f_r &= \mathcal{H}(f_t, \vec{B}, \vec{S}, \vec{V}_B, \vec{V}_S) \\ &= f_t \left(1 - \frac{(\vec{V}_S - \vec{V}_B) \cdot \vec{u}}{c} \right) \end{aligned} \quad (1.3)$$

avec c la vitesse de la lumière, $\vec{u} = \frac{\vec{S} - \vec{B}}{\|\vec{S} - \vec{B}\|} = \frac{\vec{r}}{\|\vec{r}\|}$ le vecteur unité orienté de la plateforme vers le satellite (cf. Fig. 1.2). Le produit scalaire $(\vec{V}_S - \vec{V}_B) \cdot \vec{u}$ représente la vitesse radiale relative v_r entre la plateforme et le satellite. Étant donné que $\|\vec{V}_B\| \ll \|\vec{V}_S\| \approx 7 \text{ km.s}^{-1}$, la vitesse de la plateforme est négligée dans (1.3).

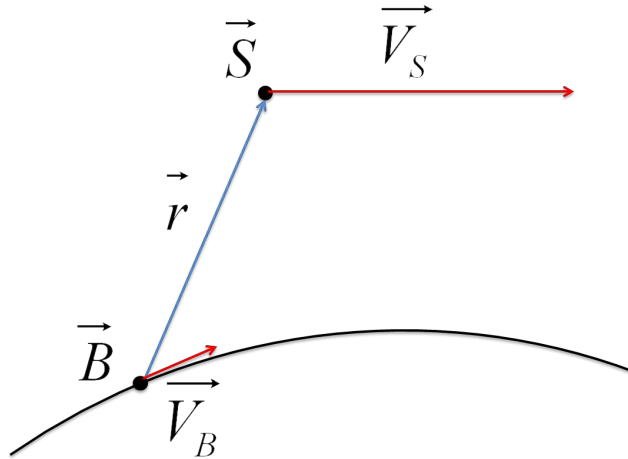


FIGURE 1.2.: Le balise $\vec{B}$ avec la vitesse $\vec{V}_B$ est observée par le satellite $\vec{S}$ avec la vitesse $\vec{V}_S$. Le vecteur radial balise-satellite est $\vec{r}$.

Dans la fonction d'observation Doppler $\mathcal{H}$, les éphémérides satellites $\vec{S}$ et $\vec{V}_S$ sont données. De plus, la plateforme est supposée se déplacer sur l'ellipsoïde de référence du WGS84, de sorte que $\vec{B}$ peut être aussi exprimé en fonction de la longitude λ , la latitude ϕ et l'altitude h . Restent donc comme inconnues la fréquence de transmission f_t ainsi que la longitude λ , la latitude ϕ et l'altitude h de la plateforme. Dans la suite, $\mathcal{H}(f_t, \vec{B}, \vec{S}, \vec{V}_B, \vec{V}_S)$ sera donc remplacé par $\mathcal{H}(\lambda, \phi, h, f_t)$.

Prise en compte de la durée de comptage τ

La mesure de la fréquence n'est pas instantanée et s'étend sur un intervalle de $\tau = 120 \text{ ms}$ appelé *durée d'intégration de la phase* ou *durée de comptage*. Pendant ce laps de temps, le satellite s'est

déplacé de plusieurs centaines de mètres. La position et la vitesse considérées dans la fonction d'observation Doppler sont la moyenne des valeurs en début et fin du comptage.

Une écriture alternative de la fonction d'observation Doppler, utilisée pour la localisation ARGOS, est obtenue en remplaçant la vitesse radiale relative v_r de (1.3) par une approximation par différences finies $\frac{R_f - R_d}{\tau}$ où $R_d = \|\vec{S}_d - \vec{B}\|$ et $R_f = \|\vec{S}_f - \vec{B}\|$ désignent respectivement la distance satellite-balise en début et en fin de comptage, avec $\vec{S}_d = (x_{s,d}, y_{s,d}, z_{s,d})^T$ et $\vec{S}_f = (x_{s,f}, y_{s,f}, z_{s,f})^T$ les positions satellites correspondantes. Quant aux éphémérides satellites à l'instant de la mesure, elles sont calculées avec une interpolation d'Everett [Davis, 1975]. La date de mesure se situe en milieu de comptage.

Forme de la courbe Doppler

L'évolution de la fréquence de réception au cours d'un passage satellite prend la forme d'une courbe strictement décroissante avec un point d'inflexion en son milieu. En ce point où le satellite et la balise sont les plus proches et où la variation de la fréquence reçue est la plus forte, la vitesse radiale relative v_r s'annule et la fréquence mesurée à bord du satellite est égale à la fréquence d'émission. Seuls quelques points de cette courbe, correspondant aux messages reçus pendant le passage, sont en réalité disponibles pour le calcul de position.

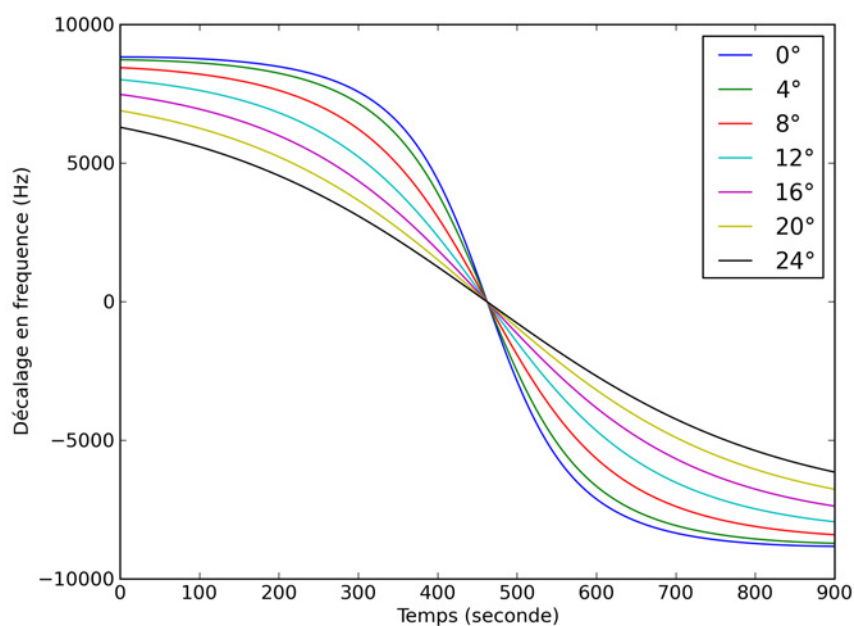
La Fig. 1.3 montre pendant la durée d'un passage la différence $f_r - f_i$ induite par l'effet Doppler, ainsi que la dérivée par rapport au temps de la fréquence de réception. On simule pour cela une balise de fréquence constante et un satellite avec une orbite parfaitement polaire. La trace au sol du satellite (ou trace sub-satellite) coupe alors perpendiculairement l'équateur. On positionne initialement la balise en ce point d'intersection et on la déplace le long de l'équateur. Les courbes sont affichées pour une distance à la position initiale de plus en plus grande.

La distance à la trace sub-satellite a un impact direct sur la forme de la courbe Doppler et la dérivée en fréquence. On distingue deux cas extrêmes : lorsque la plateforme se situe sur la trace au sol du satellite (passage sous trace, durée la plus longue) et lorsque la plateforme se situe en bordure du cercle de visibilité (passage à l'horizon, durée la plus courte). Sous trace, la dérivée en fréquence de la courbe Doppler atteint 80 Hz/s au point d'inflexion alors que pour les passages à l'horizon, elle se situe autour de 20 Hz/s.

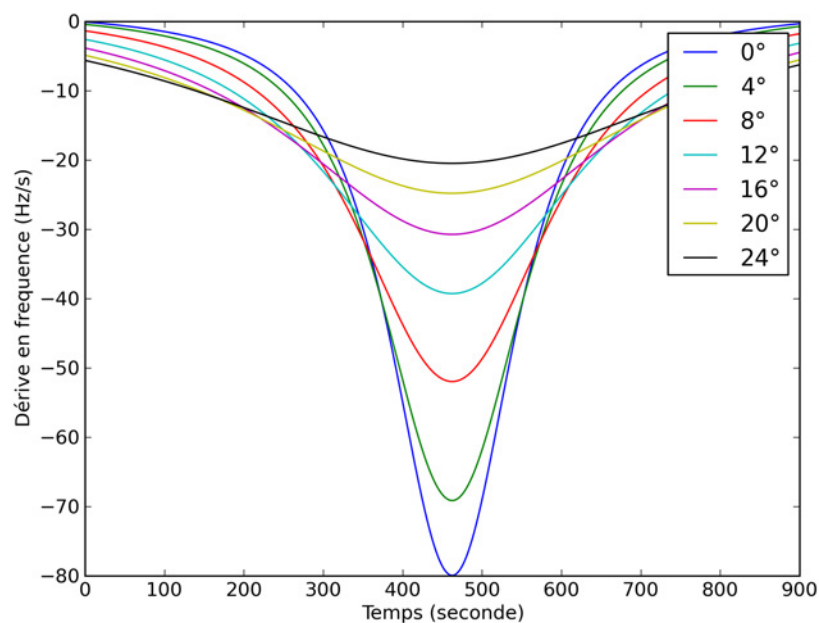
1.1.3. Solutions géométriques de l'équation d'observation

Pour une fréquence d'émission et une fréquence de réception données, la position de la plateforme sur terre peut être déterminée de manière géométrique. Cette approche permet en outre de représenter visuellement les solutions de l'équation (1.3).

Proposition 1 Soient f_i la fréquence d'émission de la balise et f_r la fréquence reçue au niveau du satellite. Les triplets (x, y, z) solutions de l'équation (1.3) se situent sur un cône de sommet $\vec{S}$ et dont l'axe de révolution coïncide avec la direction du vecteur vitesse du satellite. Si le satellite est en approche de la balise, c'est la partie avant du cône qui est solution. Si le satellite s'éloigne de la balise, c'est la partie arrière qui est solution.



(a)



(b)

FIGURE 1.3.: (a) Différence $f_r - f_t$ et (b) dérivée en fréquence en fonction du temps pour différentes distances de la balise à la trace sub-satellite (en degrés le long de l'équateur). L'origine du temps est prise en début de passage et on ne tient pas compte du cercle de visibilité.

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

Preuve Soit le repère cartésien d'origine $\vec{S}$ et dont l'axe $\vec{z}$ coïncide avec la direction et le sens de $\vec{V}_S$. Le choix des axes $\vec{x}$ et $\vec{y}$ est libre. Dans ce repère, $\vec{S} = \vec{0}$, $\vec{V}_S = (0, 0, V_{S,z})^T$, $\|\vec{V}_S\| = V_{S,z}$ et $\vec{B} = (x, y, z)^T$. L'équation (1.3) s'écrit alors

$$f_r = f_t \left(1 + \frac{1}{c} \frac{(0, 0, \|\vec{V}_S\|)^T \cdot (x, y, z)^T}{\sqrt{x^2 + y^2 + z^2}} \right) \quad (1.4)$$

qui, après ré-arrangement des membres et élévation au carré, implique

$$x^2 + y^2 = z^2 \left(\frac{1 - a^2}{a^2} \right) \text{ avec } a = \frac{c}{\|\vec{V}_S\|} \left(\frac{f_r}{f_t} - 1 \right). \quad (1.5)$$

En posant $\tan^2 \alpha = \left(\frac{1 - a^2}{a^2} \right)$, l'équation précédente définit un cône de révolution d'axe $\vec{V}_S$ dont l'angle avec la génératrice est α .

Réciproquement, l'équation (1.5) est équivalente à

$$\sqrt{x^2 + y^2} = z \frac{\sqrt{1 - a^2}}{a} \text{ ou } \sqrt{x^2 + y^2} = -z \frac{\sqrt{1 - a^2}}{a}. \quad (1.6)$$

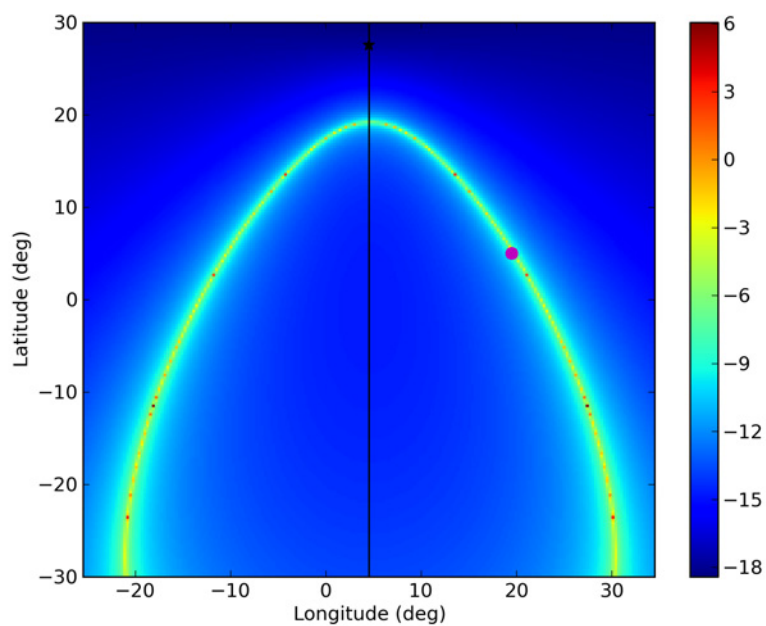
La première équation permet bien de remonter à (1.4). Quand le satellite est en approche de la plateforme, $\frac{f_r}{f_t} \geq 1$ donc $a \geq 0$ et $\sqrt{x^2 + y^2} = z \frac{\sqrt{1 - a^2}}{a}$ n'admet de solutions physiques que pour $z \geq 0$ (partie avant du cône). Pendant que le satellite s'éloigne, $a \leq 0$ et seuls les $z \leq 0$ sont possibles (partie arrière). La seconde équation aboutit par contre à

$$f_r = f_t \left(1 + \frac{1}{c} \frac{(0, 0, -\|\vec{V}_S\|)^T \cdot (x, y, z)^T}{\sqrt{x^2 + y^2 + z^2}} \right) \quad (1.7)$$

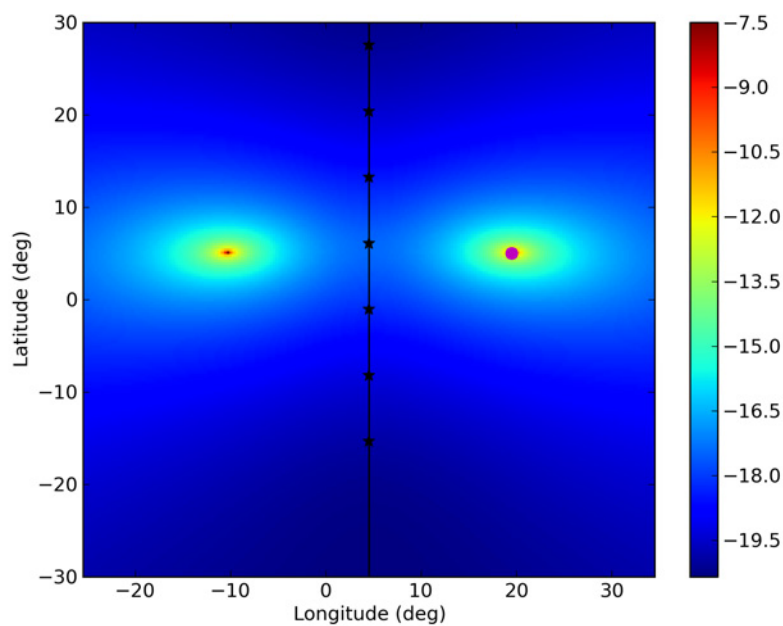
qui correspond à un satellite se déplaçant dans le sens inverse, contredisant l'hypothèse de départ selon laquelle $\vec{z}$ coïncide avec le sens de $\vec{V}_S$. $\square$

Lorsqu'un unique message est acquis pendant un passage satellite, l'ensemble des positions possibles de la balise sur terre forme une courbe symétrique par rapport à la trace au sol du satellite. Cet ensemble est obtenu par l'intersection du cône solution avec la surface de l'ellipsoïde terrestre. La Fig. 1.4 montre cet ensemble solution en traçant une carte de résidus. En revanche, quand on dénombre deux messages ou plus, il ne subsiste que deux positions sur terre, symétriques aussi par rapport à la trace au sol et situées aux intersections de tous les cônes solutions avec la terre (cf. Fig. 1.5). Dans ce dernier cas, il n'est alors pas possible de distinguer la véritable position (dite position *nominale*) de la mauvaise (dite position *image*) sans une connaissance a priori sur la trajectoire de la balise.

Bien que proposant une solution analytique, l'approche géométrique n'est cependant pas directement utilisable pour calculer une localisation car elle nécessite de fixer la fréquence d'émission de la balise. Cette dernière n'est ni connue, ni complètement stable au cours de la vie de la plateforme. En effet, son choix est laissé à la discrétion du constructeur et les variations de température influencent largement la fréquence de l'oscillateur embarqué dans la balise. Le second inconvénient est que la localisation géométrique n'est possible qu'avec des mesures non bruitées. Dans le cas contraire, l'intersection des cônes solutions avec la terre n'existe généralement pas (à moins d'utiliser seulement deux mesures du passage).



(a)



(b)

FIGURE 1.4.: Cartographie des résidus (en log) pour un passage satellite d'un seul message (a) et de plus de deux messages (b) sans prise en compte du cercle de visibilité du satellite. Le résidu est la somme des écarts entre les fréquences réellement reçues et leurs valeurs théoriques. La position réelle de la plateforme est le point violet. La courbe noire représente la trace au sol du satellite (déplacement nord-sud). Les étoiles marquent les positions au sol du satellite au moment de la réception des messages.

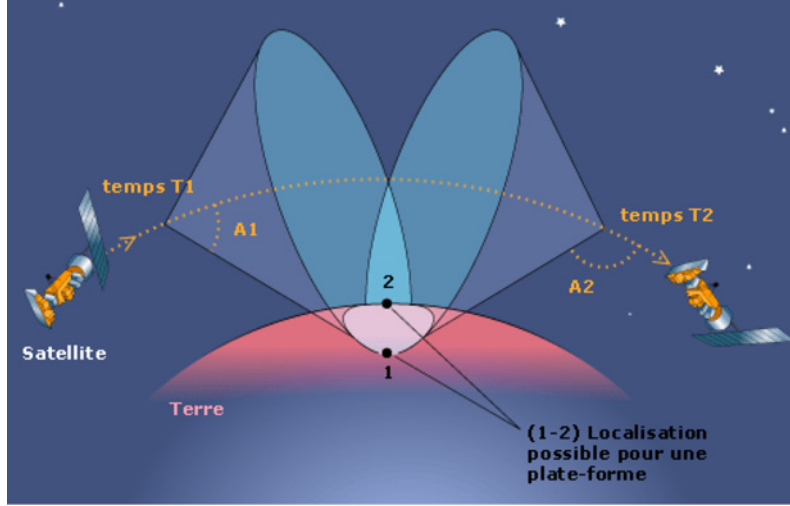


FIGURE 1.5.: Les positions nominale et image se situent aux intersections des cônes solutions avec l'ellipsoïde terrestre.

1.2. Fonctionnement de l'algorithme historique

L'estimation de la position et de la fréquence d'émission est réalisée séquentiellement au fil de l'arrivée des mesures de fréquence préalablement regroupées en passages satellites. Pour le calcul de localisation, on considère en première approximation que la fréquence d'émission de la balise et sa position sont constantes sur la durée d'un passage. L'indice temporel $k = 0, 1, \dots$ désigne le numéro du passage satellite dont la localisation correspondante est donnée à la date moyenne des mesures du passage. La position et la vitesse du satellite étant particulières à chaque mesure de chaque passage, la fonction d'observation Doppler appliquée à la $i^{\text{ème}}$ mesure du $k^{\text{ème}}$ passage sera dorénavant notée $\mathcal{H}_k^i$.

1.2.1. Ajustement par moindres carrés

De manière générale, un ajustement par moindres carrés permet d'estimer un ensemble de paramètres inconnus et constants à partir d'un ensemble de mesures bruitées. Soient $x_k \in \mathbb{R}^n$ le vecteur de paramètres à déterminer et $z_k \in \mathbb{R}^{m_k}$ le vecteur des observations à l'instant k ($m_k \geq n$). Les mesures sont par hypothèse reliées aux paramètres par le système d'équations suivant :

$$z_k = g_k(x_k) + v_k \quad (1.8)$$

où g_k est la fonction d'observation, $v_k \in \mathbb{R}^{m_k}$ représente un bruit de mesure supposé ici gaussien de moyenne $\bar{v}_k = 0$ et de matrice de covariance R_k . Si z_k désigne une réalisation du vecteur d'observations, l'estimateur des moindres carrés fournit une estimée $\hat{x}_k$ qui minimise par rapport à x_k la somme pondérée des écarts quadratiques [Zarrouati, 1987]

$$S_k(x_k) = \|z_k - g_k(x_k)\|_{R_k^{-1}}^2. \quad (1.9)$$

Application à la localisation ARGOS

Concernant le calcul de localisation, $x_k = (\lambda_k, \phi_k, f_{t,k})^T$ est le vecteur des $n = 3$ inconnues, z_k est un vecteur de m_k mesures de fréquences. L'altitude h n'est pas considérée comme une inconnue. Elle est fixée à zéro pour les plateformes évoluant en mer. Pour les balises terrestres et les oiseaux, l'altitude est approximée par le modèle numérique de terrain GTOPO30 [USGS, 2010].

Le problème consistant à déterminer la position et la fréquence de transmission s'écrit comme un système d'équations non linéaires où $g_k(x_k) = [\mathcal{H}_k^1(x_k) \dots \mathcal{H}_k^i(x_k) \dots \mathcal{H}_k^{m_k}(x_k)]^T$ et $R_k = \sigma_k^2 \mathbb{I}_{m_k}$ (avec σ_k l'écart-type du bruit de mesure). Comme nous l'avons mis en évidence plus haut (cf. 1.1.3), la conformation des valeurs de S_k est liée au nombre de messages reçus. Avec un seul message, S_k dessine une gouttière qui suit l'intersection du cône solution avec la surface terrestre. Avec au moins deux messages, il ne reste que deux puits centrés sur les positions nominale et image, symétriques par rapport à la trace au sol du satellite. La profondeur des puits (ou de la gouttière) est d'autant plus grande que la covariance du bruit de mesure est faible.

À cause de la non-linéarité de la fonction d'observation Doppler, l'estimateur des moindres carrés n'admet pas de solution analytique. C'est pourquoi on procède à une implémentation itérative en approximant la fonction d'observation au premier ordre :

$$z_k = g_k(x_k) + v_k \approx g_k(\hat{x}_k^0) + J_{g_k}^0 \delta x_k^0 + v_k \quad (1.10)$$

avec $\hat{x}_k^0$ une estimée initiale de la solution réelle, $\delta x_k^0 = x_k - \hat{x}_k^0$ l'écart entre la solution réelle et son estimation, et $J_{g_k}^0 = \frac{\partial g_k}{\partial x_k} \big|_{\hat{x}_k^0} \in \mathbb{R}^{m_k \times n}$ la jacobienne évaluée en $\hat{x}_k^0$ de la fonction d'observation. Ainsi, la somme des écarts quadratiques se réécrit $S'_k(\delta x_k^0) = \|z_k - g_k(\hat{x}_k^0) - J_{g_k}^0 \delta x_k^0\|_{R_k^{-1}}^2$ dont le minimum par rapport à δx_k^0 est atteint en

$$\delta \hat{x}_k^0 = \left(J_{g_k}^{0T} R_k^{-1} J_{g_k}^0 \right)^{-1} J_{g_k}^{0T} R_k^{-1} (z_k - g_k(\hat{x}_k^0)) \quad (1.11)$$

avec la covariance de l'erreur d'estimation associée

$$\text{cov}(\delta x_k^0 - \delta \hat{x}_k^0) = \text{cov}(x_k - \hat{x}_k^0) = \left(J_{g_k}^{0T} R_k^{-1} J_{g_k}^0 \right)^{-1}. \quad (1.12)$$

La prochaine estimation de la solution est finalement $\hat{x}_k^1 = \hat{x}_k^0 + \delta \hat{x}_k^0$. L'estimée $\hat{x}_k^1$ est alors réutilisée comme estimée initiale pour calculer $\hat{x}_k^2$ et ainsi de suite. Ce raffinement itératif de la solution se résume à la méthode Gauss-Newton [Kelley, 1999] et peut être stoppé à l'itération j quand la variation de S' devient suffisamment petite ou qu'un nombre maximum d'itérations est atteint. À chaque itération, la nouvelle estimation de la position permet de rafraichir la valeur de l'altitude à partir du modèle de terrain.

Expression de la jacobienne de $\mathcal{H}$

La jacobienne $J_{\mathcal{H}}$ de la fonction d'observation $\mathcal{H}$ s'exprime sous la forme

$$J_{\mathcal{H}} = \left(\frac{\partial \mathcal{H}}{\partial \lambda} \quad \frac{\partial \mathcal{H}}{\partial \phi} \quad \frac{\partial \mathcal{H}}{\partial h} \quad \frac{\partial \mathcal{H}}{\partial f_t} \right) = \left(\frac{\partial \mathcal{H}}{\partial (x,y,z)} \frac{\partial (x,y,z)}{\partial (\lambda,\phi,h)} \quad \frac{\partial \mathcal{H}}{\partial f_t} \right). \quad (1.13)$$

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

Avec les notations introduites dans la section 1.1.1, il advient que :

$$\frac{\partial(x, y, z)}{\partial(\lambda, \phi, h)} = \begin{pmatrix} \frac{\partial x}{\partial \lambda} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial h} \\ \frac{\partial y}{\partial \lambda} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial h} \\ \frac{\partial z}{\partial \lambda} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial h} \end{pmatrix} = \begin{pmatrix} -GE \cos \Theta \sin \lambda & -GE \sin \Theta \cos \lambda \frac{\partial \Theta}{\partial \phi} & \cos \Theta \sin \lambda \\ GE \cos \Theta \cos \lambda & -GE \sin \Theta \sin \lambda \frac{\partial \Theta}{\partial \phi} & \cos \Theta \sin \lambda \\ 0 & GP \cos \Theta \frac{\partial \Theta}{\partial \phi} & \sin \Theta \end{pmatrix}$$

où $\frac{\partial \Theta}{\partial \phi} = \frac{GP}{GE} / \left[\left(1 - \frac{GP^2}{GE^2} \right) \cos^2 \phi + \frac{GP^2}{GE^2} \right]$. Encore une fois, l'altitude h est une donnée du calcul et la dernière colonne de $\frac{\partial(x, y, z)}{\partial(\lambda, \phi, h)}$ est donc en réalité omise pour le traitement ARGOS. Pour le reste des dérivées partielles, le calcul diffère selon que l'on dérive directement depuis (1.3) :

$$\frac{\partial \mathcal{H}}{\partial(x, y, z)} = -\frac{f_t}{c \|\vec{r}\|} \vec{V}_S^T \cdot \left(\frac{\vec{r} \cdot \vec{r}^T}{\|\vec{r}\|^2} - \mathbb{I}_3 \right)$$

ou, en tenant compte de la durée de comptage τ , que l'on remplace la vitesse radiale relative v_r par $\frac{R_f - R_d}{\tau}$ dans l'équation d'observation (cf. paragraphe 1.1.2) :

$$\frac{\partial \mathcal{H}}{\partial(x, y, z)} = \frac{f_t}{c \tau} \begin{bmatrix} \left(\frac{x - x_{s,f}}{R_f} - \frac{x - x_{s,d}}{R_d} \right) & \left(\frac{y - y_{s,f}}{R_f} - \frac{y - y_{s,d}}{R_d} \right) & \left(\frac{z - z_{s,f}}{R_f} - \frac{z - z_{s,d}}{R_d} \right) \end{bmatrix}.$$

Enfin, on a dans tous les cas :

$$\frac{\partial \mathcal{H}}{\partial f_t} = 1 - \frac{v_r}{c}.$$

Pour en revenir aux équations de l'ajustement par moindres carrés, la jacobienne de la fonction d'observation pour l'ensemble des messages d'un passage est $J_{g_k} = [J_{\mathcal{H}_k^1}^T, \dots, J_{\mathcal{H}_k^{m_k}}^T]^T$, issue de l'empilement des jacobiniennes des fonctions d'observation associées à chaque message du passage.

1.2.2. Performances de localisation a priori

Les performances de localisation a priori sont affectées par la *géométrie d'observation* du passage. C'est un terme large désignant la façon dont la balise est observée par le satellite. La géométrie dépend des positions relatives de $\vec{B}$ et $\vec{S}$ ainsi que de la quantité de messages reçus et de leur distribution dans le passage. Pour une géométrie donnée, d'autres effets liés à la qualité de l'émetteur ou à l'environnement viennent moduler les performances.

Influence de la géométrie d'observation

L'angle d'observation entre le satellite et la balise est caractérisé par la distance à la trace sub-satellite qui est une composante de leur position relative. Afin de bien saisir son influence, nous avons simulé un passage satellite (orbite parfaitement polaire) où quatre messages sont reçus. Ceux-ci sont distribués uniformément sur le passage de 15 minutes. Pour simplifier, nous n'avons pas tenu compte du cercle de visibilité du satellite qui réduit le nombre de messages reçus lorsque la balise est à l'horizon. La figure 1.6 présente les écarts-types des erreurs sur chaque variable estimée en fonction de la distance à la trace. Ces chiffres sont extraits de la matrice $(J_g^T J_g)^{-1}$ qui donne les covariances σ_λ^2 , σ_ϕ^2 et $\sigma_{f_r}^2$ de l'erreur de localisation par Hertz d'erreur sur la fréquence de réception. La jacobienne J_g est évaluée au point où se trouve la balise et avec sa fréquence

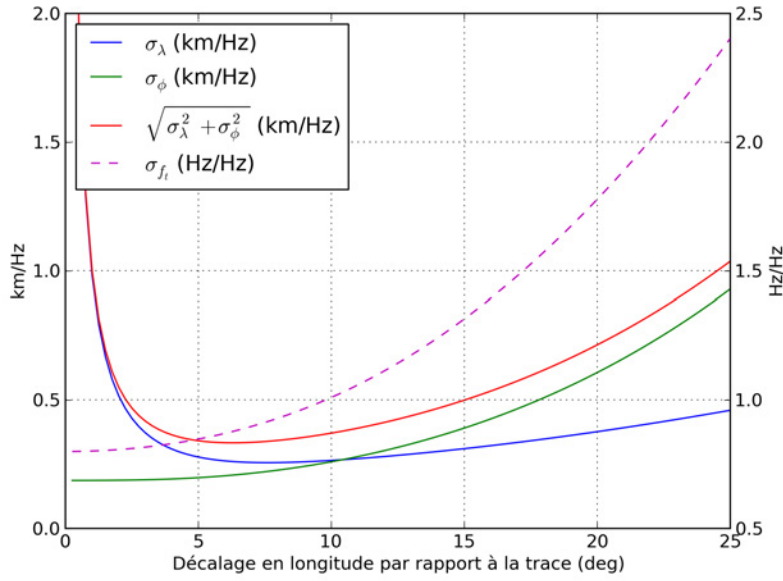


FIGURE 1.6.: Écart-types des erreurs de localisation en fonction de la distance à la trace (par Hertz de bruit à 1 sigma sur la fréquence de réception).

d'émission (cela revient à faire tendre l'indice d'itération j du raffinement itératif des moindres carrés vers l'infini).

On remarque d'emblée une singularité sous la trace avec une divergence des erreurs à cause d'une perte d'observabilité sur la longitude λ . Dans cette configuration particulière, les positions du satellite aux instants des mesures et la position de la balise sont dans le même plan. Cela se traduit par une annulation des dérivées en longitude de la jacobienne qui n'est alors plus de rang plein lors de l'inversion du système. En outre, l'erreur globale $\sqrt{\sigma_\lambda^2 + \sigma_\phi^2}$ atteint un minimum à 400 m vers 6 degrés. L'erreur s'accroît vers l'horizon jusqu'à quatre fois la valeur du minimum à cause d'une perte d'observabilité sur la latitude ϕ . En deçà de 10 degrés, l'erreur est orientée Est-Ouest (l'erreur en longitude est prépondérante), et au delà, Nord-Sud. Dans un cas où le cercle de visibilité est pris en compte, le nombre de messages reçus diminue avec la distance à la trace, et l'erreur s'accroît en réalité de manière discontinue : on observe des sauts chaque fois qu'un message en moins est reçu.

Dans la figure 1.7, nous avons fait varier le nombre de messages reçus en prenant soin de toujours les répartir uniformément sur les quinze minutes du passage. Sans surprise, un plus grand nombre de messages permet de réduire l'erreur. On atteint 250 m avec neuf messages. Mis à part à l'horizon où le faible nombre de messages est responsable d'une divergence plus rapide, l'ordre de grandeur de l'erreur ne varie pas significativement entre neuf et trois messages.

La manière dont les messages sont distribués dans le passage est en revanche bien plus déterminante. C'est une autre composante de la position relative entre le satellite et la balise. Nous avons généré quatre messages mais fait varier leur répartition dans la figure 1.8. On perd un facteur 20 sur l'erreur globale entre un passage constitué de messages reçus dans le premier tiers et un passage dont les messages sont uniformément distribués. De manière générale, plus on réduit l'intervalle

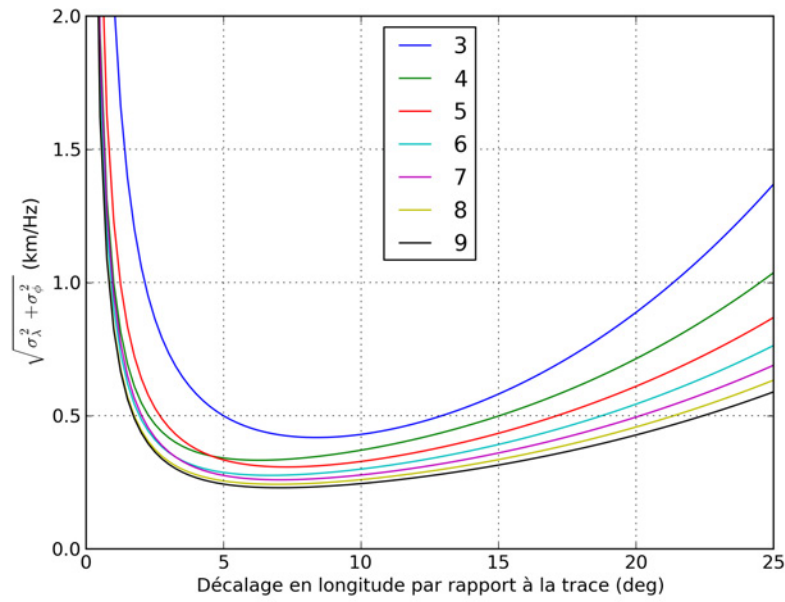


FIGURE 1.7.: Valeur de $\sqrt{\sigma_\lambda^2 + \sigma_\phi^2}$ en fonction de la distance à la trace et du nombre de messages reçus (par Hertz de bruit à 1 sigma sur la fréquence de réception).

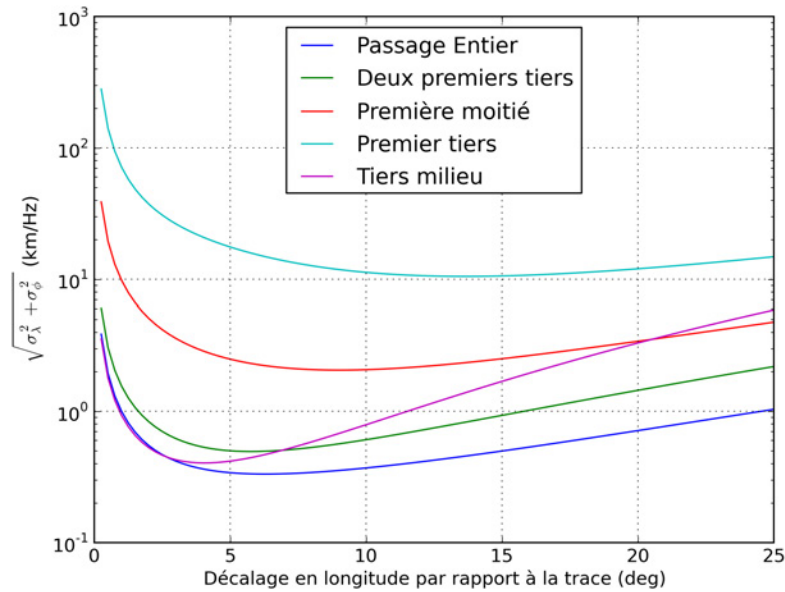


FIGURE 1.8.: Valeur de $\sqrt{\sigma_\lambda^2 + \sigma_\phi^2}$ en fonction de la distance à la trace et de la distribution dans le passage (par Hertz de bruit à 1 sigma sur la fréquence de réception). On considère quatre mesures réparties uniformément selon plusieurs configurations : sur le passage entier, les deux premiers tiers, la première moitié, le premier tiers et le tiers du milieu.

1.2. FONCTIONNEMENT DE L'ALGORITHME HISTORIQUE

de réception des messages moins la précision est bonne. Cependant, on s'aperçoit qu'il vaut mieux recevoir des messages dans le tiers du milieu plutôt que dans le premier ou dernier tiers. Les messages apportent donc plus d'information dans la partie la plus pentue de la courbe Doppler que dans les parties plus plates aux extrémités. Numériquement parlant, recevoir des messages bien répartis et/ou autour du point d'inflexion permet de renforcer l'indépendance des colonnes de la jacobienne et d'améliorer le conditionnement global du système.

En résumé, la géométrie d'observation est d'autant meilleure que la balise est vue sous un grand nombre d'angles différents par le satellite lors du passage. Elle se dégrade naturellement si peu de messages sont reçus ou s'ils sont mal répartis dans le passage (par exemple, si les messages sont reçus principalement en début de visibilité satellite à cause d'un masque). La distance à la trace sub-satellite influence aussi largement la géométrie d'observation avec un minimum d'erreur vers 6 degrés.

Effets non géométriques

Les effets géométriques sont amplifiés ou atténués selon l'intensité du bruit ambiant industriel sous le satellite. On observe par exemple des planchers de bruit très élevés dans la bande de fréquence ARGOS en Asie et autour de la Mer Noire, responsables d'un rapport signal à bruit (SNR pour *Signal to Noise Ratio*) plus faible. Ces phénomènes impactent directement la probabilité de recevoir un message à bord des satellites et conduisent à des passages tronqués où les messages ne sont pas uniformément répartis. Le comportement du porteur de balise peut également induire les mêmes effets : les animaux marins, transitant dans des mers agitées ou ne remontant que sporadiquement à la surface pour prendre leur respiration, limitent à la fois le nombre de messages et l'intervalle durant lequel ils sont reçus. La position de l'animal modifie aussi les probabilités de réception selon la forme du gain de l'antenne.

La stabilité en fréquence et en position de la balise s'ajoute à ces effets. Rappelons que le traitement suppose cette stabilité sur la durée d'un passage. De fait, une dérive en fréquence ou un déplacement importants accroissent le résidu de l'ajustement par moindres carrés car ils ne sont pas modélisés dans l'équation d'observation. Leur prise en compte est d'ailleurs une gageure car elle nécessite soit une connaissance fine des caractéristiques de l'émetteur, soit d'augmenter le nombre de variables à estimer avec la même quantité d'information (par exemple en incluant la dérive en fréquence).

Ce paragraphe se prolonge dans l'annexe A où nous analysons entre autres l'impact d'un biais sur la valeur réelle de l'altitude (cas des oiseaux) ainsi que l'effet d'un déplacement de la balise ou d'une dérive de la fréquence d'émission pendant le passage.

1.2.3. Particularités de l'implémentation

Calcul de la solution initiale $\hat{x}_k^0 = (\hat{\lambda}_k^0, \hat{\phi}_k^0, \hat{f}_{t,k}^0)$

Pour débiter le raffinement itératif de l'ajustement par moindres carrés, l'estimée initiale de la fréquence d'émission $\hat{f}_{t,k}^0$ de la plateforme est supposée égale à la dernière valeur connue. La

position initiale $(\hat{\lambda}_k^0, \hat{\phi}_k^0)$ de l'équation (1.10) est par contre estimée en utilisant le principe de la localisation géométrique de la section 1.1.3 avec un couple de mesures et la fréquence initiale $\hat{f}_{t,k}^0$. La localisation géométrique calcule deux solutions proches du nominal et de l'image que l'on ne peut encore distinguer à cette étape (cf. Fig. 1.5). Le raffinement itératif décrit ci-dessus est alors appliqué sur chacune des deux estimées initiales en utilisant l'ensemble des messages du passage.

Si la dernière fréquence estimée n'est pas disponible ou date de plus d'un mois, la localisation géométrique et les ajustements sont répétés en balayant par pas de 500 Hz une bande réduite de fréquences d'émission. Cette dernière, d'une largeur de 20 kHz, est centrée sur la moyenne des fréquences reçues pendant le passage. À l'issue du balayage de bande, la double solution de plus faible résidu est conservée. Ce processus de balayage est aussi appelé *initialisation* du traitement.

Calculs avec au moins 4 messages

Avec plus de mesures que d'inconnues, le système est surdéterminé. Une estimée de type moindres carrés de $(\lambda_k, \phi_k, f_{t,k})$ peut être calculée ainsi que la matrice de covariance de l'erreur d'estimation associée :

$$\text{cov}(x_k - \hat{x}_k^j) = (\hat{\sigma}_k^j)^2 \left(J_{gk}^j T J_{gk}^j \right)^{-1} \quad (1.14)$$

où $\hat{\sigma}_k^j$ est l'estimée non biaisée de l'écart-type du bruit de mesure déduite de

$$\hat{\sigma}_k^j = \sqrt{\frac{S(\hat{x}_k^j)}{m_k - n}}. \quad (1.15)$$

On rappelle que j est le numéro de la dernière itération de l'ajustement.

Calculs avec 3 messages

Dans ce cas, $m = n$ et la méthode des moindres carrés reste applicable. Mais, étant donné que l'estimateur du bruit d'observation (1.15) est maintenant indéfini, la seule information disponible sur l'erreur est la dilution géométrique de précision (ou GDOP pour *Geometric Dilution Of Precision*) :

$$\text{GDOP}^j \approx \sqrt{(\sigma_{\lambda,k}^j)^2 + (\sigma_{\phi,k}^j)^2} \quad (1.16)$$

où $\sigma_{\lambda,k}^j$ et $\sigma_{\phi,k}^j$ sont extraites de $\left(J_{gk}^j T J_{gk}^j \right)^{-1}$. Le GDOP est exprimé en m.Hz^{-1} après avoir effectué les conversions appropriées de degré en mètre.

Calculs avec 2 messages

Le système est sous-déterminé ce qui entrave sa résolution directe. Pour se ramener à un ensemble fini de solutions, on suppose que la fréquence de transmission de la balise est connue et égale à la dernière estimée calculée. Ainsi, pour les plateformes à faible puissance avec lesquelles deux messages sont en moyenne reçus, l'estimée de la fréquence d'émission n'est pas systématiquement rafraîchie par le traitement. Les conséquences d'un biais sur la fréquence d'émission dans le calcul sont étudiées dans l'annexe A.

Deuxième passe avec la vitesse

L'algorithme de traitement tente un second ajustement qui prend en compte la vitesse du mobile. En effet, pour des plateformes qui se déplacent rapidement en ligne droite (bateaux...), l'hypothèse de stabilité de la position pendant le passage n'est plus adaptée. Avec la première passe, une localisation a été déterminée de manière qu'une vitesse moyenne $\vec{V}_{B,k}$ depuis la dernière position puisse être calculée. La deuxième passe fait l'hypothèse que $\vec{V}_{B,k}$ est la vitesse de déplacement du mobile pendant le passage. Dans l'équation d'observation Doppler (1.3), la position de la balise devient donc spécifique à l'instant de la mesure i du passage k et elle s'écrit $\vec{B}_k^i = \vec{B}(\lambda_k, \phi_k, h_k) + (T_k^i - \bar{T}_k)\vec{V}_{B,k}$ avec $\bar{T}_k$ la date moyenne des mesures (qui est aussi la date de localisation) et T_k^i la date de la mesure i . Le système est ajusté pour déterminer une nouvelle estimation de $(\lambda_k, \phi_k, f_{t,k})$ qui n'est retenue que si le résidu correspondant est plus petit que lors de la première passe sans prise en compte de la vitesse.

1.2.4. Distribution des positions aux utilisateurs

Choix de la position nominale

Le choix de la « bonne » position entre le nominal et l'image s'opère avec l'aide de quatre tests de vraisemblance appliqués successivement aux deux solutions possibles de l'ajustement :

1. *Résidu minimum* : si la solution testée possède le plus petit résidu, le test est un succès. De fait, le test est un échec pour l'autre solution.
2. *Déplacement minimum* : l'algorithme calcule les quatre distances possibles entre les positions nominal/image du passage précédent et les positions nominal/image courantes. Si, pour la solution considérée, la distance de déplacement depuis la dernière position nominale ou image est la plus petite des quatre, le test est un succès. De fait, le test est un échec pour l'autre solution.
3. *Vitesse de déplacement* : lors de la déclaration d'une plateforme dans le système, l'utilisateur donne une vitesse maximale de déplacement. Si la vitesse de déplacement depuis la dernière position est en deçà de la vitesse maximale, le test est un succès.
4. *Continuité en fréquence* : si la variation absolue de fréquence estimée depuis le dernier passage est inférieure à 400 Hz, le test est un succès. Si ce test est un échec, l'algorithme considère au prochain traitement que la fréquence estimée n'est plus fiable et que l'algorithme doit être initialisé de nouveau par un balayage de la bande des fréquences d'émission.

La solution qui réussit le plus de tests est déclarée position nominale.

Système de classification

Si quatre messages au moins ont été reçus, les positions ARGOS sont fournies avec une caractérisation de l'erreur de localisation, supposée gaussienne, sous la forme d'une ellipse avec un demi-grand axe de longueur a , un demi-petit axe de longueur b et un angle θ (depuis le nord en direction de l'est).

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

Pour une gaussienne de moyenne $\hat{x}$ et de covariance P , on définit l'ellipsoïde à K - σ ($K > 0$) comme étant l'ensemble des vecteurs $\xi_K(\hat{x}, P) = \{ \zeta : (\zeta - \hat{x})^T P^{-1} (\zeta - \hat{x}) \leq K^2 \}$. Dans le cas d'une gaussienne bi-dimensionnelle, la probabilité de tirer un point inclus dans l'ellipse d'erreur à K - σ , centrée sur la moyenne, est égale à $1 - e^{-\frac{K^2}{2}}$. On suppose maintenant que la matrice de covariance P est tirée de (1.14), réduite à la longitude et à la latitude, de sorte que

$$P = \begin{pmatrix} \sigma_\lambda^2 & \sigma_{\lambda\phi} \\ \sigma_{\lambda\phi} & \sigma_\phi^2 \end{pmatrix}. \quad (1.17)$$

Alors (après les conversions appropriées en mètre) les paramètres a , b et θ de l'ellipse d'erreur ARGOS à K - σ sont donnés par

$$a^2 = \frac{K^2}{2} \left(\sigma_\lambda^2 + \sigma_\phi^2 + \sqrt{(\sigma_\lambda^2 + \sigma_\phi^2)^2 - 4(\sigma_\lambda^2 \sigma_\phi^2 - \sigma_{\lambda\phi}^2)} \right) \quad (1.18)$$

$$b^2 = \frac{K^2}{2} \left(\sigma_\lambda^2 + \sigma_\phi^2 - \sqrt{(\sigma_\lambda^2 + \sigma_\phi^2)^2 - 4(\sigma_\lambda^2 \sigma_\phi^2 - \sigma_{\lambda\phi}^2)} \right) \quad (1.19)$$

$$\tan 2\theta = \frac{2\sigma_{\lambda\phi}}{\sigma_\phi^2 - \sigma_\lambda^2}. \quad (1.20)$$

Les paramètres a et b sont en fait les valeurs propres de P multipliées par K , et θ est l'angle entre l'axe pointant vers le nord et le vecteur propre associé à la plus grande valeur propre de P . ARGOS délivre une ellipse à $\sqrt{2}$ - σ qui contient théoriquement 63,2 % des positions calculées. Réciproquement, on a aussi :

$$P = \begin{pmatrix} \left(\frac{a}{K}\right)^2 \sin^2 \theta + \left(\frac{b}{K}\right)^2 \cos^2 \theta & \frac{a^2 - b^2}{K^2} \cos \theta \sin \theta \\ \frac{a^2 - b^2}{K^2} \cos \theta \sin \theta & \left(\frac{a}{K}\right)^2 \cos^2 \theta + \left(\frac{b}{K}\right)^2 \sin^2 \theta \end{pmatrix}. \quad (1.21)$$

La donnée des paramètres de l'ellipse d'erreur n'est apparue qu'en 2008 pour les utilisateurs. Historiquement, l'évaluation de la qualité d'une localisation ARGOS se base sur le nombre de messages reçus pendant le passage et le *rayon d'erreur équivalent* $R = \sqrt{ab}$. Il s'agit du rayon d'un cercle de même surface que l'ellipse. Les localisations sont réparties en classes dont les caractéristiques sont disponibles dans le tableau 1.1. S'ajoute à cette liste la classe spéciale Z affectée aux

TABLE 1.1.: Classes de localisation ARGOS

Classe	Nombre de messages	Rayon d'erreur R
3	≥ 4	< 250 m
2		250 - 500 m
1		500 - 1500 m
0		≥ 1500 m
A	3	NA
B	2	

localisations pour lesquelles le calcul est impossible (pas suffisamment de mesures) ou si l'ajustement des moindres carrés n'a pas convergé (nombre maximum d'itérations atteint). Le terme de

rayon équivalent est trompeur car il n'y a pas d'équivalence mathématique entre une ellipse et son rayon d'erreur. Un même rayon R correspond à une infinité d'ellipses possibles qui se différencient par leurs rapports b/a . Bien que ce système de rayon équivalent caractérise très grossièrement l'erreur réelle, il conditionne la distribution des positions et est largement utilisé dans la communauté lors de l'analyse des données. Pour des balises à fortes puissances avec de nombreux messages reçus (bouées, bateaux), seules les localisations de classes 1, 2 ou 3 sont délivrées. Les suiveurs d'animaux, équipés de balises moins puissantes ou émettant dans des conditions plus difficiles, bénéficient généralement du *Service Plus* pour lequel toutes les positions sont distribuées.

1.3. Analyse de performance

1.3.1. Panel de test et traitement

Les performances de l'algorithme de localisation historique ont été analysées sur un large panel de balises comprenant 200 plateformes déployées aussi bien sur des animaux (terrestres, marins et oiseaux), des bateaux que des bouées dérivantes¹. Ces plateformes, toutes équipées d'un récepteur GPS, couvrent un large spectre de comportements, d'aires géographiques et de types d'émetteurs. Les positions GPS, dont les erreurs excèdent rarement quelques dizaines de mètres, sont considérées comme la vérité terrain pour estimer directement l'erreur de positionnement ARGOS.

Nous avons développé en Java un logiciel complet d'analyse chargé de :

1. récupérer les données depuis la base archive ARGOS (éphémérides satellites et mesures) ;
2. regrouper les mesures par passage satellite et calculer les positions/vitesses du satellite pour chaque mesure ;
3. à partir de ces données, recalculer les trajectoires de chaque balise en faisant appel au module de localisation moindres carrés (écrit en FORTRAN) ;
4. analyser les sorties des moindres carrés pour notamment déterminer les statistiques d'erreur par rapport à du GPS.

Le logiciel se décompose en réalité en deux projets, le premier étant capable de récupérer les données et recalculer les positions, le second procédant à l'analyse des données générées. Étant donné que le traitement est mené hors-ligne, les paramètres d'orbite sont en général de meilleure qualité qu'au moyen du traitement temps réel. Pour l'analyse d'erreur, on retient toutes les positions calculées, soit un total de plus de 100 000, excepté celles classées Z (voir Tab. 1.2). Dans la suite, l'ajustement par moindres carrés sera parfois dénommé "MC" par souci de concision.

Comme la date des positions GPS ne correspond jamais aux dates ARGOS, on interpole linéairement la trajectoire GPS aux dates des localisations ARGOS comme cela a déjà été proposé dans de précédentes études sur la précision ARGOS (par exemple [Costa et al., 2010, Witt et al., 2010]). L'erreur est la distance orthodromique² entre la position ARGOS et la position GPS interpolée. La figure 1.9 illustre ce calcul. Les erreurs de positions ARGOS estimées de cette manière incluent de fait l'erreur –souvent négligeable– de positionnement du GPS et l'erreur d'interpolation. Dans

1. Les bouées GPS fournissant les données de validation ont été financées par les bourses ONR N00014-08-1-0557 et NOAA NA10OAR4320156.

2. Distance la plus courte entre deux points de l'ellipsoïde terrestre.

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

TABLE 1.2.: Liste et caractéristiques des plateformes du jeu de données

Type de mobile	Nombre de plateformes	Nombre de localisations		Propriétaire des données et références
		LS	GPS	
Marabout <i>Leptoptilos crumeniferus</i>	5	3 472	8 250	Neil et Liz Baker (Tanzania Bird Atlas, P.O. Box 1605, Iringa, Tanzania)
Oie <i>Anser indicus</i>	55	16 097	80 980	Lucy Hawkes (Bangor University, Bangor, Gwynedd, LL57 2DG, UK) [Hawkes et al., 2011]
Gnou <i>Connochaetes taurinus</i>	10	2 029	4 144	Moses Selebatso (Western Kgalagadi Conservation Corridor Project)
Mouflons <i>Ovis canadensis</i>	13	1 819	1 704	Norv Dallin (Nevada Department of Wildlife, Eastern Region - 60 Youth Center Drive – 89801 Elko – NV, USA)
Tortue à dos plat <i>Natator depressus</i>	19	16 879	21 809	Kellie Pendoley (Pendoley Environmental Pty Ltd, 61 Kishorn Street, Mount Pleasant WA 6153)
Tortue verte <i>Chelonia mydas</i>	24	8 690	13 340	Simon Benhamou (Centre d'Ecologie Fonctionnelle et Evolutive, U.M.R. 5175 Montpellier, France) [Benhamou et al., 2011]
Lion de mer des Galapagos <i>Zalophus wollebaeki</i>	9	1 224	3 027	Daniel Costa (Department of Ecology and Evolutionary Biology Institute of Marine Sciences, Long Marine Lab University of California, Santa Cruz Santa Cruz, California, USA)
Éléphant de mer <i>Mirounga angustirostris</i> & <i>Mirounga leonina</i>	26	6 305	62 664	- Daniel Costa (cf. plus haut) [Robinson et al., 2012] - Christophe Guinet (Centre d'Études Biologiques de Chizé, 79360 Villiers-en-Bois, France) [Dragon et al., 2010, Dragon et al., 2012]
Bateau	23	20 160	36 425	CLS (8-10, rue Hermès, Parc Technologique du Canal, 31520 Ramonville Saint-Agne, France)
Bouée dérivante	44	48 733	175 633	Luca Centurioni (Scripps Institution of Oceanography, Physical Oceanography Research Division, 9500 Gilman Drive, Mail Code 0213, La Jolla, CA, 92093 USA)
Total	228	125 408	430 370	

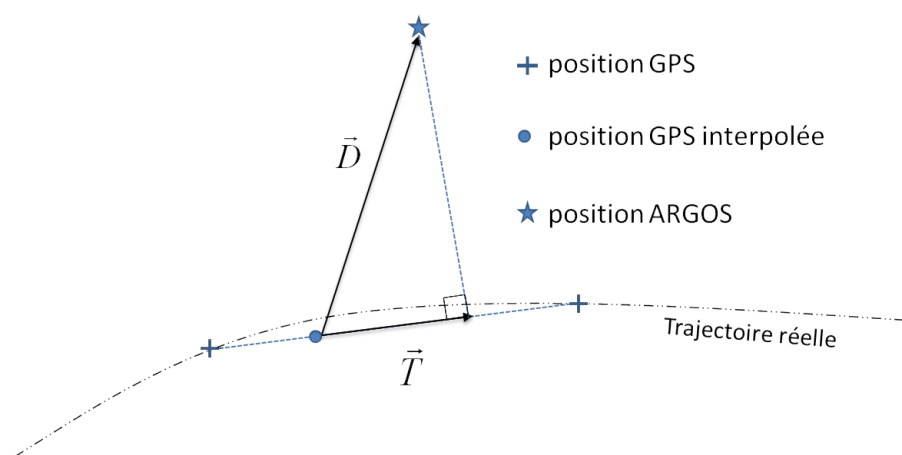


FIGURE 1.9.: Vue du dessus. L'erreur est la distance entre la localisation ARGOS et la position GPS interpolée à la même date, c'est-à-dire le module du vecteur $\vec{D}$. L'erreur tangentielle est la projection orthogonale de $\vec{D}$ sur la droite joignant les deux positions GPS successives encadrant la position GPS interpolée, c'est-à-dire le module du vecteur $\vec{T}$.

le but de minimiser cette dernière, on ne prend en compte que les erreurs des positions ARGOS datées à moins de 15 minutes d'une position GPS (excepté pour les mouflons et gnous, animaux se déplaçant en moyenne peu rapidement).

Le nombre de localisations ARGOS finalement sélectionnées est détaillé dans le tableau 1.3. Avec plus de 70 000 positions, le jeu de données final constitue une base solide pour l'évaluation de la précision du traitement moindres carrés. Seuls quelques sous-ensembles, à savoir les marabouts, les gnous et les mouflons, disposent de moins de 100 positions pour les passages à moins de 4 messages et peuvent conduire à des statistiques d'erreur empirique peu fiables.

1.3.2. Statistiques de l'erreur empirique de localisation

Le tableau 1.4 rassemble des statistiques moyennes sur les passages satellites : nombre de messages reçus, temps entre deux passages, niveau de signal reçu par le satellite et rapport signal à bruit (SNR). Le tableau 1.5 détaille des informations sur les trajectoires estimées par le traitement : GDOP, vitesse, dérive en fréquence et valeur des angles entre trois positions successives.

Lorsqu'au moins 4 messages sont reçus, les erreurs moyennes s'étendent entre 500 m et 3 km (Fig. 1.10). Ce sont les bouées, mouflons, gnous et tortues vertes qui affichent les meilleurs scores avec des erreurs en deçà de 1 km. Les valeurs les plus élevées sont atteintes par les oies, les éléphants de mer et les marabouts. Sur les passages avec deux ou trois mesures de fréquence, on retrouve un ordonnancement similaire avec des erreurs plus marquées (Fig. 1.11). Elles avoisinent les 15 km pour les oies et les éléphants de mer. Pour les meilleures balises (bouées...), la valeur moyenne de l'erreur ne descend pas en dessous de 1 km. Enfin, quel que soit le nombre de messages, les écarts-types sont du même ordre de grandeur que les erreurs moyennes.

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

TABLE 1.3.: Écart de temps maximum avec la plus proche position GPS et nombre de localisations ARGOS sélectionnées pour le calcul d'erreur

Type de mobile	Écart temps max. GPS (min)	4 mess. ou plus	2 ou 3 mess.	Total	% de localisations de Tab. 1.2
Marabout	15	378	53	431	12%
Oie	15	3 062	1 223	4 285	27%
Gnou	60	399	38	437	22%
Mouflon	60	306	94	400	22%
Tortue à dos plat	15	972	2 390	3 362	20%
Tortue verte	15	835	1 541	2 376	27%
Lion de mer	15	207	176	383	31%
Éléphant de mer	15	455	3 054	3 509	56%
Bateau	15	12 134	5 213	17 347	86%
Bouée	15	30 009	11 651	41 660	85%
Total		48 757	25 433	74 190	59%

TABLE 1.4.: Statistiques sur les passages satellites (valeurs moyennes)

Type	Nombre de message	Temps entre deux passages (heure)	Niveau de signal (dBm)	SNR (dB-Hz)
Marabout	7.4	10.1	-127	43
Oie	3.7	9.9	-125	38
Gnou	5.3	12	-125	44
Mouflon	5.4	13.4	-129	42
Tortue à dos plat	2.6	2.7	-127	45
Tortue verte	2.2	2.6	-129	42
Lion de mer	3.1	2.5	-127	44
Éléphant de mer	2.0	4.7	-131	41
Bateau	4.5	0.9	-119	47
Bouée	4.3	0.9	-119	43

Les plateformes les plus performantes se caractérisent par des vitesses moyennes de déplacement faibles (bien inférieures à 1 m/s) ainsi qu'une bonne stabilité en fréquence (principalement en dessous de 0.10 Hz/min). Les oies bougent en revanche assez rapidement (2 m/s en moyenne), avec des oscillateurs peu stables (0.31 Hz/min de dérive) et dans des environnements particulièrement bruités (rapport signal à bruit le plus faible). C'est également le cas des marabouts qui opposent malgré tout de bonnes géométries d'observation et un bon rapport signal à bruit. Rappelons aussi que les performances oiseaux sont dégradées à cause du biais entre l'altitude réelle et celle du modèle de terrain. Les éléphants de mer sont eux localisés seulement avec en moyenne 2 messages par passage (chiffre le plus bas) et des GDOP très forts de 6802 m/Hz en moyenne (chiffre le plus

1.3. ANALYSE DE PERFORMANCE

TABLE 1.5.: Statistiques sur les trajectoires calculées (valeurs moyennes)

Type	GDOP (m.Hz ⁻¹)	Vitesse (m.s ⁻¹)	Dérive en fréquence (Hz.min ⁻¹)	Angle de rotation (deg)
Marabout	985	1.6	0.32	108
Oie	3640	1.94	0.31	105
Gnou	990	0.46	0.19	108
Mouflon	1680	0.3	0.06	106
Tortue à dos plat	4242	0.58	0.08	86
Tortue verte	4921	0.65	0.06	74
Lion de mer	2588	0.97	0.20	101
Éléphant de mer	6802	1.48	0.19	72
Bateau	874	2.84	0.26	62
Bouée	720	0.55	0.05	66

haut).

1.3.3. Exemple de trajectoire

Un exemple typique d'animal est présenté en Fig. 1.12 pour un éléphant de mer (*Mirounga leonina*) suivi entre novembre 2010 et février 2011 lors d'un voyage aller-retour d'approximativement 4800 km au large des îles Kerguelen. Les éléphants de mer sont des animaux migrants dont le large nez des mâles adultes ressemble à une trompe. La plateforme émet régulièrement avec en moyenne 8,8 localisations par jour et 1,8 messages par passage. Un total de 37% des passages satellites contient 2 ou 3 messages et 56% seulement 1 message. En comparaison des lions de mer, les éléphants de mer passent peu de temps en surface et pratiquent de longues plongées [Costa et al., 2010] réduisant les opportunités de transmission. Cela se traduit par des géométries d'observation assez défavorables comme en témoigne la valeur très élevée du GDOP.

Avec le traitement moindres carrés, le calcul de la fréquence est rare car le nombre moyen de message est faible. Cela empêche le traitement d'ajuster régulièrement la fréquence d'émission de la plateforme et génère des localisations aberrantes le long de la trajectoire (cf. Fig. 1.13). Les plus grands pics (parfois dépassant le cadre de la figure) résultent d'un mauvais choix entre la localisation nominale et image.

1.3.4. Discussion

Bien que dans les meilleures conditions l'erreur de positionnement avoisine 250 m, l'analyse précédente montre que de nombreuses applications sont loin d'afficher de tels chiffres. Elle confirme les mesures de performances de [Kuhn et al., 2009, Costa et al., 2010, Witt et al., 2010] sur des balises ARGOS/GPS. En cause, la stabilité à moyen terme des oscillateurs qui est influencée par les gradients de température subis par la plateforme sur la durée d'un passage. D'autres sources

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

d'erreur s'ajoutent, à savoir le mouvement de la balise entre deux émissions ou le bruit ambiant autour de la fréquence centrale de la bande ARGOS. Ces phénomènes s'amplifient lorsque la quantité de messages reçus est faible. On a pourtant vu sur l'analyse des erreurs théoriques que l'abaissement du nombre de messages reçus ne change pas l'ordre de grandeur de l'erreur dans les mêmes conditions d'émission. La raison est sans doute à chercher du côté de la répartition des messages dans le passage : pour une période de répétition donnée, un plus faible nombre d'observations est synonyme d'une répartition moins uniforme des messages dans le passage. Cette mauvaise distribution décuple l'erreur plus que ne le fait un manque de mesures. Sur des balises difficiles, ces effets produisent des trajectoires visuellement bruitées, nécessitant des traitements complémentaires pour être exploitables [Douglas et al., 2012].

Par ailleurs, la symétrie du système par rapport à la distance à la trace conduit à la production de fausses localisations. Malgré le passage des tests de vraisemblance, il s'opère de temps à autre un mauvais choix entre les deux solutions. L'erreur liée à la distribution d'une localisation image peut atteindre le diamètre du cercle de visibilité des satellites soit 5000 km. Sur des balises à faible puissance, la fréquence des localisations images tend à s'accroître et rend plus délicate l'exploitation des données par les utilisateurs. Ils ont alors recours à des traitements hors ligne afin de diminuer cette probabilité. Sur les balises qui n'émettent que lorsqu'un satellite est visible, une localisation image peut avoir des conséquences plus importantes. Par exemple, les flotteurs sub-surface, établissant des profils de température et pression de l'océan, plongent une dizaine de jours avant de remonter à la surface pour envoyer leurs données. La remontée à lieu au moment d'un passage afin de minimiser le temps en surface. La prévision des dates des passages satellites est basée sur la dernière position transmise au flotteur par la voie descendante du système.

Enfin, quand le nombre de messages reçus est inférieur à quatre, aucune estimation d'erreur n'est disponible. Cela concerne les localisations classées A et B qui sont les plus fréquentes pour des balises animaux. Toujours pour des balises à faible puissance, les passages à un message ne sont pas exploitables pour compléter la trajectoire avec des localisations additionnelles, comme on pourrait l'espérer pour l'exemple des éléphants de mer.

Conclusions

Le traitement de localisation tirant parti d'un ajustement par moindres carrés est utilisé depuis 30 ans par les océanographes et les suiveurs d'animaux. Pour les premiers, il a permis de mieux comprendre les phénomènes océaniques. Pour les seconds, le traitement a mis en évidence des aires de reproduction et de nourrissage ainsi que les trajets afférents aux longues migrations.

La méthode souffre néanmoins de plusieurs désagréments majeurs survenant lorsque la balise opère dans des conditions difficiles : une faible précision dans les cas dégradés, la présence de localisations images, l'absence d'estimation d'erreur avec moins de quatre messages et l'impossibilité de calculer des positions avec un message. Pour ce qui est de l'erreur de localisation et des positions images, les utilisateurs ont élaboré une palette d'outils, présentés dans l'annexe B, pour nettoyer les trajectoires ARGOS moindres carrés.

Le chapitre suivant se propose d'utiliser des approches plus flexibles que l'ajustement par moindres carrés et prenant en compte la cohérence temporelle des mobiles. D'autre part, nous les appliquons directement sur les fréquences mesurées plutôt que sur les trajectoires ARGOS.

CHAPITRE 1. MOINDRES CARRÉS : L'ALGORITHME HISTORIQUE DE LOCALISATION ARGOS

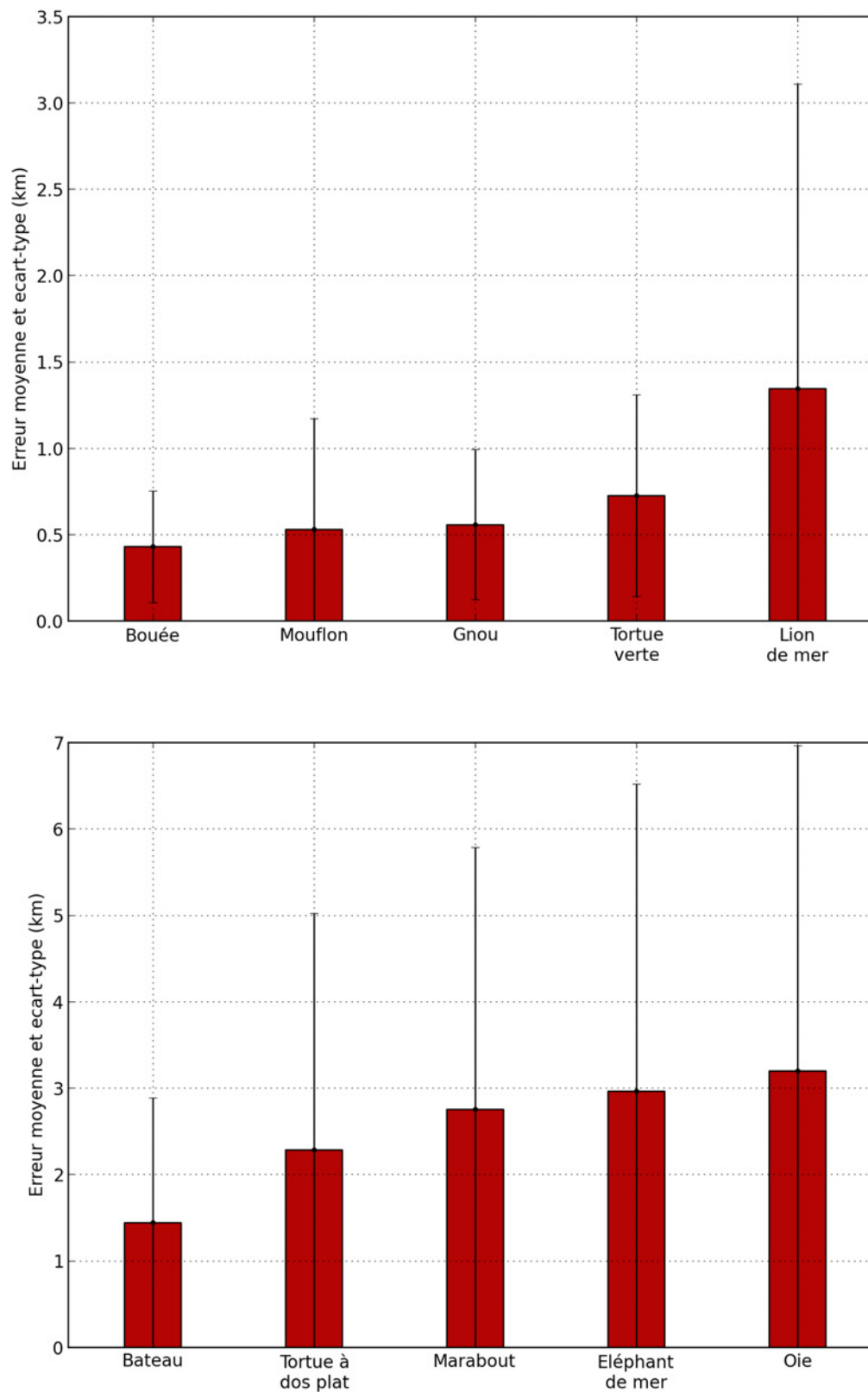


FIGURE 1.10.: Erreurs de localisation des passages à plus de 4 messages pour l'ajustement moindres carrés.

1.3. ANALYSE DE PERFORMANCE

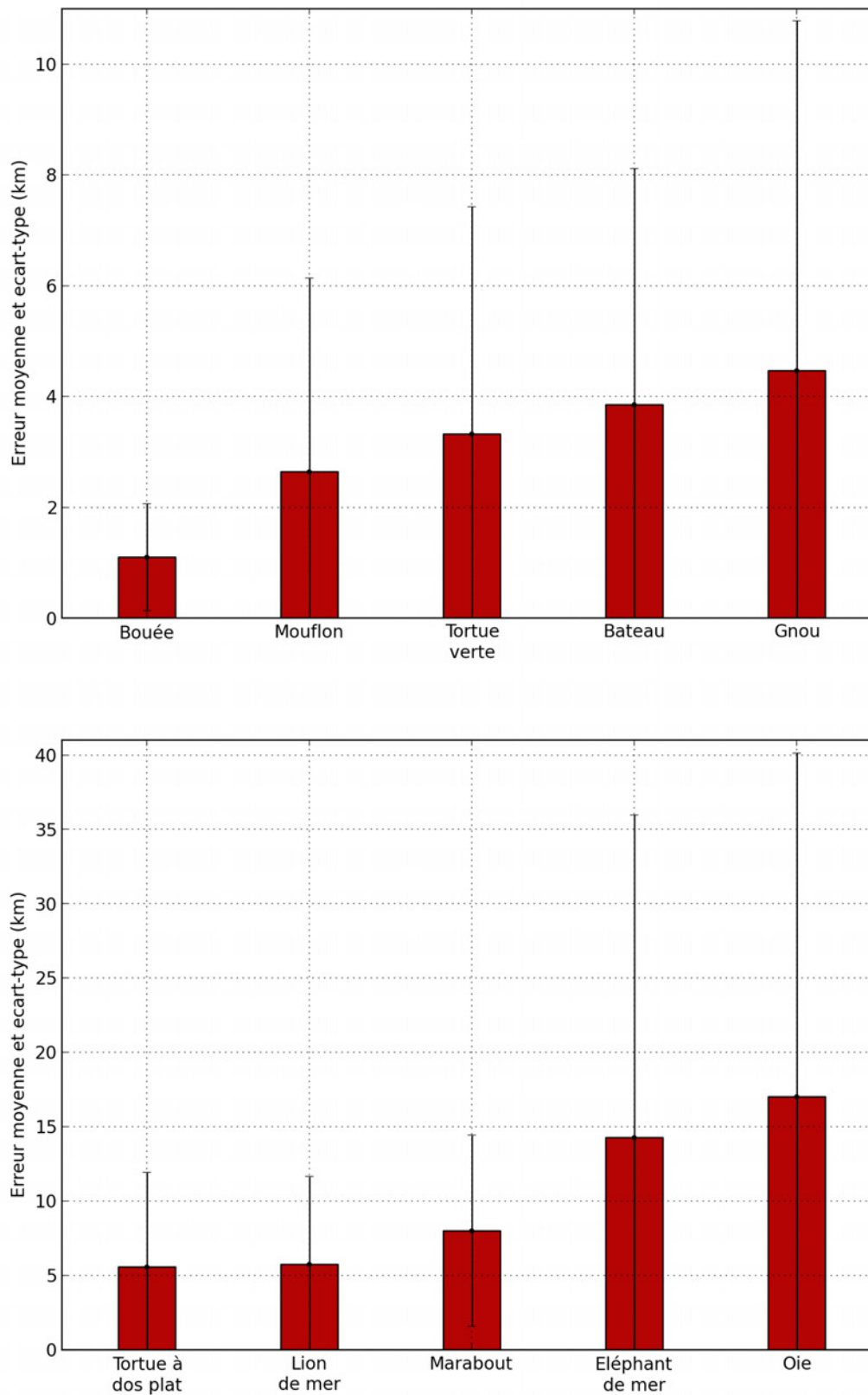


FIGURE 1.11.: Erreurs de localisation des passages à 2 ou 3 messages pour l'ajustement moindres carrés.

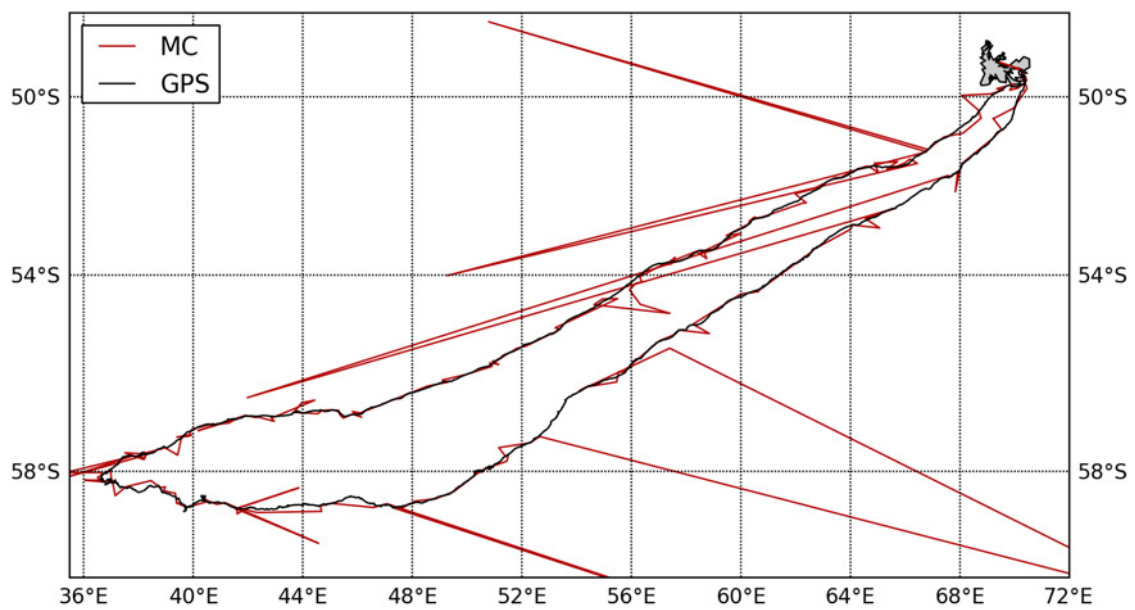


FIGURE 1.12.: Trajectoire d'éléphant de mer effectuant un aller-retour depuis les îles Kerguelen (l'aller est la partie basse du trajet).

1.3. ANALYSE DE PERFORMANCE

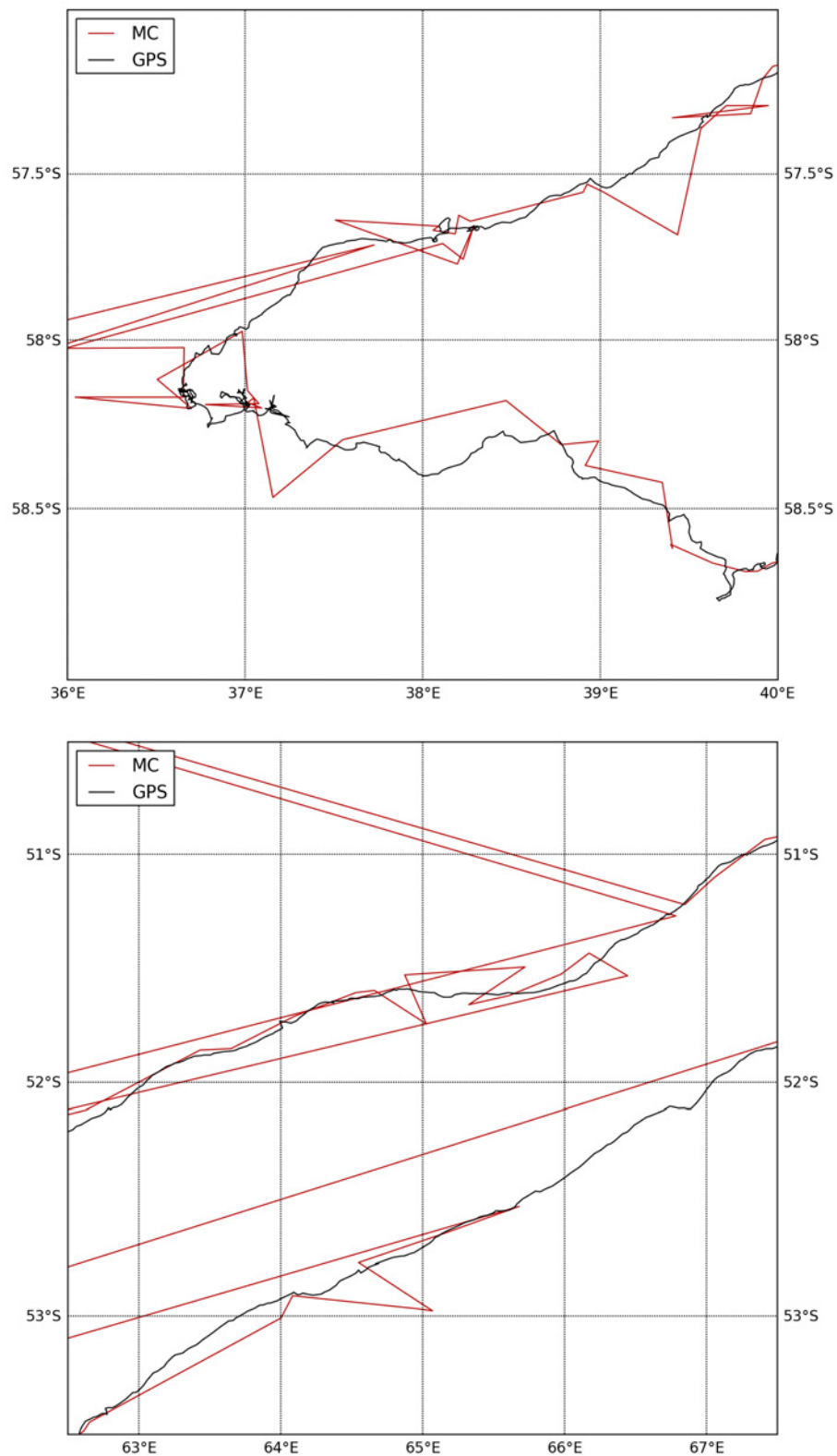


FIGURE 1.13.: Extraits de trajectoire d'éléphant de mer pour le traitement moindres carrés.

OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION



Mouflon et gnou.

Sommaire

2.1. Modélisation à espace d'état	38
2.1.1. Formulation et résolution	38
2.1.2. Avantages méthodologiques pour ARGOS	39
2.1.3. Modèles de comportement pour la localisation	40
2.2. Filtrage multi-modèle	42
2.2.1. Formulation des systèmes à sauts markoviens	42
2.2.2. Résolution par filtrage exact	43
2.2.3. Panorama des méthodes de résolution approchées	45
2.2.4. Construction du filtre IMM	48
2.3. Lissage multi-modèle	51
2.3.1. Hypothèses de base et stratégies de lissage	52
2.3.2. Solution exacte par récursion arrière	52
2.3.3. Solution exacte par double filtrage	54
2.3.4. Panorama des méthodes de résolution sous-optimales	57
Conclusions	59

LA modélisation d'état introduit une équation qui contraint a priori la dynamique de déplacement de la balise. Cette formulation délimite un cadre propice pour élaborer des solutions alternatives à l'ajustement par moindres carrés et il s'agit ici de montrer en quoi elle est pertinente pour améliorer le calcul de position ARGOS. Nous passons ensuite en revue plusieurs dynamiques de déplacement et expliquons comment il est possible de les prendre en compte simultanément dans un cadre multi-modèle.

2.1. Modélisation à espace d'état

2.1.1. Formulation et résolution

Le problème de localisation peut être reformulé comme l'estimation récursive du vecteur d'état d'un système dynamique stochastique à temps discret. L'état réel de la plateforme $x = (\lambda, \phi, f_t, \dots)^T$ inclut la position, la fréquence d'émission et éventuellement des inconnues d'ordre plus élevé (vitesse, dérivée en fréquence, ...). À l'instant initial $k = 0$, la distribution de x_0 est caractérisée par $p(x_0) = \mathcal{N}(x_0; \hat{x}_{0|0}, P_{0|0})$. La dynamique a priori $p(x_k|x_{k-1})$ du vecteur d'état est décrite par une équation stochastique :

$$x_k = f_k(x_{k-1}) + q_{k-1} \quad (2.1)$$

avec k l'indice temporel du passage satellite, f_k la fonction de transition et $q_{k-1} \sim \mathcal{N}(\bar{q}_{k-1}, Q_{k-1})$ le bruit de dynamique entre les instants $k-1$ et k . La densité $p(z_k|x_k)$, reliant les fréquences observées z_k pendant le passage avec l'état x_k , se déduit directement de l'équation de mesure Doppler :

$$z_k = g_k(x_k) + r_k \quad (2.2)$$

où $r_k \sim \mathcal{N}(\bar{r}_k, R_k)$ représente le bruit de mesure. Pour (2.1) et (2.2), les séquences $\{f, \bar{q}, Q\}$ et $\{g, \bar{r}, R\}$ sont données, et les bruits q_k et r_k sont supposés blancs, mutuellement indépendants et indépendants de x_0 .

Un estimateur bayésien permet de déterminer au cours du temps la densité $p(x_k|z_{1:k})$ du vecteur d'état x_k conditionnellement aux mesures $z_{1:k} = (z_1, \dots, z_k)$. Partant de $p(x_k|z_{1:k})$, la densité filtrée (ou a posteriori) du vecteur d'état à l'instant k , l'estimateur procède en deux étapes de prédiction et correction pour calculer la densité filtrée $p(x_{k+1}|z_{1:k+1})$ à l'instant $k+1$ avec une nouvelle mesure z_{k+1} . La prédiction (ou mise à jour en temps) s'appuie sur la dynamique a priori pour évaluer la densité prédite $p(x_{k+1}|z_{1:k})$ avec

$$p(x_{k+1}|z_{1:k}) = \int_{x_k} p(x_{k+1}|x_k, \cancel{z_{1:k}}) p(x_k|z_{1:k}) dx_k. \quad (2.3)$$

Notons que $\cancel{z_{1:k}}$ signifie que l'on peut omettre les mesures passées car, conditionnellement à x_k , la variable aléatoire x_{k+1} est indépendante de $z_{1:k}$. Nous utiliserons régulièrement la notation $\cancel{\cdot}$ dans la suite du document pour mettre en évidence les simplifications réalisées tout au long des calculs.

La correction construit la densité finale $p(x_{k+1}|z_{1:k+1})$ en assimilant la nouvelle mesure z_{k+1} selon

$$p(x_{k+1}|z_{1:k+1}) \propto p(z_{k+1}|x_{k+1}, \cancel{z_{1:k}}) p(x_{k+1}|z_{1:k}). \quad (2.4)$$

CHAPITRE 2. OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION

Dans cette dernière étape de mise à jour par la mesure, la densité filtrée est obtenue en pondérant la densité prédite par la vraisemblance de la mesure $p(z_{k+1}|x_{k+1})$.

Si les équations de dynamique et d'observation (2.1)-(2.2) sont linéaires, et avec les hypothèses ci-dessus sur les bruits de mesure et l'état initial, la densité filtrée $p(x_k|z_{1:k})$ est aussi gaussienne. Elle est alors complètement définie par ses deux premiers moments $\hat{x}_{k|k} = E[x_k|z_{1:k}]$ (la localisation estimée de la plateforme) et $P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T|z_{1:k}]$ (la covariance de l'erreur d'estimation de la localisation associée). Dans ce cas, le filtre de Kalman [Kalman, 1960] permet de déterminer analytiquement ces moments. Entre les instants k et $k+1$, le filtre propage la dernière estimée $\hat{x}_{k|k}$ et la dernière covariance $P_{k|k}$ à l'aide de la dynamique a priori. L'estimée prédite $\hat{x}_{k+1|k}$ (associée par la fonction d'observation à la mesure prédite $\hat{z}_{k+1|k}$) et la covariance $P_{k+1|k}$ ainsi obtenues sont finalement corrigées par l'assimilation de la mesure z_{k+1} pour donner $\hat{x}_{k+1|k+1}$ et $P_{k+1|k+1}$. Par construction, le filtre de Kalman est l'estimateur, non biaisé, du minimum d'erreur quadratique moyen.

Plusieurs variantes du filtre ont été développées pour appréhender la non-linéarité de la fonction d'observation de notre problème. Le filtre de Kalman étendu [Sorenson, 1985] exploite un développement de Taylor au premier ordre de f_{k+1} et de g_{k+1} au voisinage de $\hat{x}_{k|k}$ et $\hat{x}_{k+1|k}$ respectivement. Il reconstitue ainsi la moyenne prédite selon $\hat{x}_{k+1|k} = f_{k+1}(\hat{x}_{k|k})$ et la mesure prédite selon $\hat{z}_{k+1|k} = g_{k+1}(\hat{x}_{k+1|k})$, et se ramène aux équations classiques d'un filtre de Kalman linéaire pour le calcul des covariances prédite et filtrée. Au début des années 2000, le filtre de Kalman "Unscented" (en anglais *Unscented Kalman Filter* ou UKF), sous-tendu par une linéarisation statistique à base de transformées en points sigma [Julier, 2002], a montré de meilleures performances tant sur le plan théorique que pratique [Julier and Uhlmann, 1997, Wan and Van Der Merwe, 2000]. Plutôt que de linéariser f_{k+1} et g_{k+1} , les densités $p(x_k|z_{1:k})$ et $p(x_{k+1}|z_{1:k})$ sont approximées statistiquement par un ensemble de points bien choisis. Les moyennes et les covariances sont reconstruites en propageant individuellement chaque point à travers f_{k+1} et g_{k+1} . Les équations de ces deux variantes du filtre de Kalman sont présentées dans l'annexe C.

2.1.2. Avantages méthodologiques pour ARGOS

En comparaison du traitement moindres carrés, les calculs basés sur un filtre de Kalman peuvent être menés indépendamment du nombre de mesures et délivrent toujours une caractérisation de l'erreur d'estimation du vecteur d'état à travers la matrice de covariance $P_{k|k}$. En pratique, les localisations produites avec deux ou trois messages (classées B et A respectivement) bénéficient de cette caractérisation. Avec un message reçu, il devient possible de construire une position. Enfin, lorsque deux messages sont disponibles, il n'est plus nécessaire de supposer que la fréquence d'émission est fixe (cf. paragraphe 1.2.3).

L'autre avantage majeur réside dans la capacité à limiter l'apparition des localisations images. Avec le calcul traditionnel par moindres carrés, il est possible de retirer l'ambiguïté de localisation. Cela nécessite de rassembler les mesures de fréquence de plusieurs passages satellites proches dans le temps, typiquement dans un intervalle de 30 minutes maximum (traitement *multi-passage*). Au delà de cette durée, les ajustements sont plus hasardeux car la plateforme peut s'être déplacée significativement durant le passage. En comparaison, la modélisation dans l'espace d'état introduit un lien direct entre deux passages satellites et prend en compte la cohérence temporelle du mobile

à travers la dynamique de déplacement. Rappelons que la forme de l'ellipse d'erreur — aplatissement et orientation —, déduite de la matrice de covariance, est influencée par la position relative entre le satellite et la plateforme pendant le passage (cf. paragraphe sur les performances de localisation a priori du chapitre 1). Ainsi, le point de vue du dernier passage, représenté par la dernière localisation et l'erreur associée, est partagé avec le passage courant : le filtrage se rapproche des conditions d'un traitement multi-passage.

2.1.3. Modèles de comportement pour la localisation

On se donne comme objectif un traitement capable de couvrir l'ensemble des applications du système sans avoir à particulariser les calculs selon les types de plateforme. Comme ARGOS est utilisé pour une grande variété de mobiles, le choix le plus évident pour la dynamique de déplacement consiste à utiliser un modèle de marche aléatoire pour les variables de positions λ et ϕ . Les mouvements avec une direction privilégiée sont par contre mieux modélisés par une marche aléatoire corrélée (ou dirigée) qui est une marche aléatoire sur la vitesse $(\dot{\lambda}, \dot{\phi})$. On retrouve de tels modèles lorsqu'il s'agit de retraiter des données ARGOS [Johnson et al., 2008, Patterson et al., 2010] ou de localiser des animaux marins à l'aide de mesures de profondeur de plongée [Pedersen, 2007, Thygesen et al., 2009]. Un compromis entre ces deux premières dynamiques est une marche aléatoire, dite biaisée, qui subit un transport dans une direction privilégiée (dû à des courants marins par exemple). Les modèles de déplacement basés sur les marches aléatoires et leurs dérivées sont développés dans [Codling et al., 2008] pour les applications biologiques en général. Des modèles plus précis mais moins génériques ne seront pas envisagés [Li and Jilkov, 2003]. Pour les bateaux par exemple, une caractérisation des constantes hydrodynamiques liées à la forme du bateau et l'estimation d'un plus grand nombre de variables seraient alors nécessaires. Concernant la fréquence d'émission, une hypothèse majeure faite par ARGOS est la stabilité de l'oscillateur embarqué au cours du temps. Ainsi, on considère que la fréquence reste identique entre deux passages successifs mais peut être entachée d'un bruit stationnaire additif V_f pour tenir compte des gradients de température.

Les équations des trois modèles de marches aléatoires (classique, corrélée et biaisée) sont détaillées ci-dessous.

Marche aléatoire

$$x_k = x_{k-1} + q_{k-1}, \quad (2.5)$$

$$\text{où } \bar{q}_{k-1} = 0, \quad Q_{k-1} = \begin{pmatrix} 2D_\lambda \Delta t_k & 0 & 0 \\ 0 & 2D_\phi \Delta t_k & 0 \\ 0 & 0 & V_f \end{pmatrix} \quad (2.6)$$

avec l'état $x_k = (\lambda_k, \phi_k, f_{i,k})^T$ et $\Delta t_k = t_k - t_{k-1}$ le temps écoulé depuis le dernier passage satellite. La fréquence d'émission $f_{i,k}$ est exprimée en Hertz et la longitude et la latitude en degrés décimaux. Pour une marche isotrope, les paramètres de diffusion de la marche aléatoire sont $D_\lambda = D_\phi = D$ (une fois exprimés en m^2s^{-1}). Cette équation de dynamique découle de la discrétisation sur le temps de l'équation de diffusion appliquée à la densité de probabilité $p(y, t)$ (avec $y = \lambda$

CHAPITRE 2. OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION

ou $y = \phi$) au temps t :

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial y^2}. \quad (2.7)$$

La solution fondamentale de cette équation différentielle est

$$p(y, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{y^2}{4Dt}\right) \quad (2.8)$$

de sorte que $\text{cov}(y, t_{k-1} \leq t \leq t_k) = 2D\Delta t_k$. La covariance du bruit de dynamique s'accroît donc linéairement avec le temps.

Marche aléatoire corrélée

$$x_k = \begin{pmatrix} 1 & 0 & \frac{\Delta t_k}{d_e \cos(\phi_{k-1})} & 0 & 0 \\ 0 & 1 & 0 & \frac{\Delta t_k}{d_e} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + q_{k-1}, \quad (2.9)$$

$$\text{où } \bar{q}_{k-1} = 0, \quad Q_{k-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2D'_\lambda \Delta t_k & 0 & 0 \\ 0 & 0 & 0 & 2D'_\phi \Delta t_k & 0 \\ 0 & 0 & 0 & 0 & V_f \end{pmatrix}$$

avec l'état $x_k = (\lambda_k, \phi_k, \dot{\lambda}_k, \dot{\phi}_k, f_{t,k})^T$ où les vitesses correspondantes $(\dot{\lambda}_k, \dot{\phi}_k)$ sont exprimées en m.s^{-1} . Le paramètre d_e est la distance d'un degré en longitude à l'équateur (approximativement 110 km). Les paramètres de diffusion sur la vitesse sont $D'_\lambda = D'_\phi = D'$ ($\text{m}^2.\text{s}^{-3}$) pour une marche isotrope sur la vitesse.

Marche aléatoire biaisée

Plutôt que d'intégrer la vitesse $v_k = (\dot{\lambda}_k, \dot{\phi}_k)^T$ dans le vecteur d'état, on suppose que la vitesse prédite $\hat{v}_{k|k-1}$ utilisée dans la mise à jour en temps préalablement à l'assimilation de la mesure z_k est donnée. Le modèle complet s'écrit, avec $x_k = (\lambda_k, \phi_k, f_{t,k})^T$:

$$x_k = x_{k-1} + \begin{pmatrix} \frac{\Delta t_k}{d_e \cos(\phi_{k-1})} & 0 \\ 0 & \frac{\Delta t_k}{d_e} \\ 0 & 0 \end{pmatrix} \hat{v}_{k|k-1} + q_{k-1}, \quad (2.10)$$

$$\bar{q}_{k-1} = 0, \quad Q_{k-1} = \begin{pmatrix} 2D_\lambda \Delta t_k & 0 & 0 \\ 0 & 2D_\phi \Delta t_k & 0 \\ 0 & 0 & V_f \end{pmatrix}.$$

Cela équivaut à une marche aléatoire avec un biais dépendant du temps. Ce modèle est associé à l'équation d'advection-diffusion

$$\frac{\partial p}{\partial t} = -v \frac{\partial p}{\partial y} + D \frac{\partial^2 p}{\partial y^2} \quad (2.11)$$

dont la solution fondamentale est

$$p(y, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{(y - vt)^2}{4Dt}\right). \quad (2.12)$$

La moyenne du déplacement est $E[y, t_{k-1} \leq t \leq t_k] = vt$ et sa covariance $cov(y, t_{k-1} \leq t \leq t_k) = 2D\Delta t_k$.

Quel modèle choisir ?

Bien sûr, un unique modèle de marche aléatoire ou de marche dirigée ne peut reproduire fidèlement l'ensemble des comportements des plateformes ARGOS voire même d'une plateforme pendant sa durée de vie. Une meilleure approche est d'admettre que les mobiles peuvent adopter potentiellement plusieurs dynamiques : un bateau peut conduire des opérations de pêche puis revenir tout droit au port, alors qu'un animal marin peut résider dans des aires de nourrissage pendant plusieurs mois avant d'effectuer une longue migration vers une zone de reproduction. Ainsi, les vitesses observées peuvent varier de manière significative le long de la trajectoire et introduire une corrélation temporelle entre plusieurs incréments successifs de positions, en opposition avec une dynamique a priori de type marche aléatoire avec des bruits blancs additifs. Nous sommes donc poussés à envisager un ensemble de modèles dynamiques ou, de manière équivalente, un banc de filtres.

2.2. Filtrage multi-modèle

2.2.1. Formulation des systèmes à sauts markoviens

Si les équations du modèle d'évolution et/ou d'observation évoluent d'un instant sur l'autre, il est plus pertinent de reformuler le problème dans le cadre des systèmes à sauts markoviens. Par exemple, sur des systèmes dont les capteurs sont sujets aux pannes, considérer plusieurs modèles d'observation permet d'adapter le processus de mesure en modulant l'amplitude des bruits [Rapoport and Oshman, 2004]. Pour le suivi de mobile, il devient possible de tenir compte des transitions entre un modèle de marche aléatoire et un modèle avec une direction de déplacement privilégiée [Bar-Shalom and Li, 1993].

Les systèmes à sauts markoviens (en anglais *Markovian switching systems* ou *jump Markov systems*) sont décrits par un banc de *modèles* partageant le même vecteur d'état et correspondant à des *modes*¹ de comportement possibles [Li and Jilkov, 2005]. L'ensemble des M modes est noté $\mathcal{M}$ et la suite des modes forme une chaîne de Markov homogène. Pour chaque pas de temps k , $m_k = j$ ou m_k^j désigne l'évènement que le mode $j \in \mathcal{M}$ est actif durant la période $(t_{k-1}, t_k]$. Lorsque m_k^j se produit, la dynamique de l'état continu $p(x_k | x_{k-1}, m_k^j)$ et son lien avec la mesure $p(z_k | x_k, m_k^j)$ sont

1. La littérature fait la distinction entre modèle et mode. Le mode désigne le comportement réel du système alors que le modèle est sa représentation imparfaite sous forme mathématique. Par exemple, si un animal effectue une migration, il est dans un mode de migration et la représentation mathématique de ce mouvement peut être un modèle de marche aléatoire corrélée.

CHAPITRE 2. OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION

décrits par le modèle d'état

$$x_k = f_k^j(x_{k-1}) + q_{k-1}^j, \quad z_k = g_k^j(x_k) + r_k^j, \quad (2.13)$$

où $f_k^j(\cdot)$, $g_k^j(\cdot)$ sont données et q_{k-1}^j , r_k^j représentent les bruits de dynamique et de mesure. Les probabilités initiales et les probabilités de transition des modes, elles aussi connues, sont données par

$$P(m_0^j) = \mu_0^j; P(m_{k+1}^i | m_k^j) = \pi_{ji}. \quad (2.14)$$

Conditionnellement au mode j , l'état continu à l'instant initial et les bruits sont conjointement gaussiens. Leurs statistiques sont données par

$$\forall k, k', E \left[\begin{pmatrix} x_0 | m_0^j \\ q_k^j \\ r_{k'}^j \end{pmatrix} \right] = \begin{pmatrix} \hat{x}_{0|0}^j \\ 0 \\ 0 \end{pmatrix}; \quad (2.15)$$

$$E \left[\begin{pmatrix} x_0 | m_0^j - \hat{x}_{0|0}^j \\ q_k^j \\ r_k^j \end{pmatrix} \begin{pmatrix} x_0 | m_0^j - \hat{x}_{0|0}^j \\ q_{k'}^j \\ r_{k'}^j \end{pmatrix}^T \right] = \begin{pmatrix} P_{0|0}^j & 0 & 0 \\ 0 & Q_k^j \delta_{k,k'} & 0 \\ 0 & 0 & R_k^j \delta_{k,k'} \end{pmatrix} \quad (2.16)$$

avec $\delta_{k,k'}$ le symbole de Kronecker. Ainsi, la densité de l'état continu x_0 à l'instant de départ est une mixture de M gaussiennes $p(x_0) = \sum_{j \in \mathcal{M}} p(x_0 | m_0^j) P(m_0^j)$.

Le vecteur d'état complet d'un système à sauts markoviens est hybride car il inclut l'état continu x et l'état discret m . À l'instant k , il est dorénavant noté $\xi_k^j = (x_k, m_k^j)$. La dynamique a priori sur l'état complet s'écrit par définition $p(\xi_{k+1}^i | \xi_k^j) = p(x_{k+1}, m_{k+1}^i | x_k, m_k^j) \triangleq p(x_{k+1} | m_{k+1}^i, x_k, m_k^j) P(m_{k+1}^i | m_k^j, x_k)$. La figure 2.1 propose une représentation graphique du processus de génération de l'état et de la mesure. Dans le cas particulier où le nombre de modes est réduit à un, on se retrouve bien dans la configuration d'une modélisation à espace d'état classique (2.1)-(2.2).

2.2.2. Résolution par filtrage exact

Pour chaque instant k , la densité de l'état continu x_k conditionné sur l'ensemble des mesures jusqu'à l'instant k' s'écrit comme la mixture [Gustafsson, 2000] :

$$p(x_k | z_{1:k'}) = \sum_{j_{0:k} \in \mathcal{M}^{k+1}} p(x_k | m_{0:k} = j_{0:k}, z_{1:k'}) P(m_{0:k} = j_{0:k} | z_{1:k'}). \quad (2.17)$$

Les termes de la mixture sont les densités a posteriori de l'état continu conditionné sur les observations jusqu'à k' et sur toutes les séquences possibles de modes jusqu'à k . Chaque terme est pondéré par la probabilité a posteriori des séquences de modes possibles jusqu'à k conditionnées sur les observations jusqu'à k' . Dans un contexte de filtrage, on a $k = k'$ et la densité postérieure de x_k est obtenue en combinant les densités de l'état continu filtrées selon toutes les séquences de modes possibles entre l'instant initial et l'instant k . À cet égard, le nombre de termes de la mixture (2.17) croît de manière exponentielle avec k .

Il est possible de calculer de manière récursive la densité filtrée de l'état continu. Pour clarifier les développements qui suivent, on décompose la densité selon le modèle actif au temps k de

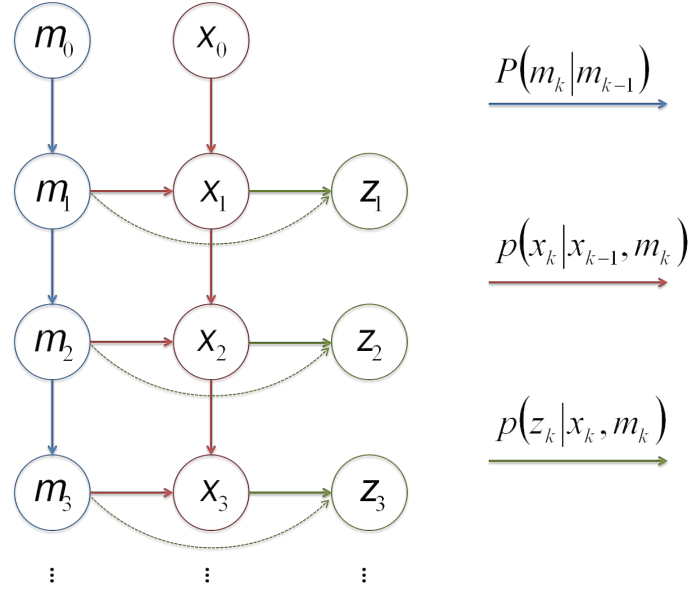


FIGURE 2.1.: Graphique de dépendance d'un système à sauts markoviens avec m l'état discret (mode), x l'état continu et z la mesure. L'état complet du système est (x, m) . Les flèches bleues sont associées aux probabilités de transition $P(m_k | m_{k-1})$, les rouges aux modèles de transition $p(x_k | x_{k-1}, m_k)$ et les vertes aux modèles d'observation $p(z_k | x_k, m_k)$.

façon que $p(x_k | z_{1:k}) = \sum_{j \in \mathcal{M}} p(x_k | m_k^j, z_{1:k}) P(m_k^j | z_{1:k})$. Au début d'un cycle de filtrage, on suppose que la mixture $p(x_k | m_k^j, z_{1:k}) = \sum_{j_{0:k-1} \in \mathcal{M}^k} p(x_k | m_{0:k} = j_{0:k}, z_{1:k}) P(m_{0:k-1} = j_{0:k-1} | m_k^j, z_{1:k})$ et la probabilité $P(m_k^j | z_{1:k}) = \sum_{j_{0:k-1} \in \mathcal{M}^k} P(m_{0:k} = j_{0:k} | z_{1:k})$ sont connues $\forall j \in \mathcal{M}$. La croissance géométrique du nombre d'hypothèses s'observe plus ouvertement lorsqu'on déroule ce cycle [Blom and Bar-Shalom, 1988] :

1. $\forall i \in \mathcal{M}, \{P(m_k^j | z_{1:k})\}_{j \in \mathcal{M}} \xrightarrow{\text{Prédiction}} P(m_{k+1}^i | z_{1:k})$
2. $\forall i \in \mathcal{M}, \{p(x_k | m_k^j, z_{1:k})\}_{j \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k | m_{k+1}^i, z_{1:k})$
3. $\forall i \in \mathcal{M}, p(x_k | m_{k+1}^i, z_{1:k}) \xrightarrow{\text{Prédiction}} p(x_{k+1} | m_{k+1}^i, z_{1:k})$
4. $\forall i \in \mathcal{M}, p(x_{k+1} | m_{k+1}^i, z_{1:k}) \xrightarrow{\text{Correction}} p(x_{k+1} | m_{k+1}^i, z_{1:k+1})$
5. $\forall i \in \mathcal{M}, P(m_{k+1}^i | z_{1:k}) \xrightarrow{\text{Correction}} P(m_{k+1}^i | z_{1:k+1})$

L'étape 1, par exemple, se lit « construire les probabilités prédites $P(m_{k+1}^i | z_{1:k})$ de l'état discret à l'instant $k+1$ pour chaque mode i à partir de l'ensemble $\{P(m_k^j | z_{1:k})\}_{j \in \mathcal{M}}$ des probabilités filtrées à l'instant k ». Les équations du cycle sont détaillés ci-dessous :

1. Prédire les probabilités des modes à l'instant $k+1$ avec

$$\forall i \in \mathcal{M}, P(m_{k+1}^i | z_{1:k}) = \sum_{j \in \mathcal{M}} P(m_{k+1}^i | m_k^j, z_{1:k}) P(m_k^j | z_{1:k}) \quad (2.18)$$

en tenant compte des hypothèses de travail sur la suite de modes.

2. Construire la *densité d'interaction* $p(x_k|m_{k+1}^i, z_{1:k})$:

$$\forall i \in \mathcal{M}, p(x_k|m_{k+1}^i, z_{1:k}) = \sum_{j \in \mathcal{M}} p(x_k|m_k^j, \cancel{m_{k+1}^i}, z_{1:k}) P(m_k^j|m_{k+1}^i, z_{1:k}) \quad (2.19)$$

en exploitant l'indépendance de m_{k+1}^i et x_k si m_k^j est connu². La densité d'interaction $p(x_k|m_{k+1}^i, z_{1:k})$ est construite dans la perspective d'y appliquer par la suite un filtre correspondant au mode i pour assimiler la mesure z_{k+1} . Par ailleurs, les *probabilités d'interaction* $P(m_k^j|m_{k+1}^i, z_{1:k})$ vérifient

$$P(m_k^j|m_{k+1}^i, z_{1:k}) = \frac{P(m_{k+1}^i|m_k^j, \cancel{z_{1:k}})P(m_k^j|z_{1:k})}{P(m_{k+1}^i|z_{1:k})}. \quad (2.20)$$

3. Prédire la densité de x_{k+1} conditionnellement à chaque mode avec

$$\forall i \in \mathcal{M}, p(x_{k+1}|m_{k+1}^i, z_{1:k}) = \int_{x_k} p(x_{k+1}|x_k, m_{k+1}^i, \cancel{z_{1:k}}) p(x_k|m_{k+1}^i, z_{1:k}) dx_k \quad (2.21)$$

en tenant compte des hypothèses de travail sur la dynamique de l'état continu.

4. et 5. Terminer la récursion en assimilant la mesure z_{k+1} selon

$$\forall i \in \mathcal{M}, p(x_{k+1}|m_{k+1}^i, z_{1:k+1}) \propto p(z_{k+1}|x_{k+1}, m_{k+1}^i, \cancel{z_{1:k}}) p(x_{k+1}|m_{k+1}^i, z_{1:k}) \quad (2.22)$$

et

$$\forall i \in \mathcal{M}, P(m_{k+1}^i|z_{1:k+1}) \propto p(z_{k+1}|m_{k+1}^i, z_{1:k}) P(m_{k+1}^i|z_{1:k}). \quad (2.23)$$

L'étape 2 forme la phase d'interaction (*mixing* en anglais) car c'est à ce moment que les densités conditionnées sur m_k se "mélangent" et que le nombre de termes grandit : à l'issue de l'équation (2.19) si $p(x_k|m_k^j, z_{1:k})$ est déjà une mixture de N densités, $p(x_k|m_{k+1}^i, z_{1:k})$ sera une mixture plus large de $N \times M$ termes. Comme la mixture de l'état continu à l'instant initial est composée de M termes, ce nombre s'élève bien à M^{k+1} à l'instant k . Les étapes 3 et 4 sont les étapes de prédiction et correction d'un estimateur bayésien. Elles sont réalisées en appliquant sur chaque terme de la densité d'interaction $p(x_k|m_{k+1}^i, z_{1:k})$ un filtre associé au mode i .

La structure du problème est illustrée de manière équivalente par un arbre dont le nombre de branches est démultiplié avec le temps (cf. Fig. 2.2). Entre l'instant initial et l'instant k , chaque branche représente une séquence de modes possibles $m_{0:k} = j_{0:k}$ de probabilité a posteriori $P(m_{0:k} = j_{0:k}|z_{1:k})$. Le filtrage de l'état continu conditionné sur la séquence de la branche donne la densité a posteriori $p(x_k|m_{0:k} = j_{0:k}, z_{1:k})$.

2.2.3. Panorama des méthodes de résolution approchées

Comme conséquence de la croissance géométrique du nombre de termes et du nombre de filtres nécessaires, l'évaluation exacte de la densité a posteriori de l'état continu est limitée par la puissance de calcul et de stockage. Des méthodes sous-optimales de résolution ont été élaborées afin de conserver une complexité constante.

2. cf. le diagramme de dépendance 2.1, ou bien remarquer que $p(x_k|m_k^j, m_{k+1}^i, z_{1:k}) = p(m_{k+1}^i|m_k^j, \cancel{z_{1:k}}) p(x_k|m_k^j, z_{1:k}) / p(m_{k+1}^i|m_k^j, \cancel{z_{1:k}}) = p(x_k|m_k^j, z_{1:k})$.

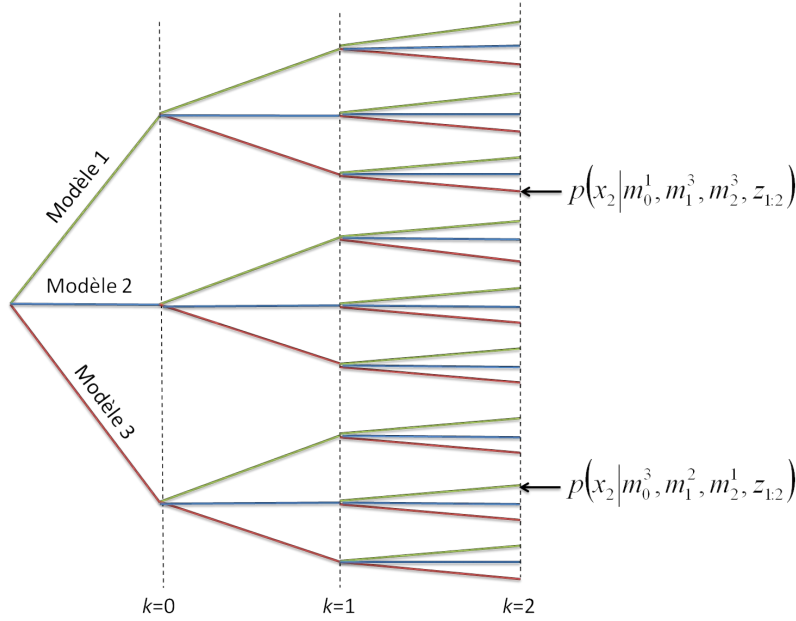


FIGURE 2.2.: Arbre des séquences possibles entre l'instant initial et l'instant $k = 2$ pour trois modèles. Sur le graphique, les densités associées aux séquences de modes (1,3,3) et (3,2,1) sont données. La densité globale de l'état à $k = 2$ est la mixture des densités associées à chaque branche.

Concevoir une méthode de résolution sous-optimale consiste à répondre en fait à une double question : *quand* et *comment* réduire le nombre de termes au sein du cycle de filtrage exact ? Dans la récursion de la section précédente, il est en effet possible de réduire le nombre d'hypothèses (ou de termes de la mixture) au sein de chaque étape manipulant la densité de l'état continu. Il existe plusieurs manières d'en limiter le nombre et elles se répartissent en deux grandes familles : la décimation ou la fusion des branches de l'arbre des séquences possibles [Li and Jilkov, 2005].

Décimation

La décimation consiste à ne garder à la toute fin du cycle de récursion qu'un nombre réduit de séquences de modes ou, de manière équivalente, de termes dans la mixture (cf. Fig. 2.3). La méthode *B-best* [Tugnait and Haddad, 1979, Tugnait, 1982, Tugnait, 1983] conserve uniquement les B_k "meilleures" séquences (au sens des termes de la mixture les plus probables a posteriori) : partant des B_{k-1} séquences de l'instant précédent, l'algorithme applique $B_{k-1}M$ filtres conditionnés sur les M modes possibles à l'instant k avant de sélectionner le sous-ensemble des B_k séquences les plus probables. Un inconvénient majeur du *B-best* est que la meilleure séquence de modèles sur l'horizon entier de temps peut avoir été éliminée plus tôt dans le traitement si elle se révèle très improbable à un moment donné. À la fin du cycle, les probabilités des termes retenus sont renormalisées afin d'avoir une somme toujours égale à un.

L'algorithme de Viterbi [Forney, 1973, Averbuch et al., 1991] identifie en revanche la meilleure séquence sur tout l'horizon en exploitant mieux les propriétés markoviennes de la suite des modes.

CHAPITRE 2. OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION

Viterbi retient à chaque instant M séquences mais pas nécessairement les M plus probables. Cela permet, à la différence du B -best, de garantir l'inclusion de la meilleure au final.

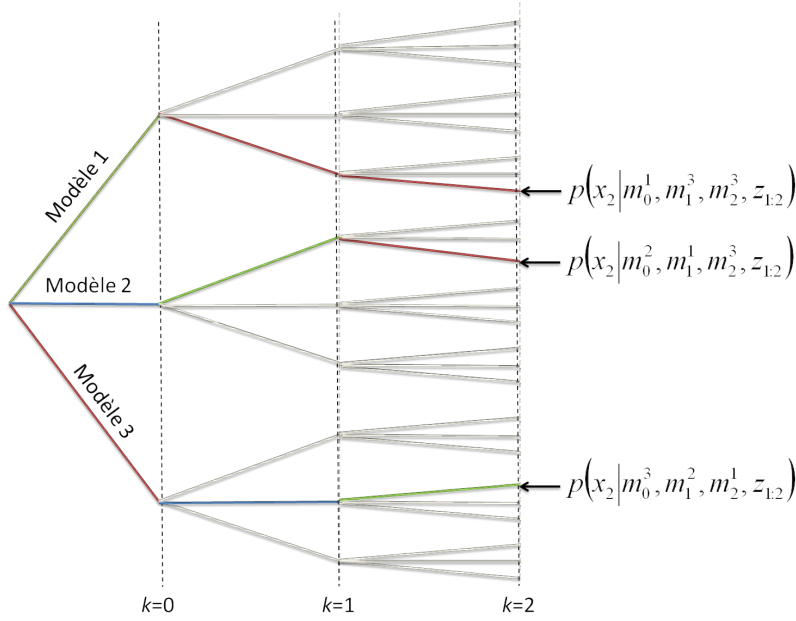


FIGURE 2.3.: Décimation des branches de l'arbre des possibles : on ne conserve qu'un sous-ensemble des termes de la mixture globale sur l'état (ici trois).

Fusion

Le nombre d'hypothèses composant la mixture (2.17) peut être réduit en *fusionnant* les densités qui sont conditionnées sur des sous-séquences de modes $m_{k-n+1:k}$ identiques entre les instants $k-n+1$ et k inclus pour $n \geq 1$ (cf. Fig. 2.4). Plus concrètement, cela revient à remplacer

$$p(x_k | m_{k-n+1:k}, z_{1:k}) = \sum_{m_{0:k-n} \in \mathcal{M}^{k-n+1}} p(x_k | m_{0:k-n}, m_{k-n+1:k}, z_{1:k}) P(m_{0:k-n}, m_{k-n+1:k} | z_{1:k}) \quad (2.24)$$

extraite de (2.17) par une densité approchée —mais plus simple à manipuler— $q(x_k | m_{k-n+1:k}, z_{1:k})$ de sorte que

$$p(x_k | z_{1:k}) \approx \sum_{m_{k-n+1:k} \in \mathcal{M}^n} q(x_k | m_{k-n+1:k}, z_{1:k}) P(m_{k-n+1:k} | z_{1:k}) \quad (2.25)$$

avec

$$P(m_{k-n+1:k} | z_{1:k}) = \sum_{m_{0:k-n} \in \mathcal{M}^{k-n+1}} P(m_{0:k-n}, m_{k-n+1:k} | z_{1:k}). \quad (2.26)$$

La nature de l'approximation est liée à la forme des densités impliquées.

Les filtres de types "Generalized Pseudo-Bayesian" d'ordre n (GPB_n) [Ackerson and Fu, 1970] s'inscrivent dans le paradigme de la fusion et, pour un banc de M modes, utilisent M^n filtres. Mais l'approche la plus répandue est sans conteste le filtre *Interacting Multiple Model* (IMM) [Blom, 1984, Blom and Bar-Shalom, 1988] qui propage au cours du temps une mixture

approximée de M gaussiennes avec seulement M filtres. Sa complexité équivaut à celle du GPB_1 mais avec des performances comparables au GPB_2 . La différence majeure entre un GPB_n et l'IMM tient dans l'instant où s'opère la réduction du nombre d'hypothèses. Pour le GPB_n , c'est en toute fin de cycle, alors que l'IMM le fait au moment de l'interaction. Mais pour les deux, la réduction utilise une approximation par moments qui remplace chaque mixture de gaussiennes considérée par une gaussienne unique ayant la même moyenne et covariance.

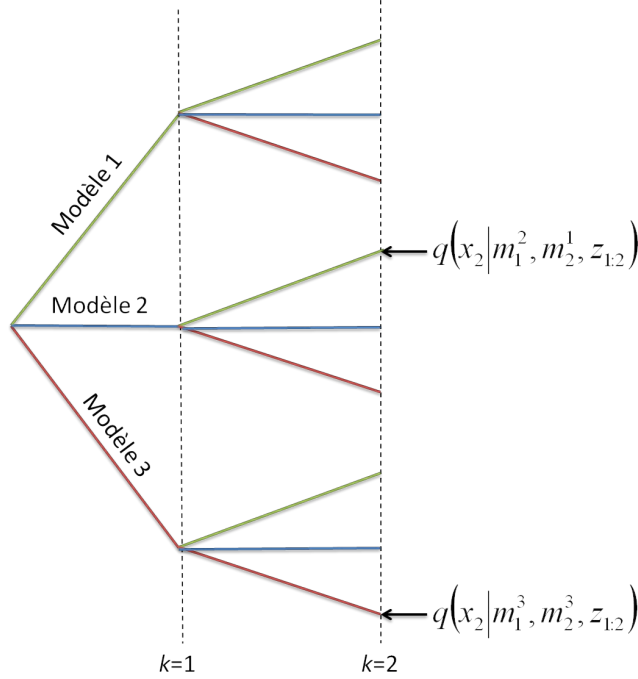


FIGURE 2.4.: Fusion des branches : on approxime par une densité plus simple q l'ensemble des densités dont la sous-séquence de modes $m_{k-n+1:k}$ est identique (ici $n = 2$). On obtient ainsi un arbre plus petit que celui de Fig. 2.2.

Décimation ou Fusion ?

Lorsque les modèles sont très proches des comportements véritables, ce sont les techniques de décimation qui calculent les meilleures estimées [Pitre et al., 2005]. En revanche, si la modélisation s'éloigne de la vérité ou n'est pas complètement fidèle, les techniques de fusion se révèlent être les plus performantes. Dans le domaine du suivi de mobile ARGOS, c'est cette dernière situation qui s'impose du fait de la grande généricité des modèles impliqués. Nous avons en outre fait le choix d'un filtre IMM pour son faible coût en calcul.

2.2.4. Construction du filtre IMM

Le filtre IMM approxime récursivement les probabilités filtrées des modes $\{\mu_{k|k}^j = P(m_k^j | z_{1:k})\}_{j \in \mathcal{M}}$ et les densités filtrées de l'état conditionné sur les modes $\{p(x_k | m_k^j, z_{1:k}) \approx \mathcal{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)\}_{j \in \mathcal{M}}$

CHAPITRE 2. OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION

de sorte que $p(x_k|z_{1:k}) \approx \sum_{j \in \mathcal{M}} \mu_{k|k}^j \mathcal{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)$. Nous allons donc construire les équations de l'IMM en traduisant le cycle de filtrage exact pour des densités gaussiennes et en réduisant le nombre d'hypothèses de la mixture à l'étape d'interaction par l'intermédiaire d'une approximation par moments.

En premier lieu, l'IMM conserve les équations (2.18), (2.20) de calcul des probabilités prédites $\{P(m_{k+1}^i|z_{1:k}) = \mu_{k+1|k}^i\}_{i \in \mathcal{M}}$ et des probabilités d'interaction $\{\mu_{k|k}^{j|i} = P(m_k^j|m_{k+1}^i, z_{1:k})\}_{(i,j) \in \mathcal{M}^2}$ qui s'expriment de manière plus concise :

$$\mu_{k+1|k}^i = \sum_{j \in \mathcal{M}} \pi_{ji} \mu_{k|k}^j \quad \text{et} \quad \mu_{k|k}^{j|i} = \frac{\pi_{ji} \mu_{k|k}^j}{\mu_{k+1|k}^i}. \quad (2.27)$$

Pour l'étape suivante d'interaction (2.19), l'IMM associe à chaque mixture $p(x_k|m_{k+1}^i, z_{1:k}) \approx \sum_{j \in \mathcal{M}} \mu_{k|k}^{j|i} \mathcal{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)$ une unique gaussienne $\mathcal{N}(x_k; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i)$ avec l'approximation par moments :

$$\bar{x}_{k|k}^i = \sum_{j \in \mathcal{M}} \mu_{k|k}^{j|i} \hat{x}_{k|k}^j \quad (2.28)$$

$$\bar{P}_{k|k}^i = \sum_{j \in \mathcal{M}} \mu_{k|k}^{j|i} [P_{k|k}^j + (\hat{x}_{k|k}^j - \bar{x}_{k|k}^i)(\hat{x}_{k|k}^j - \bar{x}_{k|k}^i)^T]. \quad (2.29)$$

C'est cette opération qui maintient constant le nombre d'hypothèses. La spécificité et l'intérêt de l'IMM ne réside cependant pas dans l'approximation ci-dessus, liée au caractère gaussien des densités, mais dans le moment où elle intervient. Son application préalable au filtrage permet de ne maintenir que M termes dans la mixture et donc de n'impliquer que M filtres.

Sur chaque densité résultante $p(x_k|m_{k+1}^i, z_{1:k}) \approx \mathcal{N}(x_k; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i)$, on applique ensuite une passe de filtrage de Kalman (2.21), (2.22) correspondant au mode i afin de déterminer $p(x_{k+1}|m_{k+1}^i, z_{1:k+1}) \approx \mathcal{N}(x_{k+1}; \hat{x}_{k+1|k+1}^i, P_{k+1|k+1}^i)$. Les probabilités des modes $\{\mu_{k+1|k+1}^i = P(m_{k+1}^i|z_{1:k+1})\}_{i \in \mathcal{M}}$ sont mises à jour selon (2.23), *i.e.* :

$$\mu_{k+1|k+1}^i = \frac{\mu_{k+1|k}^i p(z_{k+1}|m_{k+1}^i, z_{1:k})}{\sum_{l \in \mathcal{M}} \mu_{k+1|k}^l p(z_{k+1}|m_{k+1}^l, z_{1:k})}. \quad (2.30)$$

La vraisemblance $L_{k+1}^i = p(z_{k+1}|m_{k+1}^i, z_{1:k})$ du modèle i par rapport à la mesure z_{k+1} est égale à $\mathcal{N}(\tilde{z}_{k+1}^i; 0, S_{k+1}^i)$ où $\tilde{z}_{k+1}^i$ est l'innovation du filtre de Kalman associé au mode i et S_{k+1}^i sa covariance.

Finalement, la moyenne a posteriori de l'état continu $\hat{x}_{k+1|k+1}$ et sa covariance $P_{k+1|k+1}$ associée sont données avec une approximation par moments de $p(x_{k+1}|z_{1:k+1}) \approx \sum_{i \in \mathcal{M}} \mu_{k+1|k+1}^i \mathcal{N}(x_{k+1}; \hat{x}_{k+1|k+1}^i, P_{k+1|k+1}^i)$ telle que :

$$\hat{x}_{k+1|k+1} = \sum_{i \in \mathcal{M}} \mu_{k+1|k+1}^i \hat{x}_{k+1|k+1}^i \quad (2.31)$$

$$P_{k+1|k+1} = \sum_{i \in \mathcal{M}} \mu_{k+1|k+1}^i \left[P_{k+1|k+1}^i + (\hat{x}_{k+1|k+1}^i - \hat{x}_{k+1|k+1})(\hat{x}_{k+1|k+1}^i - \hat{x}_{k+1|k+1})^T \right]. \quad (2.32)$$

Les équations du filtre IMM sont résumées dans l'algorithme 1. Les densités et les probabilités manipulées par l'IMM sont rassemblées dans Tab. 2.1. Elles nous seront utiles dans toute la suite.

2.2. FILTRAGE MULTI-MODÈLE

Algorithme 1 FILTRAGE IMM

```

1: Probabilités de transition (données) :  $\pi_{ji} = P(m_{k+1}^i | m_k^j)$ 
2: Probabilités initiales des modes (données) :  $\mu_{0|0}^i$ 
3: Moyenne et covariance initiales (données) :  $\hat{x}_{0|0}^i$  et  $P_{0|0}^i$ 
4: SI  $k \geq 0$  ALORS
5:   POUR  $i \in \mathcal{M}$  {Interaction}
6:     Probabilité prédite du mode :  $\mu_{k+1|k}^i = \sum_{j \in \mathcal{M}} \pi_{ji} \mu_{k|k}^j$ 
7:     Probabilité d'interaction :  $\mu_{k|k}^{j|i} = \pi_{ji} \mu_{k|k}^j / \mu_{k+1|k}^i$ 
8:     Moyenne d'interaction  $\bar{x}_{k|k}^i = \sum_{j \in \mathcal{M}} \mu_{k|k}^{j|i} \hat{x}_{k|k}^j$ 
9:     Covariance d'interaction  $\bar{P}_{k|k}^i = \sum_{j \in \mathcal{M}} \mu_{k|k}^{j|i} \left[ P_{k|k}^j + (\hat{x}_{k|k}^j - \bar{x}_{k|k}^i)(\hat{x}_{k|k}^j - \bar{x}_{k|k}^i)^T \right]$ 
10:   FIN POUR
11:   POUR  $i \in \mathcal{M}$  {Filtrage}
12:      $(\bar{x}_{k|k}^i, \bar{P}_{k|k}^i) \xrightarrow{\text{Prédiction}} (\hat{x}_{k+1|k}^i, P_{k+1|k}^i)$ 
13:      $(\hat{x}_{k+1|k}^i, P_{k+1|k}^i) \xrightarrow{\text{Correction}} (\hat{x}_{k+1|k+1}^i, P_{k+1|k+1}^i, L_{k+1}^i)$  avec  $L_{k+1}^i$  la vraisemblance du mode
14:   FIN POUR
15:   POUR  $i \in \mathcal{M}$  {Mise à jour de la probabilité des modes}
16:     Probabilité filtrée du mode :  $\mu_{k+1|k+1}^i = \frac{\mu_{k+1|k}^i L_{k+1}^i}{\sum_{l \in \mathcal{M}} \mu_{k+1|k}^l L_{k+1}^l}$ 
17:   FIN POUR
18:   Moyenne globale  $\hat{x}_{k+1|k+1} = \sum_{i \in \mathcal{M}} \mu_{k+1|k+1}^i \hat{x}_{k+1|k+1}^i$ 
19:   Covariance globale
20:    $P_{k+1|k+1} = \sum_{i \in \mathcal{M}} \mu_{k+1|k+1}^i \left[ P_{k+1|k+1}^i + (\hat{x}_{k+1|k+1}^i - \hat{x}_{k+1|k+1})(\hat{x}_{k+1|k+1}^i - \hat{x}_{k+1|k+1})^T \right]$ 
20: FIN SI

```

TABLE 2.1.: Principales quantités utilisées par l'IMM

Quantité	Notation / Approximation	Nom
$p(x_k z_{1:k})$	$\mathcal{N}(x_k; \hat{x}_{k k}, P_{k k})$	densité filtrée globale au temps k
$p(x_k m_k^j, z_{1:k})$	$\mathcal{N}(x_k; \hat{x}_{k k}^j, P_{k k}^j)$	densité filtrée au temps k conditionnée sur m_k^j
$p(x_k m_{k+1}^i, z_{1:k})$	$\mathcal{N}(x_k; \bar{x}_{k k}^i, \bar{P}_{k k}^i)$	densité d'interaction au temps k conditionnée sur m_{k+1}^i
$p(x_{k+1} m_{k+1}^i, z_{1:k})$	$\mathcal{N}(x_{k+1}; \hat{x}_{k+1 k}^i, P_{k+1 k}^i)$	densité prédite au temps $k+1$ conditionnée sur m_{k+1}^i
$P(m_k^j z_{1:k})$	$\mu_{k k}^j$	probabilité filtrée au temps k du mode j
$P(m_{k+1}^i m_k^j)$	π_{ji}	probabilité de transition entre m_k^j et m_{k+1}^i
$P(m_k^j m_{k+1}^i, z_{1:k})$	$\mu_{k k}^{j i}$	probabilité d'interaction au temps k entre m_{k+1}^i et m_k^j
$P(m_{k+1}^i z_{1:k})$	$\mu_{k+1 k}^i$	probabilité prédite au temps $k+1$ du mode i

Comparaison avec un GPB_n

La phase d'interaction n'est pas explicitement réalisée dans le GPB_n [Li and Jilkov, 2005] et elle n'a été mise en évidence que plus tard dans [Blom and Bar-Shalom, 1988] en vue de la construction de l'IMM. Le GPB_n ne fait donc pas d'approximations à l'issue de l'interaction (2.19) de sorte qu'à ce stade la densité

$$p(x_k | m_{k+1}, z_{1:k}) = \sum_{m_{k-n+1:k} \in \mathcal{M}^n} p(x_k | m_{k-n+1:k+1}, z_{1:k}) P(m_{k-n+1:k+1} | z_{1:k}) \quad (2.33)$$

est une mixture de M^n gaussiennes. Après le filtrage, la densité est aussi une mixture de M^n gaussiennes. Ce chiffre est ramené à M^{n-1} en appliquant une approximation par moments sur chaque ensemble de termes dont la séquence de modes $m_{k-n+2:k+1}$ entre les instants $k-n+2$ et $k+1$ est identique. Notons qu'il est tout à fait loisible de construire un IMM à l'image du GPB_n , c'est-à-dire avec une profondeur d'historique plus importante lors de l'interaction [Blom and Bar-Shalom, 1988] mais au prix d'un plus grand nombre de filtres.

Extension au cas non linéaire et non gaussien

Initialement développé dans le cas linéaire, le filtre IMM s'étend de manière transparente dans un contexte non linéaire en s'appuyant sur les filtres de Kalman étendus et "unscented" [Cork and Walker, 2007, Toledo-Moreo et al., 2006]. Il s'applique aussi au cadre non gaussien avec des filtres particuliers en lieu et place des filtres de Kalman, et en adaptant l'interaction pour éviter la croissance géométrique de la taille du nuage de particules [McGinnity and Irwin, 2000, Doucet et al., 2001, Boers and Driessen, 2003].

2.3. Lissage multi-modèle

Là où le filtrage assimile les observations au fil de l'eau, le lissage intègre des mesures futures à l'instant d'estimation considéré. Dans le domaine du suivi de mobile, calculer les positions à l'aide d'observations postérieures réduit les erreurs associées mais au prix d'un délai de mise à disposition des résultats.

La croissance géométrique au cours du temps du nombre de séquences de modes possibles touche aussi la forme de la densité lissée de l'état continu, que ce soit pour du lissage à intervalle fixe (*i.e.* $p(x_k | z_{1:T})$ avec $T \geq k \geq 1$ la longueur donnée de l'intervalle), à retard fixe (*i.e.* $p(x_k | z_{1:k+n})$, avec $n \geq 1$ la longueur donnée du retard) ou à point fixe (*i.e.* $p(x_j | z_{1:k})$, avec j donné et $k \geq j$). On s'intéresse en premier lieu au lissage sur intervalle fixe afin de retraiter des trajectoires de plateformes sur l'ensemble de leur durée de vie. La densité lissée de l'état continu au temps k est ainsi une mixture de densités dont chacune est conditionnée sur les mesures jusqu'au temps T et sur les séquences de modes jusqu'au temps k :

$$p(x_k | z_{1:T}) = \sum_{j_{0:k} \in \mathcal{M}^{k+1}} p(x_k | m_{0:k} = j_{0:k}, z_{1:T}) P(m_{0:k} = j_{0:k} | z_{1:T}). \quad (2.34)$$

Nous établissons dans cette partie deux variantes des équations du lissage exact avant de présenter les approches sous-optimales qui en découlent. Notre objectif est, d'une part, de mettre en pers-

pective les différentes méthodes existantes dans la littérature et, d'autre part, de mettre en évidence les approximations auxquelles elles ont recours.

2.3.1. Hypothèses de base et stratégies de lissage

Considérons un modèle d'état de la forme (2.13) mais impliquant un unique mode (l'index $j = 1$ est donc momentanément omis). En raison des hypothèses associées (2.15)–(2.16), les propriétés markoviennes avant et arrière en temps suivantes sont valables [Särkkä and Hartikainen, 2010] :

$$p(x_{k+1}|x_k, z_{1:T}) = p(x_{k+1}|x_k, z_{k+1:T}); \quad (2.35)$$

$$p(x_k|x_{k+1}, z_{1:T}) = p(x_k|x_{k+1}, z_{1:k}). \quad (2.36)$$

Pour des systèmes à sauts markoviens, (2.36) s'applique sur le vecteur d'état complet ξ_k qui est indépendant de $z_{k+1:T}$ une fois conditionné sur ξ_{k+1} . Ainsi, $p(\xi_k|\xi_{k+1}, z_{1:T}) = p(x_k, m_k|x_{k+1}, m_{k+1}, z_{1:T})$ est égal à $p(\xi_k|\xi_{k+1}, z_{1:k}) = p(x_k, m_k|x_{k+1}, m_{k+1}, z_{1:k})$. Par marginalisation sur m_k , il vient

$$p(x_k|x_{k+1}, m_{k+1}, z_{1:T}) = p(x_k|x_{k+1}, m_{k+1}, z_{1:k}) \quad (2.37)$$

qui est conditionné seulement sur le mode actif pendant la période d'échantillonnage se terminant au temps t_{k+1} inclus. Il est aussi possible d'établir que

$$p(x_{k+1}|x_k, m_k, z_{1:T}) = p(x_{k+1}|x_k, m_k, z_{k+1:T}) \quad (2.38)$$

Comme l'état discret m suit aussi une chaîne de Markov, on a :

$$P(m_{k+1}|m_k, z_{1:T}) = P(m_{k+1}|m_k, z_{k+1:T}); \quad (2.39)$$

$$P(m_k|m_{k+1}, z_{1:T}) = P(m_k|m_{k+1}, z_{1:k}). \quad (2.40)$$

Dans un contexte mono-modèle, deux stratégies sont envisageables concernant le lissage à intervalle fixe [Gustafsson, 2000]. La densité lissée $p(x_k|z_{1:T})$ peut être obtenue avec les formules de lissage à double filtre : la densité a posteriori $p(x_k|z_{1:k})$ calculée par un filtrage exécuté de l'instant initial jusqu'à l'instant courant k est fusionnée avec la densité prédite donnée par un filtrage arrière de l'instant final T jusqu'à k et initialisé sans information a priori. D'un autre côté, un filtrage avant peut être mené afin d'établir les densités a posteriori jusqu'au temps T , préalablement à l'application d'une récursion arrière de type Rauch-Tung-Striebel qui construit les densités lissées. La question se posant alors est comment étendre ces deux principes aux systèmes à sauts Markoviens. Les paragraphes suivants établissent les solutions exactes possibles sans faire d'hypothèses sur la nature des densités manipulées.

2.3.2. Solution exacte par récursion arrière

Estimation récursive - ER1

On présume que les densités $p(x_T|m_T, z_{1:T})$ et les probabilités $P(m_T|z_{1:T})$ jusqu'au temps T sont connues par application du filtrage exact (cf. paragraphe 2.2.2). On souhaite construire récursivement $p(x_k|m_k^j, z_{1:T})$ et $P(m_k^j|z_{1:T})$, $\forall j \in \mathcal{M}$. Partant de $\{p(x_{k+1}|m_{k+1}^i, z_{1:T})\}_{i \in \mathcal{M}}$ et $\{P(m_{k+1}^i|z_{1:T})\}_{i \in \mathcal{M}}$, les étapes de la récursion sont résumées ci-dessous :

1. $\forall i \in \mathcal{M}, p(x_{k+1}|m_{k+1}^i, z_{1:T}) \xrightarrow{\text{Lissage}} p(x_k|m_{k+1}^i, z_{1:T})$
2. $\forall j \in \mathcal{M}, \{P(m_{k+1}^j|z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Lissage}} P(m_k^j|z_{1:T})$
3. $\forall j \in \mathcal{M}, \{p(x_k|m_{k+1}^j, z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k|m_k^j, z_{1:T})$

En vertu de (2.37), l'étape 1 de lissage de l'état continu s'explicite selon :

$$\begin{aligned}
 \forall i \in \mathcal{M}, p(x_k|m_{k+1}^i, z_{1:T}) &= \int_{x_{k+1}} p(x_k|x_{k+1}, m_{k+1}^i, z_{1:k}, \cancel{z_{k+1:T}}) p(x_{k+1}|m_{k+1}^i, z_{1:T}) dx_{k+1} \\
 &= \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m_{k+1}^i, \cancel{z_{1:k}}) p(x_k|m_{k+1}^i, z_{1:k})}{p(x_{k+1}|m_{k+1}^i, z_{1:k})} p(x_{k+1}|m_{k+1}^i, z_{1:T}) dx_{k+1} \\
 &= p(x_k|m_{k+1}^i, z_{1:k}) \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m_{k+1}^i)}{p(x_{k+1}|m_{k+1}^i, z_{1:k})} p(x_{k+1}|m_{k+1}^i, z_{1:T}) dx_{k+1}
 \end{aligned} \tag{2.41}$$

où apparaissent la dynamique $p(x_{k+1}|x_k, m_{k+1}^i)$ et la densité $p(x_k|m_{k+1}^i, z_{1:k})$ calculée par le filtre avant. Le lissage de l'état discret de la seconde étape s'obtient d'après (2.40) avec :

$$\forall j \in \mathcal{M}, P(m_k^j|z_{1:T}) = \sum_{i \in \mathcal{M}} P(m_k^j|m_{k+1}^i, z_{1:k}, \cancel{z_{k+1:T}}) P(m_{k+1}^i|z_{1:T}) \tag{2.42}$$

où apparaissent les probabilités d'interaction $P(m_k^j|m_{k+1}^i, z_{1:k})$ données par le filtre avant. Enfin la récursion se termine avec une étape d'interaction :

$$\forall j \in \mathcal{M}, p(x_k|m_k^j, z_{1:T}) = \sum_{i \in \mathcal{M}} p(x_k|m_k^j, m_{k+1}^i, z_{1:T}) P(m_{k+1}^i|m_k^j, z_{1:T}) \tag{2.43}$$

avec

$$P(m_{k+1}^i|m_k^j, z_{1:T}) = \frac{P(m_k^j|m_{k+1}^i, z_{1:k}) P(m_{k+1}^i|z_{1:T})}{P(m_k^j|z_{1:T})} \tag{2.44}$$

la probabilité d'interaction du lissage. La densité lissée de l'état $p(x_k|m_k^j, m_{k+1}^i, z_{1:T})$ conditionnée sur les modes de deux instants successifs est [Helmick et al., 1995] :

$$\forall (i, j) \in \mathcal{M}^2, p(x_k|m_k^j, m_{k+1}^i, z_{1:T}) \propto p(z_{k+1:T}|x_k, \cancel{m_k^j}, m_{k+1}^i, \cancel{z_{1:k}}) p(x_k|m_k^j, \cancel{m_{k+1}^i}, z_{1:k}). \tag{2.45}$$

Cette dernière équation exploite l'indépendance de $z_{k+1:T}$ avec m_k et $z_{1:k}$ si x_k et m_{k+1} sont données¹. Mais étant donné que $p(x_k|m_{k+1}^i, z_{1:T}) \propto p(z_{k+1:T}|x_k, m_{k+1}^i, \cancel{z_{1:k}}) p(x_k|m_{k+1}^i, z_{1:k})$ on peut aussi écrire

$$\forall (i, j) \in \mathcal{M}^2, p(x_k|m_k^j, m_{k+1}^i, z_{1:T}) \propto \frac{p(x_k|m_{k+1}^i, z_{1:T})}{p(x_k|m_{k+1}^i, z_{1:k})} p(x_k|m_k^j, z_{1:k}). \tag{2.46}$$

Estimation récursive - ER2

Il est aussi possible de faire :

1. cf. le diagramme de dépendance de Fig. 2.1.

1. $\forall (i, j) \in \mathcal{M}^2, p(x_{k+1}|m_{k+1}^i, z_{1:T}) \xrightarrow{\text{Lissage}} p(x_k|m_{k+1}^i, m_k^j, z_{1:T})$
2. $\forall j \in \mathcal{M}, \{P(m_{k+1}^j|z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Lissage}} P(m_k^j|z_{1:T})$
3. $\forall j \in \mathcal{M}, \{p(x_k|m_{k+1}^i, m_k^j, z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k|m_k^j, z_{1:T})$

L'étape 2 est identique à celle de ER1 et l'étape 3 reprend une partie de celle de ER1. Le lissage de l'étape 1 s'obtient par contre avec

$$\begin{aligned}
 & \forall (i, j) \in \mathcal{M}^2, p(x_k|m_{k+1}^i, m_k^j, z_{1:T}) \\
 &= \int_{x_{k+1}} p(x_k|x_{k+1}, m_{k+1}^i, m_k^j, z_{1:k}, z_{k+1:T}) p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:T}) dx_{k+1} \\
 &= \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m_{k+1}^i, m_k^j, z_{1:k}) p(x_k|m_k^j, m_{k+1}^i, z_{1:k})}{p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})} p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:T}) dx_{k+1} \\
 &= p(x_k|m_k^j, z_{1:k}) \int_{x_{k+1}} \frac{p(x_{k+1}|x_k, m_{k+1}^i)}{p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})} p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:T}) dx_{k+1}. \quad (2.47)
 \end{aligned}$$

De plus, sur la base de $p(x_{k+1}|m_k^j, m_{k+1}^i, z_{1:T}) \propto p(z_{k+1:T}|x_{k+1}, m_k^j, m_{k+1}^i, z_{1:k}) p(x_{k+1}|m_k^j, m_{k+1}^i, z_{1:k})$ et de $p(x_{k+1}|m_{k+1}^i, z_{1:T}) \propto p(z_{k+1:T}|x_{k+1}, m_{k+1}^i, z_{1:k}) p(x_{k+1}|m_{k+1}^i, z_{1:k})$, la densité de l'état conditionné sur deux modes se décompose selon

$$\forall (i, j) \in \mathcal{M}^2, p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:T}) \propto \frac{p(x_{k+1}|m_{k+1}^i, z_{1:T})}{p(x_{k+1}|m_{k+1}^i, z_{1:k})} p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k}). \quad (2.48)$$

Contrairement à l'estimation réursive 1 qui nécessite seulement M lisseurs, chacun conditionné sur un mode, le lissage de l'état continu en implique un total de M^2 , chacun conditionné sur deux modes. Cette solution de récursion est aussi moins naturelle car la densité $p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})$ n'a pas été explicitée dans le cycle de filtrage. C'est en réalité une composante de la densité prédite $p(x_{k+1}|m_{k+1}^i, z_{1:k}) = \sum_{j \in \mathcal{M}} p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k}) P(m_k^j|m_{k+1}^i, z_{1:k})$. On peut remarquer que la densité $p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})$ est une mixture qui s'obtient en regroupant les termes associés aux séquences de modes se terminant par m_{k+1}^i et m_k^j dans l'expression globale (2.17).

2.3.3. Solution exacte par double filtrage

Construction d'un filtre multi-modèle à rebours

On suppose que la dynamique du système à sauts markoviens est inversible et peut s'exprimer par les lois à rebours $p(x_k|x_{k+1}, m_{k+1})$ sur l'état continu et $P(m_k|m_{k+1})$ sur l'état discret, vérifiant $p(x_k, m_k|x_{k+1}, m_{k+1}) \triangleq p(x_k|x_{k+1}, m_{k+1})P(m_k|m_{k+1})$. La génération de l'état et de la mesure d'un système à sauts markoviens évoluant à rebours est illustrée dans le graphique de dépendance 2.5 (à comparer avec Fig. 2.1). Connaissant $P(m_{k+1}^i|z_{k+1:T})$ et $p(x_{k+1}|m_{k+1}^i, z_{k+1:T})$, $\forall i \in \mathcal{M}$, un cycle de filtrage multi-modèle à rebours depuis l'instant terminal T est mené de la façon suivante :

1. $\forall j \in \mathcal{M}, \{P(m_{k+1}^j|z_{k+1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Prédiction arrière}} P(m_k^j|z_{k+1:T})$
2. $\forall i \in \mathcal{M}, p(x_{k+1}|m_{k+1}^i, z_{k+1:T}) \xrightarrow{\text{Prédiction arrière}} p(x_k|m_{k+1}^i, z_{k+1:T})$

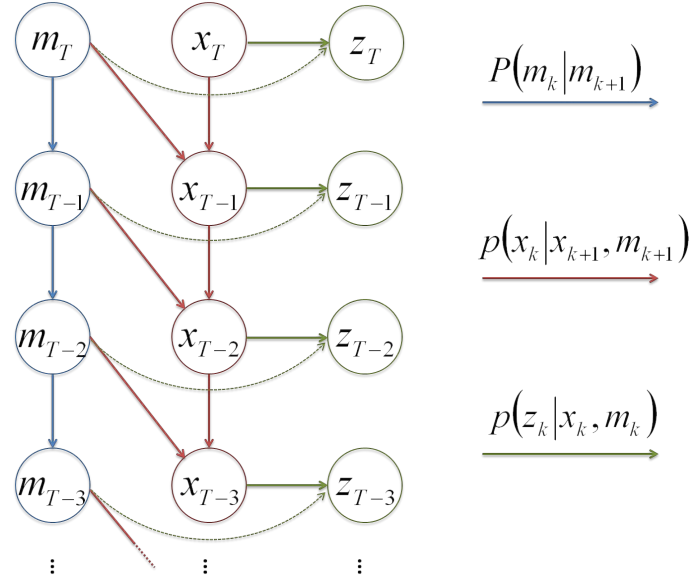


FIGURE 2.5.: Graphique de dépendance d'un système à sauts markoviens se propageant en temps inverse. Les flèches bleues sont associées aux probabilités de transition à rebours $P(m_k|m_{k+1})$, les rouges aux modèles de transition à rebours $p(x_k|x_{k+1}, m_{k+1})$ et les vertes aux modèles d'observation $p(z_k|x_k, m_k)$.

3. $\forall j \in \mathcal{M}, \{p(x_k|m_{k+1}^i, z_{k+1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k|m_k^j, z_{k+1:T})$
4. $\forall j \in \mathcal{M}, p(x_k|m_k^j, z_{k+1:T}) \xrightarrow{\text{Correction}} p(x_k|m_k^j, z_{k:T})$
5. $\forall j \in \mathcal{M}, P(m_k^j|z_{k+1:T}) \xrightarrow{\text{Correction}} P(m_k^j|z_{k:T})$

La prédiction arrière s'écrit en fonction des probabilités de transition à rebours :

$$\forall j \in \mathcal{M}, P(m_k^j|z_{k+1:T}) = \sum_{i \in \mathcal{M}} P(m_{k+1}^i|m_k^j, z_{k+1:T}) P(m_{k+1}^i|z_{k+1:T}). \quad (2.49)$$

Si la suite des modes est une chaîne de Markov, alors la suite en temps inverse l'est aussi et les probabilités de transition sont données par [Helmick et al., 1995] :

$$P(m_k^j|m_{k+1}^i) = \frac{P(m_{k+1}^i|m_k^j)P(m_k^j)}{\sum_{l \in \mathcal{M}} P(m_{k+1}^i|m_k^l)P(m_k^l)} \quad (2.50)$$

qui fait intervenir les probabilités de transition $P(m_{k+1}^i|m_k^j)$. La probabilité a priori $P(m_k^j)$ est déduite de $P(m_k^j) = \sum_l P(m_k^j|m_{k-1}^l)P(m_{k-1}^l)$.

La prédiction arrière utilise la dynamique à rebours $p(x_k|x_{k+1}, m_{k+1}^i)$ sur l'état continu :

$$\forall i \in \mathcal{M}, p(x_k|m_{k+1}^i, z_{k+1:T}) = \int_{x_{k+1}} p(x_k|x_{k+1}, m_{k+1}^i, z_{k+1:T}) p(x_{k+1}|m_{k+1}^i, z_{k+1:T}) dx_{k+1}. \quad (2.51)$$

Le calcul de la densité d'interaction s'effectue selon

$$p(x_k|m_k^j, z_{k+1:T}) = \sum_{i \in \mathcal{M}} p(x_k|m_{k+1}^i, m_k^j, z_{k+1:T}) P(m_{k+1}^i|m_k^j, z_{k+1:T}) \quad (2.52)$$

avec la probabilité d'interaction à rebours

$$P(m_{k+1}^i | m_k^j, z_{k+1:T}) = \frac{P(m_k^j | m_{k+1}^i) P(m_{k+1}^i | z_{k+1:T})}{P(m_k^j | z_{k+1:T})}. \quad (2.53)$$

D'autre part,

$$\begin{aligned} p(x_k | m_{k+1}^i, m_k^j, z_{k+1:T}) &= \int_{x_{k+1}} p(x_k | x_{k+1}, m_{k+1}^i, \cancel{m_k^j, z_{k+1:T}}) p(x_{k+1} | m_{k+1}^i, \cancel{m_k^j, z_{k+1:T}}) dx_{k+1} \\ &= p(x_k | m_{k+1}^i, z_{k+1:T}). \end{aligned} \quad (2.54)$$

La simplification du premier membre de l'intégrale exploite notamment le fait que la transition entre x_{k+1} et x_k ne dépend que de m_{k+1} . Pour le second membre, il suffit de remarquer que $p(x_{k+1} | m_{k+1}^i, m_k^j, z_{k+1:T}) = P(m_k^j | m_{k+1}^i, \cancel{x_{k+1}, z_{k+1:T}}) p(x_{k+1} | m_{k+1}^i, z_{k+1:T}) / P(m_k^j | m_{k+1}^i, \cancel{z_{k+1:T}})$. Pour finir, les étapes 4 et 5 sont des étapes de correction classiques.

Par rapport au filtrage avant, l'interaction sur l'état continu s'intercale entre la prédiction à rebours et l'assimilation de la mesure. La prédiction arrière relie en effet l'état continu dans l'intervalle de temps $(t_k, t_{k+1}]$ où le modèle m_{k+1} est actif par définition. La mesure z_k , assimilée pendant le cycle, est en revanche reliée à x_k par la densité $p(z_k | m_k, x_k)$, c'est à dire par le modèle m_k actif dans $(t_{k-1}, t_k]$ par définition. La correction requiert donc la densité de l'état continu conditionnée sur le modèle m_k et obtenue par une phase préalable d'interaction.

Lisseur à deux filtres

On suppose que les densités issues d'un filtre multi-modèle avant et d'un filtre arrière sont disponibles. On fait aussi l'hypothèse que les deux filtres sont initialisés de manière totalement indépendante. Les probabilités lissées des modes sont obtenues avec (2.42) de la section précédente. Les densités lissées $p(x_k | m_k^j, z_{1:T})$ se décomposent en appliquant le théorème de Bayes

$$p(x_k | m_k^j, z_{1:T}) \propto p(z_{k+1:T} | m_k^j, x_k, \cancel{z_{1:k}}) p(x_k | m_k^j, z_{1:k}). \quad (2.55)$$

La densité $p(x_k | m_k^j, z_{1:k})$ a bien été calculée par le filtre avant. Pour appliquer la formule de lissage (2.55), il faudrait en toute rigueur construire un filtre arrière amorcé avec $p(z_T | x_T, m_T)$ et calculant récursivement la vraisemblance $p(z_{k+1:T} | m_k^j, x_k), \forall k = T-1 \dots 0$. Cependant si $p(z_{k+1:T} | m_k^j, x_k)$, prise comme une fonction de x_k , est une densité non dégénérée ($\int_{x_k} p(z_{k+1:T} | m_k^j, x_k) dx_k < +\infty$), il est possible d'utiliser directement la densité $p(x_k | m_k^j, z_{k+1:T})$ issue du filtrage multi-modèle à rebours, initialisé sans information a priori [Kitagawa, 1994].

Une autre alternative pour construire $p(x_k | m_k^j, z_{1:T})$ est d'appliquer la formule des probabilités totales :

$$p(x_k | m_k^j, z_{1:T}) = \sum_{i \in \mathcal{M}} p(x_k | m_{k+1}^i, m_k^j, z_{1:T}) P(m_{k+1}^i | m_k^j, z_{1:T}) \quad (2.56)$$

avec les probabilités d'interaction $P(m_{k+1}^i | m_k^j, z_{1:T})$ établies en (2.44) et

$$p(x_k | m_{k+1}^i, m_k^j, z_{1:T}) \propto p(z_{k+1:T} | x_k, m_{k+1}^i, \cancel{m_k^j, z_{1:k}}) p(x_k | \cancel{m_{k+1}^i}, m_k^j, z_{1:k}). \quad (2.57)$$

Chaque densité lissée $p(x_k | m_{k+1}^i, m_k^j, z_{1:T})$ conditionnée sur deux modes est alors issue de la fusion de la densité filtrée avant et de la densité prédite arrière. Cela nécessite l'intervention de M^2

CHAPITRE 2. OUTILS FONDAMENTAUX DE FILTRAGE ET DE LISSAGE POUR LA LOCALISATION

processus de fusion, conditionnés sur m_k et m_{k+1} . Remarquons enfin que l'équation (2.56) est identique à la phase d'interaction (2.43) des récursions arrières ER1 et ER2. L'équation 2.57 est aussi utilisée dans ER1.

2.3.4. Panorama des méthodes de résolution sous-optimales

Lisseur à double filtre IMM

La méthode [Helmick et al., 1995] s'inscrit dans l'approche avec double filtrage dans le sens où elle fusionne les moyennes et les covariances produites par un filtre IMM avant classique et un filtre IMM arrière qui se dérive des équations de filtrage multi-modèle à rebours ci-dessus. La fusion utilise (2.57) plutôt que (2.55), qui est moins coûteuse mais affiche des performances légèrement moins bonnes d'après les auteurs. Autre particularité lors du développement de l'IMM arrière, $p(x_k | m_{k+1}^i, m_k^j, z_{k+1:T})$ est présentée comme une approximation de $p(x_k | m_{k+1}^i, z_{k+1:T})$ alors qu'au vu des développements précédents, il s'agit bien d'une égalité. Enfin, le calcul des probabilités d'interaction n'utilise pas la récursion sur l'état discret (2.42) mais fait intervenir l'état continu à travers une approximation sur laquelle nous reviendrons.

Cette méthode nécessite un modèle de dynamique inverse qui peut se révéler difficile à établir avec une équation de transition non linéaire par exemple. En outre, le filtre arrière impose l'utilisation d'une densité a priori plate à l'initialisation afin de garantir la validité des équations de fusion. C'est ce dernier point qui nous semble rédhibitoire pour une application à ARGOS car il faudrait lors du filtrage à rebours attendre un passage à quatre messages pour démarrer le filtre arrière avec un ajustement par moindres carrés. Pour du suivi d'animaux, le nombre de messages peut être très faible du fait des conditions d'émission. Ce comportement s'amplifie si la balise est en fin de vie, c'est à dire si sa batterie est faible.

Lisseur GPB₂-RTS

Plus récemment, [Barber, 2005, Mesot and Barber, 2009, Barber, 2012] ont proposé un schéma de lissage exploitant un GPB₂ à M^2 filtres dont les estimées sont recombinaées lors d'une passe arrière impliquant M^2 lisseurs de type Rauch-Tung-Striebel. En comparaison de l'approche à double filtre, cette méthode permet d'utiliser de manière transparente des dynamiques non-inversibles et la passe arrière s'initialise simplement avec la densité filtrée de l'instant final. C'est une solution qui s'inscrit dans ER2 mais qui comporte des différences notables :

- La densité $p(x_{k+1} | m_{k+1}^i, m_k^j, z_{1:T})$ n'est pas construite selon (2.48) mais en pratiquant l'approximation $p(x_{k+1} | m_{k+1}^i, m_k^j, z_{1:T}) \approx p(x_{k+1} | m_{k+1}^i, z_{1:T})$. L'auteur indique qu'elle est valable car l'état continu au temps $k+1$ est peu influencé par m_k . Cette approximation nous semble abusive car cela revient à supposer que la probabilité de certains termes de $p(x_{k+1} | m_{k+1}^i, z_{1:T})$ est négligeable.
- Les probabilités lissées des modes sont calculées en utilisant (2.42) mais avec $p(m_k^j | m_{k+1}^i, z_{1:T}) = \int_{x_{k+1}} p(m_k^j | x_{k+1}, m_{k+1}^i, z_{1:k}) p(x_{k+1} | m_{k+1}^i, z_{1:T}) dx_{k+1}$. Cette dernière intégrale est proportionnelle à $\int_{x_{k+1}} p(x_{k+1} | m_k^j, m_{k+1}^i, z_{1:k}) p(x_{k+1} | m_{k+1}^i, z_{1:T}) dx_{k+1}$. Elle est approximée en évaluant $p(x_{k+1} | m_k^j, m_{k+1}^i, z_{1:k})$ —dont l'estimée gaussienne est fournie par un filtre GPB₂— en

la moyenne de $p(x_{k+1}|m_{k+1}^i, z_{1:T})$. En d'autres mots, $p(x_{k+1}|m_{k+1}^i, z_{1:T})$ est assimilée à un Dirac, ce qui constitue là aussi une forte approximation. L'auteur signale toutefois qu'il est aussi possible d'utiliser des approximations qui tiennent compte de la covariance de $p(x_{k+1}|m_{k+1}^i, z_{1:T})$, notamment en utilisant une transformée en points sigma.

La solution intermédiaire de Koch

ER2 a aussi été exploitée par [Koch, 2000] qui utilise une équation de lissage obtenue par l'injection de (2.48) dans (2.47) :

$$\begin{aligned} & p(x_k|m_{k+1}^i, m_k^j, z_{1:T}) \\ & \propto \int_{x_{k+1}} c_{k+1}^{ij}(x_{k+1}) \frac{p(x_{k+1}|x_k, m_{k+1}^i) p(x_k|m_k^j, z_{1:k})}{p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})} p(x_{k+1}|m_{k+1}^i, z_{1:T}) dx_{k+1} \end{aligned} \quad (2.58)$$

avec

$$c_{k+1}^{ij}(x_{k+1}) = \frac{p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})}{p(x_{k+1}|m_{k+1}^i, z_{1:k})} = \frac{p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})}{\sum_{l \in \mathcal{M}} P(m_k^l|m_{k+1}^i, z_{1:k}) p(x_{k+1}|m_{k+1}^i, m_k^l, z_{1:k})}. \quad (2.59)$$

Pour faciliter le calcul de l'intégrale, il pratique l'approximation $c_{k+1}^{ij}(x_{k+1}) \approx c_{k+1}^{ij}(\hat{x}_{k+1|T}^i)$ de sorte que les équations du lissage RTS et des probabilités lissées (mode et interaction) qui en découlent sont identiques à [Barber, 2005].

Lors de la passe avant, la méthode de Koch applique à la différence de [Barber, 2005] un filtre multi-modèle qui est une solution intermédiaire entre un IMM et un GPB₂. En effet, la réduction du nombre d'hypothèses a lieu immédiatement après la prédiction afin de calculer explicitement les densités $p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:k})$ nécessaires aux lisseurs (rappelons que pour l'IMM c'est avant la prédiction et pour le GPB₂ après la mise à jour par la mesure). Du point de vue de la complexité, on exécute M^2 prédictions et M mises à jour.

Et un lisseur IMM-RTS ?

Curieusement, l'application de ER1 n'a pas été investiguée. Elle présente l'indéniable avantage de n'utiliser que M lisseurs.

Conclusions

La reformulation du problème de localisation dans le contexte des systèmes dynamiques ouvre la voie à des techniques de filtrage qui résolvent d'emblée plusieurs inconvénients de l'algorithme historique. Celles-ci offrent en effet la possibilité d'estimer une position et de caractériser l'erreur associée quel que soit le nombre de messages reçus ainsi que de limiter l'apparition des localisations images.

Nous avons en outre introduit la formulation d'un système à sauts markoviens et les techniques de résolution possibles. C'est un sujet qui a été largement traité dans la littérature avec la mise à jour du filtre l'IMM et de nombreuses variantes de lissage³. Néanmoins, il ressort que l'IMM n'est pas capable de manipuler un banc de modèles avec des vecteurs d'état de constitutions différentes. Ce serait une possibilité intéressante pour faire coopérer une marche aléatoire classique avec une marche aléatoire corrélée dans notre banc de modèles. D'autre part, les méthodes de lissage n'ont pas exploré les équations de la récursion arrière ER1 que nous avons mis en évidence et qui constitue pourtant une piste sérieuse pour réduire la complexité des calculs. La partie suivante s'attaque à ces deux questions.

3. Nous aurions pu aussi citer l'article [Nadarajah et al., 2012] qui propose un lisseur IMM avec une passe à rebours mais dont le développement nous semble comporter des erreurs de calcul. Les auteurs indiquent que le filtrage multi-modèle entre $k - 1$ et k s'appuie par hypothèse sur des filtres associés aux modes m_k actifs au temps k , c'est-à-dire sur les densités de transition $p(x_k|x_{k-1}, m_k)$ comme il est d'usage de le faire (Eqs. (1)-(2) de l'article). Pour mener la passe arrière de lissage (Eq. (11) de l'article), les auteurs ont besoin de $p(x_{k+1}|x_k, m_k)$, la densité de transition entre les instants k et $k + 1$ associée aux modes m_k actifs au temps k . Cette densité n'est jamais définie et est curieusement assimilée à $p(x_{k+1}|x_k, m_{k+1})$ lors de l'établissement de la solution analytique. En outre, ils posent (Eq. (12) de l'article) que $p(x_{k+1}|m_{k+1}^i, m_k^j, z_{1:T}) = p(x_{k+1}|m_{k+1}^i, z_{1:T})$ où m_k^j est « ignoré en raison des propriétés markoviennes » (alors que [Barber, 2005] en reste à une approximation). Si cette égalité était valable, la densité de l'état continu ne serait pas de complexité exponentiellement croissante avec le temps. Il n'y aurait alors pas besoin de filtre ou de lisseur IMM !

NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES



Tortue verte et tortue à dos plat.

Sommaire

3.1. Lissage IMM-RTS	62
3.1.1. Vue d'ensemble	62
3.1.2. Phase 1 : lissage sur l'état continu	63
3.1.3. Phase 2 : interaction des modes avec M^2 combinaisons	65
3.1.4. Phase 2 (bis) : interaction des modes avec M combinaisons	67
3.1.5. Algorithme complet	67
3.1.6. Extension au lissage à retard fixe	71
3.2. Filtrage multi-modèle hétérogène	72
3.2.1. Formulation d'un système à sauts markoviens hétérogène	72
3.2.2. Stratégie d'interaction par composante	73
3.2.3. Algorithme de filtrage complet	76
3.3. Lissage multi-modèle hétérogène	77
3.3.1. Lisseur IMMHO-RTS	77
3.3.2. Lisseur IMMHO à double filtre	84
3.4. Application sur des données simulées	84
3.4.1. Lissage sur intervalle fixe	84
3.4.2. Lissage à retard fixe	88
3.4.3. Filtrage et lissage hétérogènes	89
Conclusions	94

COMME première contribution de ce chapitre, nous proposons une nouvelle méthode de lissage multi-modèle. Elle montre comment les quantités estimées par un filtre IMM jusqu'au temps T permettent de calculer analytiquement une approximation de la densité lissée de l'état aux instants $k < T$ avec seulement M filtres/lisseurs pour un ensemble de M modèles. C'est la traduction, sous l'hypothèse de densités gaussiennes, de la récursion de lissage ER1 que nous avons détaillée dans le précédent chapitre mais qui n'avait pas été exploitée dans la littérature.

Dans un second temps, nous étendons le cadre des systèmes à sauts markoviens à des ensembles de modèles dont les vecteurs d'état sont de constitutions différentes. Il s'agit de déduire une formulation plus générale de l'IMM dans l'optique de faire cohabiter des marches aléatoires simples et corrélées. L'algorithme qui en découle est étendu aux solutions de lissage sur intervalle fixe.

Le chapitre se termine par une évaluation de ces nouveaux outils sur des données simulées.

3.1. Lissage IMM-RTS

3.1.1. Vue d'ensemble

De la même manière que l'IMM propage les approximations des probabilités a posteriori des modes $\{P(m_k^j|z_{1:k})\}_{j \in \mathcal{M}}$ et les approximations gaussiennes $\{\mathcal{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)\}_{j \in \mathcal{M}}$ des densités filtrées conditionnées sur les modes, de sorte que $p(x_k|z_{1:k}) = \sum_{j \in \mathcal{M}} P(m_k^j|z_{1:k}) p(x_k|m_k^j, z_{1:k}) \approx \sum_{j \in \mathcal{M}} P(m_k^j|z_{1:k}) \mathcal{N}(x_k; \hat{x}_{k|k}^j, P_{k|k}^j)$, l'objectif est de calculer les M termes de la mixture de gaussiennes qui approxime la densité lissée $p(x_k|z_{1:T})$ selon

$$\begin{aligned} p(x_k|z_{1:T}) &= \sum_{j \in \mathcal{M}} P(m_k^j|z_{1:T}) p(x_k|m_k^j, z_{1:T}) \\ &\approx \sum_{j \in \mathcal{M}} P(m_k^j|z_{1:T}) \mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j). \end{aligned} \quad (3.1)$$

À partir des statistiques $\{\hat{x}_{k|k}^j, P_{k|k}^j, \mu_{k|k}^j\}_{j \in \mathcal{M}}$ et $\{\bar{x}_{k|k}^i, \bar{P}_{k|k}^i\}_{i \in \mathcal{M}}$ données par le filtre IMM aux instants $k = 0, \dots, T$, et des statistiques $\{\hat{x}_{k+1|k}^i, P_{k+1|k}^i\}_{i \in \mathcal{M}}$ produites aux instants $k+1 = 1, \dots, T$ dans l'IMM, nous déterminons récursivement les densités lissées conditionnées sur les modes $\{p(x_k|m_k^j, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j)\}_{j \in \mathcal{M}}$ et les probabilités lissées des modes $\{\mu_{k|T}^j = P(m_k^j|z_{1:T})\}_{j \in \mathcal{M}}$ pour $k = T-1, \dots, 0$.

Nous allons pour cela transposer dans le cas gaussien les équations de la récursion de lissage arrière ER1 en procédant en deux phases. La première, qui utilise comme point de départ les densités $p(x_{k+1}|m_{k+1}^i, z_{1:T})$, fournit une solution analytique aux équations de lissage sur l'état continu de l'étape 1 p. 52, *i.e.* :

$$\forall i \in \mathcal{M}, p(x_{k+1}|m_{k+1}^i, z_{1:T}) \xrightarrow{\text{Lissage}} p(x_k|m_{k+1}^i, z_{1:T})$$

avec $p(x_{k+1}|m_{k+1}^i, z_{1:T}) \approx \mathcal{N}(x_{k+1}; \hat{x}_{k+1|T}^i, P_{k+1|T}^i)$ et $p(x_k|m_{k+1}^i, z_{1:T}) \approx \mathcal{N}(x_k; \bar{x}_{k|T}^i, \bar{P}_{k|T}^i)$. La deuxième phase explicite l'interaction sur l'état continu de l'étape 3 à partir des densités $p(x_k|m_{k+1}^i, z_{1:T})$ calculées dans la phase de lissage, *i.e.* :

$$\forall j \in \mathcal{M}, \{p(x_k|m_{k+1}^i, z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k|m_k^j, z_{1:T})$$

avec $p(x_k|m_k^j, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j)$. Nous établirons pour cette interaction deux variantes qui se différencient par leur charge en calcul.

L'algorithme de lissage que nous présentons ici est une version améliorée (notamment en termes de complexité) de notre article [Lopez and Danès, 2012]. Tout au long des développements, nous ferons le lien, d'une part, avec les méthodes de lissage équivalentes dans la cas mono-modèle et, d'autre part, avec la solution de lissage à base de double filtre de [Helmick et al., 1995].

3.1.2. Phase 1 : lissage sur l'état continu

Replaçons nous d'abord dans le contexte mono-modèle (2.1)-(2.2) et supposons qu'une variante non linéaire du filtre de Kalman permette le calcul des approximations gaussiennes $\mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k}) \approx p(x_k|z_{1:k})$ et $\mathcal{N}(x_{k+1}; \hat{x}_{k+1|k}, P_{k+1|k}) \approx p(x_{k+1}|z_{1:k})$ des densités filtrées et prédites jusqu'au temps T . L'approximation $\mathcal{N}(x_k; \hat{x}_{k|T}, P_{k|T})$ de la densité lissée $p(x_k|z_{1:T})$ est déduite de ces quantités en appliquant récursivement le lemme ci-dessous.

Lemme 1 (Formules de Rauch-Tung-Striebel) *Soit*

$$C_{k,k+1} = \int_{x_k} (x_k - \hat{x}_{k|k}) (f_{k+1}(x_k) - \hat{x}_{k+1|k})^T \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k}) dx_k. \quad (3.2)$$

Les moments de la densité lissée $p(x_k|z_{1:T})$ se déduisent de

$$G_k = C_{k,k+1} P_{k+1|k}^{-1} \quad (3.3)$$

$$\hat{x}_{k|T} = \hat{x}_{k|k} + G_k (\hat{x}_{k+1|T} - \hat{x}_{k+1|k}) \quad (3.4)$$

$$P_{k|T} = P_{k|k} + G_k (P_{k+1|T} - P_{k+1|k}) G_k^T. \quad (3.5)$$

Preuve Les expressions précédentes de $\hat{x}_{k|T}$ et $P_{k|T}$ ont été démontrées par [Rauch et al., 1965] dans le cas linéaire gaussien comme étant le maximum a posteriori (MAP) de la densité lissée et la covariance de l'erreur d'estimation associée. Une preuve alternative, tirée de [Särkkä and Hartikainen, 2010] et détaillée ci-après, introduit $\hat{x}_{k|T}$ et $P_{k|T}$ comme la moyenne et la covariance a posteriori calculées sur la base d'un modèle d'état stochastique non linéaire. La preuve se déroule en quatre étapes.

1. Écrire la densité conjointe a posteriori de x_k, x_{k+1} conditionnée sur $z_{1:k}$, *i.e.*

$$p(x_k, x_{k+1}|z_{1:k}) \approx \mathcal{N}\left(\begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix}; \begin{pmatrix} \hat{x}_{k|k} \\ \hat{x}_{k+1|k} \end{pmatrix}, \begin{pmatrix} P_{k|k} & C_{k,k+1} \\ C_{k,k+1}^T & P_{k+1|k} \end{pmatrix}\right) \quad (3.6)$$

avec la covariance croisée¹ $C_{k,k+1} = E[(x_k - \hat{x}_{k|k})(x_{k+1} - \hat{x}_{k+1|k})^T | z_{1:k}]$.

2. Dédire de (3.6) la densité $p(x_k|x_{k+1}, z_{1:k})$ en exploitant les formules classiques pour l'obtention des densités conditionnelles de vecteurs conjointement gaussiens. Remarquer que la densité résultante $p(x_k|x_{k+1}, z_{1:k})$ est égale à $p(x_k|x_{k+1}, z_{1:T})$ par (2.36).
3. Par un raisonnement inverse, déduire $p(x_k, x_{k+1}|z_{1:T})$ à partir de $p(x_k|x_{k+1}, z_{1:T}) = p(x_k|x_{k+1}, z_{1:k})$ et $p(x_{k+1}|z_{1:T})$.

1. Dans l'expression de $C_{k,k+1}$ donnée par (3.2), remarquons que l'inter-covariance entre x_k et q_k est 0. Cette propriété découle de (2.16) où le bruit de dynamique est blanc et indépendant de l'état initial.

4. Calculer $p(x_k|z_{1:T})$ par marginalisation de $p(x_k, x_{k+1}|z_{1:T})$ par rapport à x_{k+1} .

□

Les équations de lissage sont complètement linéaires hormis l'évaluation de l'intégrale (3.2) qui peut être approximée en utilisant une transformée en points sigma (ou "unscented") [Julier, 2002, Julier and Uhlmann, 1997]. Autre fait notable, ces équations s'étendent au lissage à retard fixe.

Dans un contexte multi-modèle, nous établissons alors le théorème suivant.

Théorème 1 *Connaissant $p(x_{k+1}|m_{k+1}^i, z_{1:T}) \approx \mathcal{N}(x_{k+1}; \hat{x}_{k+1|T}^i, P_{k+1|T}^i)$ à l'instant $k+1$ et $p(x_k|m_{k+1}^i, z_{1:k}) \approx \mathcal{N}(x_k; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i)$ à l'instant k , la moyenne et la covariance de la densité d'interaction lissée $p(x_k|m_{k+1}^i, z_{1:T}) \approx \mathcal{N}(x_k; \bar{x}_{k|T}^i, \bar{P}_{k|T}^i)$ sont données par les formules de Rauch-Tung-Striebel*

$$G_k^i = C_{k,k+1}^i (P_{k+1|k}^i)^{-1} \quad (3.7)$$

$$\bar{x}_{k|T}^i = \bar{x}_{k|k}^i + G_k^i (\hat{x}_{k+1|T}^i - \hat{x}_{k+1|k}^i) \quad (3.8)$$

$$\bar{P}_{k|T}^i = \bar{P}_{k|k}^i + G_k^i (P_{k+1|T}^i - P_{k+1|k}^i) (G_k^i)^T \quad (3.9)$$

où

$$C_{k,k+1}^i = \int_{x_k} (x_k - \bar{x}_{k|k}^i) (f_{k+1}^i(x_k) - \hat{x}_{k+1|k}^i)^T \mathcal{N}(x_k; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i) dx_k. \quad (3.10)$$

Preuve Les équations (3.7)-(3.10) se démontrent en suivant exactement la preuve du Lemme 1 avec toutes les densités conditionnées sur m_{k+1}^i et en s'appuyant sur la propriété (2.37). □

Ce théorème fournit bien une expression analytique à l'équation de lissage multi-modèle (2.41) avec des densités gaussiennes.

À ce stade, il est déjà possible de calculer la densité globale de l'état continu en la décomposant ainsi : $p(x_k|z_{1:T}) = \sum_{i \in \mathcal{M}} P(m_{k+1}^i|z_{1:T}) p(x_k|m_{k+1}^i, z_{1:T}) \approx \sum_{i \in \mathcal{M}} \mu_{k+1|T}^i \mathcal{N}(x_k; \bar{x}_{k|T}^i, \bar{P}_{k|T}^i)$. La moyenne $\hat{x}_{k|T}$ et la covariance $P_{k|T}$ de cette mixture de gaussiennes sont données par

$$\hat{x}_{k|T} = \sum_{i \in \mathcal{M}} \mu_{k+1|T}^i \bar{x}_{k|T}^i \quad (3.11)$$

$$P_{k|T} = \sum_{i \in \mathcal{M}} \mu_{k+1|T}^i \left[\bar{P}_{k|T}^i + (\bar{x}_{k|T}^i - \hat{x}_{k|T})(\bar{x}_{k|T}^i - \hat{x}_{k|T})^T \right]. \quad (3.12)$$

En cas d'hypothèse gaussienne sur la densité globale de l'état continu, on a $p(x_k|z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|T}, P_{k|T})$.

À l'inverse d'un lisseur mono-modèle, les équations (3.7), (3.8), (3.9) ne concluent pas le cycle de récursion car la densité lissée de x_k est conditionnée sur m_{k+1}^i au lieu de m_k^j . Comme on l'a vu, il faut fermer la récursion avec l'étape d'interaction qui permet de relier les approximations des densités lissées $\mathcal{N}(x_k; \bar{x}_{k|T}^i, \bar{P}_{k|T}^i) \approx p(x_k|m_{k+1}^i, z_{1:T})$ avec $\mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j) \approx p(x_k|m_k^j, z_{1:T})$. Nous envisageons deux variantes de l'interaction, la première étant de complexité quadratique en inversions de matrice et la seconde, un peu moins évidente, mais de complexité linéaire.

3.1.3. Phase 2 : interaction des modes avec M^2 combinaisons

Comme on l'a vu lors de la mise en place des équations exactes de lissage, la densité lissée $p(x_k|m_k^j, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j)$ s'exprime comme une mixture de densités lissées conditionnées sur les modes de deux instants consécutifs, à savoir

$$p(x_k|m_k^j, z_{1:T}) = \sum_{i \in \mathcal{M}} p(x_k|m_k^j, m_{k+1}^i, z_{1:T}) P(m_{k+1}^i|m_k^j, z_{1:T}). \quad (3.13)$$

Les deux théorèmes suivants explicitent son calcul.

Théorème 2 *Les deux premiers moments de $p(x_k|m_k^j, m_{k+1}^i, z_{1:T})$ sont déduits d'un filtrage IMM avant et d'une récursion à rebours de type Rauch-Tung-Striebel selon :*

$$\hat{x}_{k|T}^{ji} = P_{k|T}^{ji} \left[(\bar{P}_{k|T}^i)^{-1} \bar{x}_{k|T}^i - (\bar{P}_{k|k}^i)^{-1} \bar{x}_{k|k}^i + (P_{k|k}^j)^{-1} \hat{x}_{k|k}^j \right], \quad (3.14)$$

avec

$$P_{k|T}^{ji} = \left[(\bar{P}_{k|T}^i)^{-1} - (\bar{P}_{k|k}^i)^{-1} + (P_{k|k}^j)^{-1} \right]^{-1}. \quad (3.15)$$

Preuve Dans le chapitre précédent, on a établi que

$$p(x_k|m_k^j, m_{k+1}^i, z_{1:T}) \propto \frac{p(x_k|m_{k+1}^i, z_{1:T})}{p(x_k|m_{k+1}^i, z_{1:k})} p(x_k|m_k^j, z_{1:k}). \quad (3.16)$$

Toutes les densités impliquées sont approximées par des gaussiennes. De ce fait, le logarithme de $p(x_k|m_k^j, m_{k+1}^i, z_{1:T})$ s'écrit $C - \frac{1}{2}J(x_k)$, avec C une constante et

$$J(x_k) = \|x_k - \bar{x}_{k|T}^i\|_{(\bar{P}_{k|T}^i)^{-1}}^2 - \|x_k - \bar{x}_{k|k}^i\|_{(\bar{P}_{k|k}^i)^{-1}}^2 + \|x_k - \hat{x}_{k|k}^j\|_{(P_{k|k}^j)^{-1}}^2. \quad (3.17)$$

La moyenne $\hat{x}_{k|T}^{ji}$ est le mode de $p(x_k|m_k^j, m_{k+1}^i, z_{1:T})$ et aussi le minimum de $J(x_k)$, ce qui conduit à (3.14)–(3.15). $\square$

Du fait que $p(z_{k+1:T}|x_k, m_{k+1}^i) \propto \frac{p(x_k|m_{k+1}^i, z_{1:T})}{p(x_k|m_{k+1}^i, z_{1:k})}$, l'estimée du maximum de vraisemblance de x_k s'écrit

$$\hat{x}_{k|k+1}^{b,i} = P_{k|k+1}^{b,i} \left[(\bar{P}_{k|T}^i)^{-1} \bar{x}_{k|T}^i - (\bar{P}_{k|k}^i)^{-1} \bar{x}_{k|k}^i \right] \quad (3.18)$$

avec

$$P_{k|k+1}^{b,i} = \left[(\bar{P}_{k|T}^i)^{-1} - (\bar{P}_{k|k}^i)^{-1} \right]^{-1}, \quad (3.19)$$

et se révèle être la prédiction arrière (avec $P_{k|k+1}^{b,i}$ la covariance d'erreur associée) calculée par le filtre IMM à rebours de [Helmick et al., 1995]. Dans (3.14), elle est aussi fusionnée avec l'estimée filtrée $\hat{x}_{k|k}^j$ d'un IMM avant pour former l'estimée lissée $\hat{x}_{k|T}^{ji}$.

En ce qui concerne les probabilités d'interaction lissées $P(m_{k+1}^i|m_k^j, z_{1:T})$, on présente une approximation faisant intervenir l'état continu dans son évaluation.

Théorème 3 Les probabilités d'interaction $\{\bar{\mu}_{k+1|T}^{i|j} = P(m_{k+1}^i | m_k^j, z_{1:T})\}_{(i,j) \in \mathcal{M}^2}$ intervenant dans (3.13) s'écrivent aussi

$$\bar{\mu}_{k+1|T}^{i|j} = \frac{P(m_{k+1}^i | m_k^j) p(z_{k+1:T} | m_k^j, m_{k+1}^i, z_{1:k})}{p(z_{k+1:T} | m_k^j, z_{1:k})} = \frac{\pi_{ji} \Lambda_{ji}}{d_j}, \quad (3.20)$$

où la vraisemblance

$$\Lambda_{ji} = p(z_{k+1:T} | m_k^j, m_{k+1}^i, z_{1:k}) \quad (3.21)$$

peut être approximée par

$$\begin{aligned} \Lambda_{ji} &\approx \mathcal{N}(\Delta_k^{ji}; 0, D_k^{ji}), \\ \Delta_k^{ji} &= \hat{x}_{k|k+1}^{b,i} - \hat{x}_{k|k}^j, \quad D_k^{ji} = P_{k|k+1}^{b,i} + P_{k|k}^j, \end{aligned} \quad (3.22)$$

et

$$d_j = p(z_{k+1:T} | m_k^j, z_{1:k}) = \sum_{i \in \mathcal{M}} \pi_{ji} \Lambda_{ji} \quad (3.23)$$

désigne une constante de normalisation.

Preuve Eq. (3.20) est immédiate. D'autre part, (3.22) a été démontrée dans [Helmick et al., 1995] en approximant dans Λ_{ji} les mesures $z_{k+1:T}$ par $\hat{x}_{k|k+1}^{b,i}$ et $z_{1:k}$ par $\hat{x}_{k|k}^j$. Une démonstration alternative (et nous semblant plus élégante) est d'appliquer le théorème des probabilités totales sur Λ_{ji} de sorte que

$$p(z_{k+1:T} | m_k^j, m_{k+1}^i, z_{1:k}) = \int_{x_k} p(z_{k+1:T} | x_k, \cancel{m_k^j}, m_{k+1}^i, \cancel{z_{1:k}}) p(x_k | m_k^j, \cancel{m_{k+1}^i}, z_{1:k}) dx_k. \quad (3.24)$$

Le produit de deux densités gaussiennes est une gaussienne non normalisée. Comme $p(z_{k+1:T} | x_k, m_{k+1}^i)$ peut être considérée égale à la gaussienne $\mathcal{N}(x_k; \hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i})$ (cf. paragraphe 2.3.3) dont les moments sont explicités plus haut, l'intégrale (3.24) est alors égale² à une constante de normalisation valant $\mathcal{N}(\Delta_k^{ji}; 0, D_k^{ji})$. $\square$

Nous avons choisi de mettre en exergue cette approximation pour les probabilités d'interaction lissées au détriment de (2.44). Elle offre dans la suite de meilleures performances et en particulier une plus grande stabilité numérique de l'algorithme complet quand on fait varier les paramètres des modèles.

La moyenne lissée $\hat{x}_{k|T}^j$ et la covariance lissée $P_{k|T}^j$ (3.13) conditionnées sur les modes sont en fin de compte calculées via une approximation par moments :

$$\hat{x}_{k|T}^j = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{i|j} \hat{x}_{k|T}^{ji} \quad (3.25)$$

$$P_{k|T}^j = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{i|j} \left[P_{k|T}^{ji} + (\hat{x}_{k|T}^{ji} - \hat{x}_{k|T}^j)(\hat{x}_{k|T}^{ji} - \hat{x}_{k|T}^j)^T \right]. \quad (3.26)$$

Ces dernières équations ferment la récursion sur l'état continu.

2. On exploite la propriété $\int_x \mathcal{N}(x; a, P_a) \mathcal{N}(x; b, P_b) dx = \mathcal{N}(a - b; 0, P_a + P_b)$ [Williams and Rasmussen, 2006, Eq. (A.8)].

3.1.4. Phase 2 (bis) : interaction des modes avec M combinaisons

Plutôt que de combiner les M^2 densités filtrées $p(x_k|m_k^j, z_{1:k})$ et $p(x_k|m_{k+1}^j, z_{k+1:T})$, une autre solution pour déterminer $p(x_k|m_k^j, z_{1:T})$ consiste à fusionner les M densités filtrées $p(x_k|m_k^j, z_{1:k})$ et $p(x_k|m_k^j, z_{k+1:T})$. Sa mise en œuvre est expliquée dans les théorèmes suivants.

Théorème 4 *La densité d'interaction arrière $p(x_k|m_k^j, z_{k+1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|k+1}^{b,j}, \bar{P}_{k|k+1}^{b,j})$ est issue de l'étape d'interaction arrière*

$$p(x_k|m_k^j, z_{k+1:T}) = \sum_{i \in \mathcal{M}} p(x_k|m_{k+1}^i, z_{k+1:T}) P(m_{k+1}^i|m_k^j, z_{k+1:T}) \quad (3.27)$$

où la densité prédite arrière $p(x_k|m_{k+1}^i, z_{k+1:T})$ est approximée par $\mathcal{N}(x_k; \hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i})$ et les probabilités d'interaction arrière $\mu_{k+1|k+1}^{b,i,j} = P(m_{k+1}^i|m_k^j, z_{k+1:T})$ sont égales à $\bar{\mu}_{k+1|T}^{i,j} = P(m_{k+1}^i|m_k^j, z_{1:T})$. Les deux premiers moments de $p(x_k|m_k^j, z_{k+1:T})$ sont :

$$\hat{x}_{k|k+1}^{b,j} = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{i,j} \hat{x}_{k|k+1}^{b,i} \quad (3.28)$$

$$\bar{P}_{k|k+1}^{b,j} = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{i,j} [P_{k|k+1}^{b,i} + (\hat{x}_{k|k+1}^{b,i} - \hat{x}_{k|k+1}^{b,j})(\hat{x}_{k|k+1}^{b,i} - \hat{x}_{k|k+1}^{b,j})^T]. \quad (3.29)$$

Preuve L'équation (3.27) et les moments (3.28)-(3.29) ont été introduits dans [Helmick et al., 1995] lors de l'étape d'interaction de l'IMM arrière. L'égalité $P(m_{k+1}^i|m_k^j, z_{1:T}) = P(m_{k+1}^i|m_k^j, z_{k+1:T})$ découle des propriétés markoviennes de la séquence de modes (2.40). $\square$

Pour terminer, la récursion est bouclée avec la combinaison de $p(x_k|m_k^j, z_{1:k})$ et $p(x_k|m_k^j, z_{k+1:T})$.

Théorème 5 *La moyenne et la covariance de la densité lissée $p(x_k|m_k^j, z_{1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|T}^j, P_{k|T}^j)$ sont*

$$\hat{x}_{k|T}^j = P_{k|T}^j \left[(P_{k|k}^j)^{-1} x_{k|k}^j + (\bar{P}_{k|k+1}^{b,j})^{-1} \hat{x}_{k|k+1}^{b,j} \right] \quad (3.30)$$

$$P_{k|T}^j = \left[(P_{k|k}^j)^{-1} + (\bar{P}_{k|k+1}^{b,j})^{-1} \right]^{-1}. \quad (3.31)$$

Preuve On a vu précédemment que la densité $p(x_k|m_k^j, z_{1:T})$ se réécrit

$$p(x_k|m_k^j, z_{1:T}) \propto p(z_{k+1:T}|x_k, m_k^j) p(x_k|m_k^j, z_{1:k}).$$

En suivant un raisonnement proche de la preuve du théorème 2 et en tirant parti du fait que la vraisemblance $p(z_{k+1:T}|x_k, m_k^j)$ conditionnée sur m_k^j est égale à $\mathcal{N}(x_k; \hat{x}_{k|k+1}^{b,j}, \bar{P}_{k|k+1}^{b,j})$, on obtient les équations (3.30) et (3.31). $\square$

3.1.5. Algorithme complet

La probabilité lissée $\mu_{k|T}^j = P(m_k^j|z_{1:T})$ du mode j à l'instant k est donnée avec (2.42) que nous répétons ci-dessous :

$$P(m_k^j|z_{1:T}) = \sum_{i \in \mathcal{M}} P(m_k^j|m_{k+1}^i, z_{1:k}) P(m_{k+1}^i|z_{1:T}).$$

Si l'on s'intéresse à la détection du mode actif, l'estimée $\hat{j}_k$ du mode au temps t_k au sens du maximum a posteriori (ou MAP) s'écrit

$$\hat{j}_{k|T} = \arg \max_{j=1,\dots,M} \mu_{k|T}^j. \quad (3.32)$$

En sortie, la densité globale de l'état continu $p(x_k|z_{1:T}) = \sum_{i \in \mathcal{M}} p(x_k|m_k^i, z_{1:T})P(m_k^i|z_{1:T})$ dans (3.1) peut être approximée par une densité gaussienne $\mathcal{N}(x_k; \hat{x}_{k|T}, P_{k|T})$ de moments

$$\hat{x}_{k|T} = \sum_{j \in \mathcal{M}} \mu_{k|T}^j \hat{x}_{k|T}^j \quad (3.33)$$

$$P_{k|T} = \sum_{j \in \mathcal{M}} \mu_{k|T}^j \left[P_{k|T}^j + (\hat{x}_{k|T}^j - \hat{x}_{k|T})(\hat{x}_{k|T}^j - \hat{x}_{k|T})^T \right]. \quad (3.34)$$

Point important, la covariance $P_{k|k+1}^{b,i}$ peut ne pas être définie durant les premiers cycles de récursion arrière car $(P_{k|k+1}^{b,i})^{-1} = (\bar{P}_{k|T}^i)^{-1} - (\bar{P}_{k|k}^i)^{-1}$ peut ne pas être inversible. Cette situation se produit quand la taille du vecteur de mesure z_k est plus petite que celle du vecteur d'état x_k ou si les mesures assimilées ne contiennent pas suffisamment d'information pour estimer $\hat{x}_{k|k+1}^{b,i}$ et doter $P_{k|k+1}^{b,i}$ de valeurs propres finies. Dans ce cas, la densité $\mathcal{N}(\cdot; 0, D_k^{ji})$ peut être envisagée comme une densité a priori « plate » et les probabilités d'interaction lissées peuvent être approximées par $\bar{\mu}_{k+1|T}^{i,j} \approx P(m_{k+1}^i|m_k^j) = \pi_{ji}, \forall (i, j) \in \mathcal{M}^2$ jusqu'à ce que $(P_{k|k+1}^{b,i})^{-1}$ ne soit plus singulière.

Les étapes de lissage et d'interaction des théorèmes 1 à 5 sont rassemblées dans l'algorithme 2. Pour des raisons de clarté, les équations détaillées de l'étape d'interaction sont présentées séparément dans l'algorithme 3 pour M^2 combinaisons (« Interaction M^2 ») et dans l'algorithme 4 pour M combinaisons (« Interaction M »). Enfin, dans Tab. 3.1, nous avons rassemblé les quantités utilisées par le lisseur IMM-RTS et qui n'avaient pas déjà été employées pour le filtre IMM dans Tab. 2.1.

L'interaction M^2 utilise $O(M^2)$ inversions de matrices alors que l'interaction M en nécessite seulement $O(M)$ pour construire $p(x_k|m_k^j, z_{1:T})$. Néanmoins, l'interaction M impose que $\hat{x}_{k|k+1}^{b,i}$ et $P_{k|k+1}^{b,i}$ soient explicitement définies. Ainsi, lorsqu'on choisit cette variante pour le lisseur, il est nécessaire de mener les premiers cycles de récursion avec l'interaction M^2 jusqu'à ce que $(P_{k|k+1}^{b,i})^{-1}$ soit inversible.

Par rapport aux solutions équivalentes de [Barber, 2005] et [Koch, 2000], l'étape de filtrage est plus économe puisqu'elle s'appuie exclusivement sur les M filtres de l'IMM. Le nombre de lisseurs RTS est égal au nombre de modes. La phase d'interaction reste de complexité linéaire (en inversions de matrices) si on utilise la deuxième variante.

Par rapport à [Helmick et al., 1995] la complexité est la même : les M filtres de l'IMM arrière remplacent les M lisseurs RTS et la phase d'interaction se fait avec M ou M^2 combinaisons. Les équations des deux types d'interaction correspondent aux équations des deux types de fusion mis en évidence par [Helmick et al., 1995]. Par contre, il n'y a maintenant plus besoin de déterminer une dynamique à rebours et la passe arrière s'initialise de manière transparente. En effet, il suffit de connaître la moyenne et la covariance associées à chaque mode à l'instant final pour démarrer la récursion arrière. [Helmick et al., 1995] doit par contre faire l'effort d'initialiser le filtre IMM arrière sans information a priori ou de s'approcher de cette situation comme nous le verrons plus tard dans un exemple sur des données simulées.

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

Algorithme 2 UN CYCLE DE LISSEUR IMM À INTERVALLE FIXE

$\left[\{\hat{x}_{k|T}^j, P_{k|T}^j, \mu_{k|T}^j\}_{j \in \mathcal{M}} \right]$
 = **LISSEUR-IMM** $\left(\{\hat{x}_{k+1|T}^i, P_{k+1|T}^i, \mu_{k+1|T}^i\}_{i \in \mathcal{M}}, \{\hat{x}_{k|k}^j, P_{k|k}^j, \mu_{k|k}^j\}_{j \in \mathcal{M}}, \{\bar{x}_{k|k}^i, \bar{P}_{k|k}^i\}_{i \in \mathcal{M}}, \{\hat{x}_{k+1|k}^i, P_{k+1|k}^i\}_{i \in \mathcal{M}} \right)$

- 1: **POUR** $i \in \mathcal{M}$ {1. Lissage conditionné sur le mode}
- 2: Gain du lisseur : $G_k^i = C_{k,k+1}^i (P_{k+1|k}^i)^{-1}$ où $C_{k,k+1}^i = \int_{x_k} (x_k - \bar{x}_{k|k}^i) (f_{k+1}^i(x_k) - \hat{x}_{k+1|k}^i)^T \mathcal{N}(x_k; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i) dx_k$.
- 3: Moyenne lissée d'interaction : $\bar{x}_{k|T}^i = \bar{x}_{k|k}^i + G_k^i (\hat{x}_{k+1|T}^i - \hat{x}_{k+1|k}^i)$.
- 4: Covariance lissée d'interaction : $\bar{P}_{k|T}^i = \bar{P}_{k|k}^i + G_k^i (P_{k+1|T}^i - P_{k+1|k}^i) (G_k^i)^T$.
- 5: **FIN POUR**
- 6: **POUR** $i \in \mathcal{M}$ {2. Interaction}
- 7: Matrice d'information prédite arrière : $(P_{k|k+1}^{b,i})^{-1} = (\bar{P}_{k|T}^i)^{-1} - (\bar{P}_{k|k}^i)^{-1}$.
- 8: Vecteur d'information prédit arrière : $(P_{k|k+1}^{b,i})^{-1} \hat{x}_{k+1}^{b,i} = (\bar{P}_{k|T}^i)^{-1} \bar{x}_{k|T}^i - (\bar{P}_{k|k}^i)^{-1} \bar{x}_{k|k}^i$.
- 9: **POUR** $j \in \mathcal{M}$
- 10: Vraisemblance sur deux modes : $\Lambda_{ji} \approx \mathcal{N}(\Delta_k^{ji}; 0, D_k^{ji})$ avec $\Delta_k^{ji} = \hat{x}_{k|k+1}^{b,i} - \hat{x}_{k|k}^j$ et $D_k^{ji} = P_{k|k+1}^{b,i} + P_{k|k}^j$.
- 11: Probabilité d'interaction lissée : $\bar{\mu}_{k+1|T}^{ij} = \frac{\pi_{ji} \Lambda_{ji}}{\sum_{i \in \mathcal{M}} \pi_{ji} \Lambda_{ji}} \left(\approx \pi_{ji} \text{ si } (P_{k|k+1}^{b,i})^{-1} \text{ est singulière} \right)$.
- 12: **FIN POUR**
- 13: **FIN POUR**
- 14: SI l'interaction M2 est choisie OU $(P_{k|k+1}^{b,i})^{-1}$ est singulière ALORS
- 15: $\left[\{\hat{x}_{k|T}^j, P_{k|T}^j\}_{j \in \mathcal{M}} \right]$
 = **INTERACTION-M2** $\left(\{(P_{k|k}^j)^{-1} \hat{x}_{k|k}^j, (P_{k|k}^j)^{-1}\}_{j \in \mathcal{M}}, \{(P_{k|k+1}^{b,i})^{-1} \hat{x}_{k+1}^{b,i}, (P_{k|k+1}^{b,i})^{-1}\}_{i \in \mathcal{M}}, \{\bar{\mu}_{k+1|T}^{ij}\}_{(i,j) \in \mathcal{M}^2} \right)$
- 16: SINON
- 17: $\left[\{\hat{x}_{k|T}^j, P_{k|T}^j\}_{j \in \mathcal{M}} \right]$
 = **INTERACTION-M** $\left(\{(P_{k|k}^j)^{-1} \hat{x}_{k|k}^j, (P_{k|k}^j)^{-1}\}_{j \in \mathcal{M}}, \{\hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i}\}_{i \in \mathcal{M}}, \{\bar{\mu}_{k+1|T}^{ij}\}_{(i,j) \in \mathcal{M}^2} \right)$
- 18: **POUR** $j \in \mathcal{M}$
- 19: Probabilité lissée des modes : $\mu_{k|T}^j = \sum_{i \in \mathcal{M}} \bar{\mu}_{k|k}^{ij} \mu_{k+1|T}^i$
- 20: **FIN POUR**
- 21: Moyenne lissée globale : $\hat{x}_{k|T} = \sum_{j \in \mathcal{M}} \mu_{k|T}^j \hat{x}_{k|T}^j$.
- 22: Covariance globale de l'erreur de lissage : $P_{k|T} = \sum_{j \in \mathcal{M}} \mu_{k|T}^j \left[P_{k|T}^j + (\hat{x}_{k|T}^j - \hat{x}_{k|T}) (\hat{x}_{k|T}^j - \hat{x}_{k|T})^T \right]$.

Algorithme 3 INTERACTION AVEC M^2 COMBINAISONS

$\left[\{\hat{x}_{k|T}^j, P_{k|T}^j\}_{j \in \mathcal{M}} \right]$
 = **INTERACTION-M2** $\left(\{(P_{k|k}^j)^{-1} \hat{x}_{k|k}^j, (P_{k|k}^j)^{-1}\}_{j \in \mathcal{M}}, \{(P_{k|k+1}^{b,i})^{-1} \hat{x}_{k+1}^{b,i}, (P_{k|k+1}^{b,i})^{-1}\}_{i \in \mathcal{M}}, \{\bar{\mu}_{k+1|T}^{ij}\}_{(i,j) \in \mathcal{M}^2} \right)$

- 1: **POUR** $i \in \mathcal{M}$ {1. Combinaison}
- 2: **POUR** $j \in \mathcal{M}$
- 3: Moyenne lissée conditionnée sur deux modes : $\hat{x}_{k|T}^{ji} = P_{k|T}^{ji} \left[(P_{k|k+1}^{b,i})^{-1} \hat{x}_{k+1}^{b,i} + (P_{k|k}^j)^{-1} \hat{x}_{k|k}^j \right]$.
- 4: Covariance lissée conditionnée sur deux modes : $P_{k|T}^{ji} = \left[(P_{k|k+1}^{b,i})^{-1} + (P_{k|k}^j)^{-1} \right]^{-1}$.
- 5: **FIN POUR**
- 6: **FIN POUR**
- 7: **POUR** $j \in \mathcal{M}$ {2. Interaction}
- 8: Moyenne lissée conditionnée sur un mode : $\hat{x}_{k|T}^j = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{ij} \hat{x}_{k|T}^{ji}$.
- 9: Covariance lissée conditionnée sur un mode : $P_{k|T}^j = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{ij} \left[P_{k|T}^{ji} + (\hat{x}_{k|T}^{ji} - \hat{x}_{k|T}^j) (\hat{x}_{k|T}^{ji} - \hat{x}_{k|T}^j)^T \right]$.
- 10: **FIN POUR**

Algorithme 4 INTERACTION AVEC M COMBINAISONS

$\left[\{\hat{x}_{k|T}^j, P_{k|T}^j\}_{j \in \mathcal{M}} \right]$
 $= \text{INTERACTION-M} \left(\left\{ (P_{k|k}^j)^{-1} \hat{x}_{k|k}^j, (P_{k|k}^j)^{-1} \right\}_{j \in \mathcal{M}}, \left\{ \hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i} \right\}_{i \in \mathcal{M}}, \left\{ \bar{\mu}_{k+1|T}^{i|j} \right\}_{(i,j) \in \mathcal{M}^2} \right)$

1: **POUR** $j \in \mathcal{M}$ {1. Interaction IMM arrière}
 2: Moyenne arrière d'interaction : $\bar{x}_{k|k+1}^{b,j} = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{i|j} \hat{x}_{k|k+1}^{b,i}$.
 3: Covariance arrière d'interaction : $\bar{P}_{k|k+1}^{b,j} = \sum_{i \in \mathcal{M}} \bar{\mu}_{k+1|T}^{i|j} [P_{k|k+1}^{b,i} + (\hat{x}_{k|k+1}^{b,i} - \bar{x}_{k|k+1}^{b,j})(\hat{x}_{k|k+1}^{b,i} - \bar{x}_{k|k+1}^{b,j})^T]$.
 4: **FIN POUR**
 5: **POUR** $j \in \mathcal{M}$ {2. Combinaison}
 6: Moyenne lissée conditionnée sur un mode : $\hat{x}_{k|T}^j = P_{k|T}^j \left[(\bar{P}_{k|k+1}^{b,j})^{-1} \bar{x}_{k|k+1}^{b,j} + (P_{k|k}^j)^{-1} \hat{x}_{k|k}^j \right]$.
 7: Covariance lissée conditionnée sur un mode : $P_{k|T}^j = \left[(\bar{P}_{k|k+1}^{b,j})^{-1} + (P_{k|k}^j)^{-1} \right]^{-1}$.
 8: **FIN POUR**

TABLE 3.1.: Principales quantités manipulées par le lisseur IMM-RTS (en plus de celles de l'IMM de Tab. 2.1)

Quantité	Notation / Approximation	Nom
$p(x_k z_{1:T})$	$\mathcal{N}(x_k; \hat{x}_{k T}, P_{k T})$	densité lissée globale au temps k
$p(x_{k+1} m_{k+1}^i, z_{1:T})$	$\mathcal{N}(x_{k+1}; \hat{x}_{k+1 T}^i, P_{k+1 T}^i)$	densité lissée au temps $k+1$ conditionnée sur m_{k+1}^i
$p(x_k m_{k+1}^i, z_{1:T})$	$\mathcal{N}(x_k; \bar{x}_{k T}^i, \bar{P}_{k T}^i)$	densité d'interaction lissée au temps k conditionnée sur m_{k+1}^i
$p(x_k m_{k+1}^i, z_{k+1:T})$	$\mathcal{N}(x_k; \hat{x}_{k k+1}^{b,i}, P_{k k+1}^{b,i})$	densité prédite arrière au temps k conditionnée sur m_{k+1}^i
$p(x_k m_k^j, z_{k+1:T})$	$\mathcal{N}(x_k; \bar{x}_{k k+1}^{b,j}, \bar{P}_{k k+1}^{b,j})$	densité d'interaction arrière au temps k conditionnée sur m_k^j
$p(x_k m_{k+1}^i, m_k^j, z_{1:T})$	$\mathcal{N}(x_k; \hat{x}_{k T}^{ji}, P_{k T}^{ji})$	densité lissée au temps k conditionnée sur m_{k+1}^i et m_k^j
$P(m_{k+1}^i z_{1:T})$	$\mu_{k+1 T}^i$	probabilité lissée au temps $k+1$ du mode i
$P(m_{k+1}^i m_k^j, z_{1:T})$	$\bar{\mu}_{k+1 T}^{i j}$	probabilité d'interaction lissée au temps $k+1$ entre m_k^j et m_{k+1}^i

3.1.6. Extension au lissage à retard fixe

Même quand un traitement temps réel est indispensable, une estimation lissée du vecteur d'état du système peut être délivrée avec un délai constant à l'aide d'un lisseur à retard fixe. Pour le cas mono-modèle, la densité lissée $p(x_k|z_{1:k+n})$ au temps k , avec $n \geq 1$ la taille du retard, peut être obtenue selon deux approches. La première méthode se résume à construire un vecteur d'état augmenté X_{k+n} en concaténant les vecteurs d'état entre les instants k et $k+n$, *i.e.* en posant $X_{k+n} = (x_{k+n}^T, \dots, x_k^T)^T$ [Gustafsson, 2000]. L'estimée lissée de l'état est alors calculée en appliquant un filtrage sur cet état augmenté, avant de sélectionner la partie correspondant à x_k . C'est une solution particulièrement intrusive. En effet, les modèles de dynamique et d'observation doivent être eux aussi étendus pour se conformer au vecteur d'état augmenté et le filtre de Kalman doit être modifié comme proposé dans [Gustafsson, 2000, Eqs. (8.66)-(8.70)] pour conserver une complexité augmentant linéairement avec n . La seconde solution [Särkkä and Hartikainen, 2010] est d'utiliser directement les équations de lissage de Rauch-Tung-Striebel sur l'intervalle de temps commençant à k et se terminant à $k+n$. Une fois que l'estimée filtrée de l'état au temps $k+n$ est connue, une passe arrière de type Rauch-Tung-Striebel est appliquée pour $k+n-1, \dots, k$ comme un simple post-traitement. Pour cette stratégie, la charge de calcul s'accroît aussi linéairement avec la taille du retard.

Dans le contexte multi-modèle, la solution la plus répandue est sans doute l'augmentation du vecteur d'état. Cette technique est proposée dans [Chen and Tugnait, 2000] avec un filtre IMM linéaire agissant sur un système à sauts markoviens augmenté. La densité de l'état augmenté $p(X_{k+n}|m_{k+n}, z_{1:k+n})$ est par contre conditionnée sur le modèle m_{k+n} actif à $k+n$ de sorte que l'estimée lissée de x_k est aussi conditionnée sur m_{k+n} (et non sur m_k comme on pourrait le souhaiter). Une extension non linéaire est détaillée dans [Morelande and Ristic, 2008] avec un filtre IMM basé sur un banc d'UKF. Les auteurs montrent que la transformée en points sigma de l'UKF peut être réalisée avec le même nombre de points quelle que soit la taille du retard étant donné que les modèles augmentés sont partiellement linéaires. Un lisseur exploitant un filtre IMM sans en augmenter l'état a été introduit dans [Helmick et al., 1996] pour $n = 1$ seulement et par [Jilkov et al., 2002] pour un retard quelconque, mais sous l'hypothèse qu'il n'y a pas de changement de mode dans l'intervalle $(k, k+n]$.

Avec les développements ci-dessus, l'estimée lissée à l'instant k est obtenue avec un retard constant simplement en appliquant la récursion en deux étapes —lissage multi-modèle de type Rauch-Tung-Striebel et interaction— pour $k+n-1, \dots, k$ après qu'un filtrage IMM a été mené jusqu'au temps $k+n$. Comme évoqué précédemment, cette technique peut être considérée comme un traitement additionnel sur les estimées filtrées en temps réel et ne nécessite pas de modifier les modèles existants. L'estimée lissée au temps k est bien conditionnée sur m_k , au lieu de m_{k+n} pour l'approche avec état augmenté et tient compte des sauts de modes dans l'intervalle $]k, k+n]$ dans la phase d'interaction. Remarquons enfin que dans l'algorithme 2, les gains de lissage G_k^i , les inverses $(\bar{P}_{k|k}^i)^{-1}$ et $(P_{k|k}^j)^{-1}$, les produits $(\bar{P}_{k|k}^i)^{-1} \bar{x}_{k|k}^i$ et $(P_{k|k}^j)^{-1} \hat{x}_{k|k}^j$ peuvent être précalculés et stockés pendant le filtrage IMM afin d'obtenir une exécution plus rapide des récursions successives du lisseur à retard fixe.

3.2. Filtrage multi-modèle hétérogène

L'emploi d'un ensemble *hétérogène* de modèles, c'est-à-dire avec des vecteurs d'état de taille et de sémantique différentes est un sujet peu investigué dans la littérature. Le problème est souvent contourné en rallongeant avec des zéros les vecteurs d'état pour que chacun soit ramené à une même taille préalablement à l'étape d'interaction de l'IMM. Par exemple, lors de l'estimation de la position et de la vitesse avec deux modèles dont l'un n'estime que la position, il est courant d'augmenter le plus petit vecteur d'état avec des zéros [Bar-Shalom et al., 2011]. Cette approche n'est pas très satisfaisante car elle génère des biais vers zéro lors de l'interaction [Yuan et al., 2012].

Dans cette partie, nous allons d'abord reformuler la problématique des systèmes à sauts markoviens pour des ensembles de modèles hétérogènes puis construire étape par étape une version généralisée de l'IMM. Celle-ci a fait l'objet de la publication [Lopez et al., 2010].

3.2.1. Formulation d'un système à sauts markoviens hétérogène

On considère toujours un ensemble de modes $\mathcal{M}$ mais avec des vecteurs d'état spécifiques à chaque mode. Soit x^j le vecteur d'état continu relatif au mode j . Soit $x = (x_1, \dots, x_p, \dots, x_n)^T$ le vecteur d'état continu complet issu de l'union sans doublons des variables impliquées dans chaque mode. On suppose que les vecteurs d'état de chaque mode partagent un ensemble de composantes communes. Par exemple, on peut imaginer un ensemble de modèles qui estiment tous la position d'un mobile mais qui peuvent chacun estimer des variables supplémentaires de vitesse, d'accélération, de rotation... On fait aussi l'hypothèse que la sortie observée z_k est de forme identique quel que soit le mode actif. Lorsque m_k^j se réalise, la dynamique de l'état continu $p(x_k^j | x_{k-1}^j, m_k^j)$ et son lien avec la mesure $p(z_k | x_k^j, m_k^j)$ se réécrivent maintenant de la façon suivante :

$$x_k^j = f_k^j(x_{k-1}^j) + q_{k-1}^j, \quad z_k = g_k^j(x_k^j) + r_k^j. \quad (3.35)$$

Autrement dit, pendant l'intervalle où m_k^j est actif, les variables de x^j suivent la dynamique du mode j et la mesure dépend directement de ces variables au travers de la fonction d'observation du mode j . Les variables appartenant à x mais pas à x^j évoluent librement a priori car elles ne sont pas explicitement modélisées par la dynamique a priori de m_k^j . Les statistiques des bruits sont données par

$$\forall k, k', E \begin{bmatrix} x_0^j | m_0^j \\ q_k^j \\ r_{k'}^j \end{bmatrix} = \begin{pmatrix} \hat{x}_{0|0}^j \\ 0 \\ 0 \end{pmatrix}; \quad (3.36)$$

$$E \left[\begin{pmatrix} x_0^j | m_0^j - \hat{x}_{0|0}^j \\ q_k^j \\ r_k^j \end{pmatrix} \begin{pmatrix} x_0^j | m_0^j - \hat{x}_{0|0}^j \\ q_{k'}^j \\ r_{k'}^j \end{pmatrix}^T \right] = \begin{pmatrix} P_{0|0}^j & 0 & 0 \\ 0 & Q_k^j \delta_{k,k'} & 0 \\ 0 & 0 & R_k^j \delta_{k,k'} \end{pmatrix}. \quad (3.37)$$

En transposant tout cela dans le cycle des équations de filtrage exact entre k et $k+1$, on obtient :

1. $\forall i \in \mathcal{M}, \{P(m_k^i | z_{1:k})\}_{j \in \mathcal{M}} \xrightarrow{\text{Prédiction}} P(m_{k+1}^i | z_{1:k})$
2. $\forall i \in \mathcal{M}, \{p(x_k^j | m_k^j, z_{1:k})\}_{j \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k^i | m_{k+1}^i, z_{1:k})$
3. $\forall i \in \mathcal{M}, p(x_k^i | m_{k+1}^i, z_{1:k}) \xrightarrow{\text{Prédiction}} p(x_{k+1}^i | m_{k+1}^i, z_{1:k})$

TABLE 3.2.: Notations spécifiques de l'IMMHO

Vecteur d'état complet	$x = (x_1, \dots, x_p, \dots, x_n)^T$
Composante p de x	x_p
Vecteur d'état du mode i	x^i
Composante p de x^i	$x_p^i = x_p$
Ensemble des modes impliquant x_p	$\mathcal{M}_p \subseteq \mathcal{M}$ (cardinalité M_p)
Ensemble des modes impliquant x_p et x_q	$\mathcal{M}_{pq} = \mathcal{M}_p \cap \mathcal{M}_q$ (cardinalité M_{pq})
Évènement $m_k = j$ et $j \in \mathcal{M}_{pq}$	$m_k^{j \in \mathcal{M}_{pq}}$

4. $\forall i \in \mathcal{M}, p(x_{k+1}^i | m_{k+1}^i, z_{1:k}) \xrightarrow{\text{Correction}} p(x_{k+1}^i | m_{k+1}^i, z_{1:k+1})$.
5. $\forall i \in \mathcal{M}, P(m_{k+1}^i | z_{1:k}) \xrightarrow{\text{Correction}} P(m_{k+1}^i | z_{1:k+1})$

Les étapes 3-4 de prédiction et correction sur l'état continu se développent de manière classique. Elles se résument à appliquer un filtre associé au mode m_{k+1}^i sur le vecteur d'état x_k^i . Par contre, l'interaction suppose de construire la densité de x_k^j conditionnée sur le mode m_{k+1}^i à partir des densités de x_k^j conditionnées sur les modes m_k^j . Ces densités font intervenir des vecteurs d'état de formes potentiellement différentes et la construction de la densité d'interaction n'est de manière générale plus aussi directe que l'équation (2.19). Néanmoins, nous allons proposer un moyen de mener cette étape dans un contexte gaussien et sous le paradigme de la fusion des branches de l'arbre des séquences possibles.

3.2.2. Stratégie d'interaction par composante

L'idée centrale de l'IMM *Heterogeneous-Order* (IMMHO) est de développer une étape d'interaction propre à chaque variable manipulée dans les densités. Ceci est légitime car chaque composante x_p du vecteur d'état continu complet x est impliquée dans un sous-ensemble spécifique de modèles $\mathcal{M}_p \subseteq \mathcal{M}$. Par abus de notation, si x^j est formé avec les entrées $j_1, j_2, j_3, \dots$ de x , alors il sera noté soit $x^j = (x_{j_1}, x_{j_2}, x_{j_3}, \dots)^T$ ou $x^j = (x_{j_1}^j, x_{j_2}^j, x_{j_3}^j, \dots)^T$ (en d'autres termes $x_p^j = x_p$). Cela permet de clarifier sensiblement les développements qui suivent. Toutes les nouvelles notations qui nous serviront dans la suite sont rassemblées dans le tableau 3.2.

Équivalence des densités

Pour construire l'IMMHO, on suppose que les densités conditionnées sur les modes sont gaussiennes et peuvent donc être réduites à leurs deux premiers moments. Ainsi, la densité a posteriori $p(x_k^j | m_k^j, z_{1:k})$ conditionnée sur le mode j à l'instant k est équivalente à l'ensemble des densités conjointes

$$p(x_k^j | m_k^j, z_{1:k}) \iff \left\{ p(x_{k,p}, x_{k,q} | m_k^j, j \in \mathcal{M}_p \cap \mathcal{M}_q, z_{1:k}) \right\}_{(p,q)} \quad (3.38)$$

où les indices p et q sont définis de sorte que $x_{k,p}$ et $x_{k,q}$ correspondent aux composantes possibles de x_k^j . Ci-dessus, $\mathcal{M}_{pq} \triangleq \mathcal{M}_p \cap \mathcal{M}_q$ est l'ensemble des modèles mettant en jeu à la fois $x_{k,p}$ et $x_{k,q}$. Les densités $p(x_{k,p}, x_{k,q} | m_k^j, j \in \mathcal{M}_{pq}, z_{1:k}) = p(x_{k,p} | m_k^j, j \in \mathcal{M}_p, z_{1:k})$ reconstruisent les espérances et les variances de $p(x_k^j | m_k^j, z_{1:k})$. Les densités $p(x_{k,p}, x_{k,q} | m_k^j, j \in \mathcal{M}_{pq}, z_{1:k})$ avec $p \neq q$ restituent les termes de covariance. Cette équivalence est essentielle pour garantir qu'il n'y a pas de perte d'information et sera exploitée dans toutes les densités manipulées dans l'IMMHO. Si l'hypothèse gaussienne n'est pas valable et si des moments d'ordre supérieur sont requis, alors $p(x_k^j | m_k^j, z_{1:k})$ doit être subdivisée en un ensemble de densités conjointes sur tous les n -uplets possibles de variables estimées. Dorénavant, on désigne l'évènement $(m_k^j, j \in \mathcal{M}_{pq})$ par $m_k^{j \in \mathcal{M}_{pq}}$.

Reconstruction de la moyenne et de la covariance d'interaction

Partant de l'approximation gaussienne des densités $\{p(x_k^j | m_k^j, z_{1:k})\}_{j \in \mathcal{M}}$, on rappelle que l'IMM applique un filtre associé au mode m_{k+1}^i sur la moyenne $\bar{x}_{k|k}^i = E[x_k^i | m_{k+1}^i, z_{1:k}]$ et la covariance $\bar{P}_{k|k}^i = E[(x_k^i - \bar{x}_{k|k}^i)(x_k^i - \bar{x}_{k|k}^i)^T | m_{k+1}^i, z_{1:k}]$ de la densité d'interaction $p(x_k^i | m_{k+1}^i, z_{1:k})$. Leur détermination dans le cas hétérogène s'effectue selon le théorème suivant.

Théorème 6 (Interaction hétérogène) *Pour chaque couple (p, q) de composantes, on suppose que la séquence des modes appartenant à $\mathcal{M}_{pq}$ suit une chaîne de Markov avec $\pi_{ji, pq} = P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}})$ la probabilité de transition du mode j vers i . Dans un contexte hétérogène et gaussien, la moyenne $\bar{x}_{k|k}^i$ et la covariance $\bar{P}_{k|k}^i$ de $p(x_k^i | m_{k+1}^i, z_{1:k}) \approx \mathcal{N}(x_k^i; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i)$ sont données*

1. pour chaque composante p de la moyenne $\bar{x}_{k|k}^i$

$$\begin{aligned} \bar{x}_{k|k,p}^i &= E \left[x_{k,p}^i | m_{k+1}^{i \in \mathcal{M}_p}, z_{1:k} \right] \\ &= \sum_{j \in \mathcal{M}_p} \bar{x}_{k|k,p}^j \mu_{k|k,p}^{j|i} \end{aligned} \quad (3.39)$$

$$\text{où } \mu_{k|k,p}^{j|i} = \mu_{k|k,pp}^{j|i} = P(m_k^{j \in \mathcal{M}_p} | m_{k+1}^{i \in \mathcal{M}_p}, z_{1:k});$$

2. pour chaque entrée (p, q) de la covariance $\bar{P}_{k|k}^i$

$$\begin{aligned} \bar{P}_{k|k,pq}^i &= E \left[\left(\bar{x}_{k|k,p}^i - x_{k,p}^i \right) \left(\bar{x}_{k|k,q}^i - x_{k,q}^i \right) | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k} \right] \\ &= \sum_{j \in \mathcal{M}_{pq}} \mu_{k|k,pq}^{j|i} [P_{k|k,pq}^j + (\bar{x}_{k|k,p}^j - \hat{x}_{k|k,p}^j)(\bar{x}_{k|k,q}^j - \hat{x}_{k|k,q}^j)]. \end{aligned} \quad (3.40)$$

$$\text{où } \mu_{k|k,pq}^{j|i} = \mu_{k|k,qp}^{j|i} = P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}).$$

Preuve Nous reconstituons les deux premiers moments de $p(x_k^i | m_{k+1}^i, z_{1:k})$ en appliquant l'interaction sur chaque élément de l'ensemble équivalent, à savoir $\left\{ p(x_{k,p}, x_{k,q} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) \right\}_{(p,q)}$ avec

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

p et q définis de sorte que $x_{k,p}$ et $x_{k,q}$ correspondent aux composantes de x_k^i . L'interaction sur un élément est :

$$P(x_{k,p}, x_{k,q} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) = \sum_{j \in \mathcal{M}_{pq}} P(x_{k,p}, x_{k,q} | m_{k+1}^{i \in \mathcal{M}_{pq}}, m_k^{j \in \mathcal{M}_{pq}}, z_{1:k}) \underbrace{P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k})}_{\mu_{k|k,pq}^{ji}}. \quad (3.41)$$

Enfin, les éléments de la moyenne et de la covariance se déduisent de (3.41) à partir d'une approximation par moment. L'IMMHO exploite par ailleurs l'indépendance de m_{k+1} et $(x_{k,p}, x_{k,q})$ sous la condition que m_k est connu (de manière similaire à l'IMM, cette indépendance découle des propriétés markoviennes de la suite des modes spécifiques au couple (p, q)). Seuls les modes de l'ensemble $\mathcal{M}_{pq} \subset \mathcal{M}$ estimant les composantes p et q sont pris en compte alors que les autres modes sont des événements impossibles d'un point de vue probabiliste. $\square$

L'interaction permet de partager de l'information entre les modes. Ainsi, un mode peut influencer les variables d'un autre mode, sans pour autant les estimer, à travers les valeurs des covariances.

Probabilités d'interaction par couple de variables

Les équations (3.39)–(3.40) introduisent des probabilités d'interaction $\mu_{k|k,pq}^{ji}$ spécifiques à une paire de variable $(x_{k,p}, x_{k,q})$. Leur définition inclut les événements $j \in \mathcal{M}_{pq}$ et $i \in \mathcal{M}_{pq}$ de manière à réduire l'ensemble complet des modèles à une portion estimant x_p et x_q simultanément.

Théorème 7 (Probabilités d'interaction hétérogène) *De manière similaire à l'établissement de l'IMM, la probabilité d'interaction par couple se décline selon*

$$\mu_{k|k,pq}^{ji} = \frac{\pi_{ji,pq} \mu_{k|k,pq}^j}{\mu_{k+1|k,pq}^i} \quad (3.42)$$

où $\mu_{k|k,pq}^j$ est la probabilité a posteriori $P(m_k^{j \in \mathcal{M}_{pq}} | z_{1:k})$ du mode j à l'instant k et $\mu_{k+1|k,pq}^i$ est la probabilité prédite $P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k})$. En outre, $\mu_{k+1|k,pq}^i = \sum_{l \in \mathcal{M}_{pq}} \pi_{li,pq} \mu_{k|k,pq}^l$

Preuve À l'image de l'IMM classique, la probabilité d'interaction hétérogène se décompose selon :

$$P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) = \frac{P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:k}) P(m_k^{j \in \mathcal{M}_{pq}} | z_{1:k})}{P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k})} \quad (3.43)$$

avec $P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k}) = \sum_{l \in \mathcal{M}_{pq}} P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{l \in \mathcal{M}_{pq}}, z_{1:k}) P(m_k^{l \in \mathcal{M}_{pq}} | z_{1:k})$. $\square$

Pour j et k donnés, seule une matrice symétrique des probabilités a posteriori $\mu_{k|k,pq}^j$ est maintenue pour la récursion, du fait que les probabilités de transition $\pi_{ij,pq}$ sont des données de l'algorithme et que les autres termes se déduisent de l'interaction.

Mise à jour des probabilités des modes

Théorème 8 (Probabilités filtrées) Soit $L_{k+1}^i = p(z_{k+1}|m_{k+1}^i, z_{1:k})$ la vraisemblance du mode i lors de l'assimilation de la mesure z_{k+1} . L_{k+1}^i se déduit de manière classique à partir du filtre associé au mode i . La probabilité filtrée $\mu_{k+1|k+1,pq}^i = P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k+1})$ du mode i au temps $k+1$ est donnée par :

$$\mu_{k+1|k+1,pq}^i = \frac{\mu_{k+1|k,pq}^i L_{k+1}^i}{\sum_{l \in \mathcal{M}_{pq}} \mu_{k+1|k,pq}^l L_{k+1}^l} \text{ si } i \in \mathcal{M}_{pq} \quad (3.44)$$

$$= 0 \text{ sinon.} \quad (3.45)$$

Preuve La probabilité $\mu_{k+1|k+1,pq}^i$ se décompose selon

$$P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k+1}) = \frac{P(z_{k+1}|m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k})}{\sum_{l \in \mathcal{M}_{pq}} P(z_{k+1}|m_{k+1}^{l \in \mathcal{M}_{pq}}, z_{1:k}) P(m_{k+1}^{l \in \mathcal{M}_{pq}} | z_{1:k})}. \quad (3.46)$$

Si $i \in \mathcal{M}_{pq}$ alors $p(z_{k+1}|m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) = p(z_{k+1}|m_{k+1}^i, z_{1:k})$. Si $i \notin \mathcal{M}_{pq}$ alors la densité $p(z_{k+1}|m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) \equiv 0$ car elle est conditionnée sur un évènement impossible. $\square$

3.2.3. Algorithme de filtrage complet

Le filtre IMMHO est synthétisé dans l'algorithme 5. Par rapport à l'IMM, les équations de filtrage sont inchangées et s'appliquent à chaque vecteur d'état continu. Nous avons modifié les phases d'interaction, de mise à jour des probabilités et de calcul de l'estimée globale. Pour construire la moyenne $\hat{x}_{k|k}$ et la covariance $P_{k|k}$ globales de l'état complet x , on procède aussi par couple de composantes avec $\hat{x}_{k|k,p} = E[x_{k,p} | z_{1:k}]$ et $P_{k|k,pq} = E[(x_{k,p} - \hat{x}_{k|k,p})(x_{k,q} - \hat{x}_{k|k,q}) | z_{1:k}]$.

L'IMMHO constitue au final une version généralisée de l'IMM pour des ensembles de modèles hétérogènes. Nous avons fait tout au début l'hypothèse qu'il existe un ensemble de variables communes à tous les modèles. On peut toutefois remarquer que les démonstrations des théorèmes qui mènent aux équations de l'IMMHO ne reposent pas sur cette supposition. Par contre, les probabilités de transition entre les modes, les probabilités prédite et filtrée des modes n'ont de sens, indépendamment d'un couple (p, q) , que s'il existe un ensemble de variables communes à tous les modes. Si les composantes p et q font partie de l'ensemble des variables communes à tous les modèles, on a $\mathcal{M}_{pq} = \mathcal{M}$ et il devient alors possible de se rattacher aux équations de filtrage exact, établies dans le paragraphe 3.2.1, avec :

$$P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}) = P(m_{k+1}^i | m_k^j) \quad (3.47)$$

$$P(m_k^{j \in \mathcal{M}_{pq}} | z_{1:k}) = P(m_k^j | z_{1:k}) \quad (3.48)$$

$$P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:k}) = P(m_{k+1}^i | z_{1:k}). \quad (3.49)$$

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

Enfin, s'il suffit de connaître les probabilités de transition $P\left(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}\right)$ pour tous les couples (p, q) possibles afin de déterminer $P\left(m_{k+1}^i | m_k^j\right)$, la réciproque est fautive. Le développement d'un filtre multi-modèle hétérogène en exploitant le cadre de l'IMM demande ainsi de définir des probabilités de transition, plus générales, par couple de variables.

Algorithme 5 FILTRAGE IMMHO

- 1: Probabilités de transition (données) : $\pi_{ji,pq} = P\left(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}\right)$ ($= 0$ si $i \notin \mathcal{M}_{pq}$ ou $j \notin \mathcal{M}_{pq}$)
 - 2: Probabilités initiales des modes pour (x_p, x_q) (données) : $\mu_{0|0,pq}^i$ ($= 0$ si $i \notin \mathcal{M}_{pq}$ ou $j \notin \mathcal{M}_{pq}$).
 - 3: Moyenne et covariance initiales (données) : $\hat{x}_{0|0}^i$ et $P_{0|0}^i$.
 - 4: **SI** $k \geq 0$ **ALORS**
 - 5: **POUR** $i \in \mathcal{M}$ {Interaction}
 - 6: Probabilité prédite du mode de (x_p, x_q) : $\mu_{k+1|k,pq}^i = \sum_{j \in \mathcal{M}_{pq}} \pi_{ji,pq} \mu_{k|k,pq}^j$ si $i \in \mathcal{M}_{pq}$, 0 sinon.
 - 7: Probabilité d'interaction de (x_p, x_q) : $\mu_{k|k,pq}^{ji} = \pi_{ji,pq} \mu_{k|k,pq}^j / \mu_{k+1|k,pq}^i$ si $i \in \mathcal{M}_{pq}$ et $j \in \mathcal{M}_{pq}$, 0 sinon.
 - 8: Moyenne d'interaction $\bar{x}_{k|k,p}^i : \bar{x}_{k|k,p}^i = \sum_{j \in \mathcal{M}_p} \mu_{k|k,p}^{ji} \hat{x}_{k|k,p}^j$
 - 9: Covariance d'interaction $\bar{P}_{k|k,p}^i : \bar{P}_{k|k,p}^i = \sum_{j \in \mathcal{M}_{pq}} \mu_{k|k,p}^{ji} \left[P_{k|k,pq}^j + \left(\hat{x}_{k|k,p}^j - \bar{x}_{k|k,p}^i \right) \left(\hat{x}_{k|k,q}^j - \bar{x}_{k|k,q}^i \right) \right]$
 - 10: **FIN POUR**
 - 11: **POUR** $i \in \mathcal{M}$ {Filtrage (équations classiques)}
 - 12: $(\bar{x}_{k|k}^i, \bar{P}_{k|k}^i) \xrightarrow{\text{Prédiction}} (\hat{x}_{k+1|k}^i, P_{k+1|k}^i)$
 - 13: $(\hat{x}_{k+1|k}^i, P_{k+1|k}^i) \xrightarrow{\text{Correction}} (\hat{x}_{k+1|k+1}^i, P_{k+1|k+1}^i, L_{k+1}^i)$ avec L_{k+1}^i la vraisemblance du mode
 - 14: **FIN POUR**
 - 15: **POUR** $i \in \mathcal{M}$ {Mise à jour de la probabilité des modes}
 - 16: Probabilité filtrée du mode pour (x_p, x_q) : $\mu_{k+1|k+1,pq}^i = \frac{\mu_{k+1|k,pq}^i L_{k+1}^i}{\sum_{l \in \mathcal{M}_{pq}} \mu_{k+1|k,pq}^l L_{k+1}^l}$ si $i \in \mathcal{M}_{pq}$, 0 sinon.
 - 17: **FIN POUR**
 - 18: Moyenne globale $\hat{x}_{k+1|k+1}$ de x : $\hat{x}_{k+1|k+1,p} = \sum_{i \in \mathcal{M}_p} \mu_{k+1|k+1,p}^i \hat{x}_{k+1|k+1,p}^i$
 - 19: Covariance globale $P_{k+1|k+1}$:
 - 20: $P_{k+1|k+1,pq} = \sum_{i \in \mathcal{M}_{pq}} \mu_{k+1|k+1,pq}^i \left[P_{k+1|k+1,pq}^i + \left(\hat{x}_{k+1|k+1,p}^i - \hat{x}_{k+1|k+1,p} \right) \left(\hat{x}_{k+1|k+1,q}^i - \hat{x}_{k+1|k+1,q} \right) \right]$
 - 20: **FIN SI**
-

3.3. Lissage multi-modèle hétérogène

Très naturellement, la question d'étendre l'IMMHO pour le lissage sur intervalle fixe se pose. Les deux solutions classiques sont toujours opportunes à savoir l'utilisation d'un lisseur avec double filtre IMMHO ou la mise en œuvre d'une récursion arrière de type Rauch-Tung-Striebel. Nous présentons d'abord celle-ci.

3.3.1. Lisseur IMMHO-RTS

Les équations du lissage exact sont la transposition de ER1 dans un cadre hétérogène.

1. $\forall i \in \mathcal{M}, P(x_{k+1}^i | m_{k+1}^i, z_{1:T}) \xrightarrow{\text{Lissage}} P(x_k^i | m_{k+1}^i, z_{1:T})$
2. $\forall j \in \mathcal{M}, \{P(m_{k+1}^j | z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Lissage}} P(m_k^j | z_{1:T})$

$$3. \forall j \in \mathcal{M}, \{p(x_k^i | m_{k+1}^i, z_{1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k^j | m_k^j, z_{1:T})$$

Comme pour la construction de l'IMM-RTS, nous allons d'abord nous intéresser aux étapes 1 et 3 qui concernent l'état continu (équations de lissage et d'interaction) avant d'explicitier le calcul des probabilités lissées des modes. Bon nombre des quantités de Tab. 2.1 et de Tab. 3.1 seront réutilisées dans ce qui suit.

Phase 1 : lissage sur l'état continu

Le lissage conditionné sur le mode est inchangé : les équations de lissage du théorème 1 (Rauch-Tung-Striebel) sont toujours d'actualité pour calculer les moments de $p(x_k^i | m_{k+1}^i, z_{1:T}) \approx \mathcal{N}(x_k^i; \bar{x}_{k|T}^i, \bar{P}_{k|T}^i)$ à partir de $p(x_{k+1}^i | m_{k+1}^i, z_{1:T}) \approx \mathcal{N}(x_{k+1}^i; \hat{x}_{k+1|T}^i, P_{k+1|T}^i)$. Les équations de lissage sur l'état continu de l'algorithme 2 restent donc applicables.

Phase 2 : interaction des modes avec M^2 combinaisons

Notre objectif est maintenant de calculer les deux premiers moments de $p(x_k^j | m_k^j, z_{1:T})$ à partir de l'ensemble des densités $\{p(x_k^i | m_{k+1}^i, z_{1:T})\}_{i \in \mathcal{M}}$. La première possibilité d'interaction introduite par l'équation (2.56) prend donc la forme d'une interaction hétérogène. Elle consiste à décomposer l'approximation gaussienne de $p(x_k^j | m_k^j, z_{1:T})$ selon son ensemble de densités équivalentes $\left\{ p(x_{k,p}, x_{k,q} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T}) \right\}_{(p,q)}$ et à appliquer sur chacune le théorème des probabilités totales :

$$\begin{aligned} & p(x_{k,p}, x_{k,q} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T}) \\ &= \sum_{i \in \mathcal{M}_{pq}} p(x_{k,p}, x_{k,q} | m_k^{j \in \mathcal{M}_{pq}}, m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:T}) P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T}). \end{aligned} \quad (3.50)$$

où

$$P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T}) = \frac{P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:T})}{P(m_k^{j \in \mathcal{M}_{pq}} | z_{1:T})} \quad (3.51)$$

est la probabilité d'interaction lissée par couple de variables. Les probabilités d'interaction $P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k})$ sont calculées par un filtre IMMHO et on suppose connues les probabilités lissées par couple de variables $P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:T})$. Pour ces dernières si $\mathcal{M}_{pq} = \mathcal{M}$, il est possible de faire le lien avec les probabilités des modes $P(m_{k+1}^i | z_{1:T})$ des équations exactes de lissage.

Les densités $\left\{ p(x_{k,p}, x_{k,q} | m_k^{j \in \mathcal{M}_{pq}}, m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:T}) \right\}_{(p,q)}$ sont équivalentes à la gaussienne $p(x_k^j, x_k^i | m_k^j, m_{k+1}^i, z_{1:T})$ si p et q sont définis de sorte que $x_{k,p}$ et $x_{k,q}$ correspondent à l'union sans doublon des composantes de x_k^i et x_k^j . Cette gaussienne fait intervenir les variables des modes i et j —avec des recouvrements possibles entre x^j et x^i — et se décompose selon :

$$p(x_k^j, x_k^i | m_k^j, m_{k+1}^i, z_{1:T}) \propto p(z_{k+1:T} | x_k^j, x_k^i, m_k^j, m_{k+1}^i, z_{1:k}) p(x_k^j, x_k^i | m_k^j, m_{k+1}^i, z_{1:k}). \quad (3.52)$$

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

Les densités $p(x_k^j, x_k^i | m_k^j, z_{1:k})$ et $p(z_{k+1:T} | x_k^j, x_k^i, m_{k+1}^i)$ n'ont jamais été complètement explicitées. La densité $p(x_k^j, x_k^i | m_k^j, z_{1:k})$, proche de $p(x_k^j | m_k^j, z_{1:k})$, fait intervenir des variables possiblement absentes du modèle j . Sous la condition que le mode j est actif, ces variables sont libres car non modélisées. Leurs valeurs sont supposées inconnues, ce qu'on traduit par des moyennes quelconques et des covariances associées infinies. En termes mathématiques, la moyenne $\hat{\mathbf{x}}_{k|k}^{j \cup i}$ et la covariance $\mathbf{P}_{k|k}^{j \cup i}$ de $p(x_k^j, x_k^i | m_k^j, z_{1:k})$ s'écrivent en étendant respectivement la moyenne $\hat{x}_{k|k}^j$ et la covariance $P_{k|k}^j$ de $p(x_k^j | m_k^j, z_{1:k})$ selon :

$$\begin{aligned} \hat{\mathbf{x}}_{k|k}^{j \cup i} &= \begin{pmatrix} \hat{x}_{k|k}^j \\ \hat{x}_{k|k}^{i \setminus j} \end{pmatrix} = \begin{pmatrix} \hat{x}_{k|k}^{j \setminus i} \\ \hat{x}_{k|k}^{i \cap j} \\ \hat{x}_{k|k}^{i \setminus j} \end{pmatrix} \text{ avec } \hat{x}_{k|k}^{i \setminus j} \text{ quelconque,} \\ \mathbf{P}_{k|k}^{j \cup i} &= \begin{pmatrix} P_{k|k}^j & 0 \\ 0 & \infty \end{pmatrix} = \begin{pmatrix} P_{k|k}^{j \setminus i} & C_{k|k}^{ji} & 0 \\ (C_{k|k}^{ji})^T & P_{k|k}^{j \cap i} & 0 \\ 0 & 0 & \infty \end{pmatrix}. \end{aligned} \quad (3.53)$$

L'exposant $j \setminus i$ signifie que les composantes de la moyenne ou de la covariance sont celles des composantes estimées par le modèle j mais pas par i , $j \setminus i$ pour les composantes estimées par i mais pas par j et $j \cap i$ pour les composantes estimées par i et j simultanément. Remarquons que dans $\mathbf{P}_{k|k}^{j \cup i}$, les termes de corrélation de $P_{k|k}^j$ avec le reste des entrées sont arbitrairement fixés à 0 sans perte de généralité pour la suite des calculs. Le même raisonnement conduit à l'estimée $\hat{\mathbf{x}}_{k|k+1}^{b, i \cup j}$ du maximum de vraisemblance de (x_k^j, x_k^i) relativement à $p(z_{k+1:T} | x_k^j, x_k^i, m_{k+1}^i)$ avec la covariance d'erreur associée $\mathbf{P}_{k|k+1}^{b, i \cup j}$:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k+1}^{b, i \cup j} &= \begin{pmatrix} \hat{x}_{k|k+1}^{b, j \setminus i} \\ \hat{x}_{k|k+1}^{b, i} \\ \hat{x}_{k|k+1}^{b, i \setminus j} \end{pmatrix} = \begin{pmatrix} \hat{x}_{k|k+1}^{b, j \setminus i} \\ \hat{x}_{k|k+1}^{b, i \cap j} \\ \hat{x}_{k|k+1}^{b, i \setminus j} \end{pmatrix} \text{ avec } \hat{x}_{k|k+1}^{b, j \setminus i} \text{ quelconque,} \\ \mathbf{P}_{k|k+1}^{b, i \cup j} &= \begin{pmatrix} \infty & 0 \\ 0 & P_{k|k+1}^{b, i} \end{pmatrix} = \begin{pmatrix} \infty & 0 & 0 \\ 0 & P_{k|k+1}^{b, i \cap j} & C_{k|k+1}^{b, ij} \\ 0 & (C_{k|k+1}^{b, ij})^T & P_{k|k+1}^{b, i \setminus j} \end{pmatrix}. \end{aligned} \quad (3.54)$$

La moyenne $\hat{\mathbf{x}}_{k|k+1}^{b, i \cup j}$ est construite en étendant l'estimée $\hat{x}_{k|k+1}^{b, i}$ de maximum de vraisemblance de x_k^i relativement à $p(z_{k+1:T} | x_k^i, m_{k+1}^i)$. C'est aussi valable pour la covariance associée $\mathbf{P}_{k|k+1}^{b, i \cup j}$ qui étend $P_{k|k+1}^{b, i}$. Les moments $\hat{x}_{k|k+1}^{b, i}$ et $P_{k|k+1}^{b, i}$ sont obtenus avec (3.18)-(3.19).

En conclusion, les opérations d'interaction sur M^2 modes se résument dans les deux théorèmes suivants.

Théorème 9 (Fusion hétérogène) *Les moments $\hat{\mathbf{x}}_{k|T}^{j \cup i}$ et $\mathbf{P}_{k|T}^{j \cup i}$ de $p(x_k^j, x_k^i | m_k^j, m_{k+1}^i, z_{1:T})$ se déduisent de la fusion hétérogène :*

$$\hat{\mathbf{x}}_{k|T}^{j \cup i} = \mathbf{P}_{k|T}^{j \cup i} \left[\left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} \hat{\mathbf{x}}_{k|k}^{j \cup i} + \left(\mathbf{P}_{k|k+1}^{b, i \cup j} \right)^{-1} \hat{\mathbf{x}}_{k|k+1}^{b, i \cup j} \right] \quad (3.55)$$

et

$$\mathbf{P}_{k|T}^{j \cup i} = \left[\left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} + \left(\mathbf{P}_{k|k+1}^{b, i \cup j} \right)^{-1} \right]^{-1}. \quad (3.56)$$

Preuve C'est l'application directe des équations (3.52), (3.53) et (3.54). $\square$

Théorème 10 (Interaction de lissage hétérogène) Dans un contexte hétérogène et gaussien, la moyenne $\hat{x}_{k|T}^j$ et la covariance $P_{k|T}^j$ de $p(x_k^j | m_k^j, z_{1:T}) \approx \mathcal{N}(x_k^j; \hat{x}_{k|T}^j, P_{k|T}^j)$ sont données

1. pour chaque composante p de la moyenne par

$$\begin{aligned} \hat{x}_{k|T,p}^j &= E \left[x_{k,p}^j | m_k^{j \in \mathcal{M}_p}, z_{1:T} \right] \\ &= \sum_{i \in \mathcal{M}_p} \hat{\mathbf{x}}_{k|T,p}^{j \cup i} \bar{\mu}_{k+1|T,p}^{i|j} \end{aligned} \quad (3.57)$$

avec $\bar{\mu}_{k+1|T,p}^{i|j} = P \left(m_{k+1}^{i \in \mathcal{M}_p} | m_k^{j \in \mathcal{M}_p}, z_{1:T} \right)$ obtenue en (3.51);

2. pour chaque entrée (p, q) de la covariance par

$$\begin{aligned} P_{k|T,pq}^j &= E \left[\left(x_{k,p}^j - \hat{x}_{k|T,p}^j \right) \left(x_{k,q}^j - \hat{x}_{k|T,q}^j \right) | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T} \right] \\ &= \sum_{i \in \mathcal{M}_{pq}} \bar{\mu}_{k+1|T,pq}^{i|j} \left[\mathbf{P}_{k|T,pq}^{j \cup i} + (\hat{\mathbf{x}}_{k|T,p}^{j \cup i} - \hat{x}_{k|T,p}^j)(\hat{\mathbf{x}}_{k|T,q}^{j \cup i} - \hat{x}_{k|T,q}^j) \right]. \end{aligned} \quad (3.58)$$

avec $\bar{\mu}_{k+1|T,pq}^{i|j} = P \left(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T} \right)$.

Preuve C'est l'application de (3.50) avec les moments obtenus du théorème précédent. $\square$

Nous avons donc construit une approximation gaussienne des densités $p(x_k^j | m_k^j, z_{1:T})$, ce qui clôt la récursion sur l'état continu.

Phase 2 (bis) : interaction des modes avec M combinaisons

De manière alternative, les densités lissées $p(x_k^j | m_k^j, z_{1:T})$ se décomposent toujours

$$p(x_k^j | m_k^j, z_{1:T}) \propto p(z_{k+1:T} | m_k^j, x_k^j, z_{1:k}) p(x_k^j | m_k^j, z_{1:k}). \quad (3.59)$$

Il nous faut d'abord adapter le théorème 4 au cadre hétérogène afin de construire l'estimée de maximum de vraisemblance $\bar{x}_{k|k+1}^{b,j}$ de $p(z_{k+1:T} | m_k^j, x_k^j)$ avec $\bar{P}_{k|k+1}^{b,j}$ la covariance associée. On rappelle qu'il est possible d'utiliser $p(x_k^j | m_k^j, z_{k+1:T}) \approx \mathcal{N}(x_k; \bar{x}_{k|k+1}^{b,j}, \bar{P}_{k|k+1}^{b,j})$ au lieu de la vraisemblance $p(z_{k+1:T} | m_k^j, x_k^j)$. La densité d'interaction arrière $p(x_k^j | m_k^j, z_{k+1:T}) \approx \mathcal{N}(x_k^j; \bar{x}_{k|k+1}^{b,j}, \bar{P}_{k|k+1}^{b,j})$ est issue de l'étape d'interaction hétérogène arrière

$$\forall j \in \mathcal{M}, \{p(x_k^i | m_{k+1}^i, z_{k+1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k^j | m_k^j, z_{k+1:T}) \quad (3.60)$$

avec la densité prédite arrière $p(x_k^i | m_{k+1}^i, z_{k+1:T}) \approx \mathcal{N}(x_k; \hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i})$ et les probabilités d'interaction arrière $\mu_{k+1|k+1,pq}^{b,i|j} = P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{k+1:T})$, égales à $\bar{\mu}_{k+1|T,pq}^{i|j} = P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T})$.

Théorème 11 Les deux premiers moments de $p(x_k^j | m_k^j, z_{k+1:T})$ sont :

1. pour chaque composante p de la moyenne $\bar{x}_{k|k+1}^{b,j}$:

$$\bar{x}_{k|k+1,p}^{b,j} = \sum_{i \in \mathcal{M}_p} \bar{\mu}_{k+1|T,p}^{i|j} \bar{x}_{k|k+1,p}^{b,i} \quad (3.61)$$

2. pour chaque entrée (p, q) de la covariance $\bar{P}_{k|k+1}^{b,j}$:

$$\bar{P}_{k|k+1,pq}^{b,j} = \sum_{i \in \mathcal{M}_{pq}} \bar{\mu}_{k+1|T,pq}^{i|j} [P_{k|k+1,pq}^{b,i} + (\bar{x}_{k|k+1,p}^{b,i} - \bar{x}_{k|k+1,p}^{b,j})(\bar{x}_{k|k+1,q}^{b,i} - \bar{x}_{k|k+1,q}^{b,j})]. \quad (3.62)$$

Preuve Elle se déroule d'une manière similaire à la preuve du théorème 4 en tenant compte du caractère hétérogène de l'interaction. $\square$

Les deux premiers moments de $p(x_k^j | m_k^j, z_{1:T})$ sont finalement obtenus en appliquant directement le théorème 5 qui ferme la récursion sur l'état continu.

Calcul des probabilités lissées des modes et des probabilités d'interaction

Les probabilités lissées des modes sont également obtenues avec les équations récursives (2.42) prises par couple de composantes :

$$P(m_k^{j \in \mathcal{M}_{pq}} | z_{1:T}) = \sum_{i \in \mathcal{M}_{pq}} P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}, z_{k+1:T}) P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:T}). \quad (3.63)$$

Dans la version homogène du lissage multi-modèle RTS, nous avons introduit une écriture alternative des probabilités d'interaction (cf. (3.20)) qui faisait intervenir l'état continu. Elle avait l'avantage d'offrir une meilleure stabilité de l'algorithme quels que soient les paramètres des modèles de dynamique et d'observation que l'on choisit. Néanmoins, son extension au cadre hétérogène n'a pas pu être établie car elle est basée sur la différence de deux moyennes conditionnées sur les modes de deux instants différents et de tailles potentiellement différentes. Nous avons donc été contraints d'en rester aux équations de récursion directes sur l'état discret (3.51) :

$$P(m_{k+1}^{i \in \mathcal{M}_{pq}} | m_k^{j \in \mathcal{M}_{pq}}, z_{1:T}) = \frac{P(m_k^{j \in \mathcal{M}_{pq}} | m_{k+1}^{i \in \mathcal{M}_{pq}}, z_{1:k}) P(m_{k+1}^{i \in \mathcal{M}_{pq}} | z_{1:T})}{P(m_k^{j \in \mathcal{M}_{pq}} | z_{1:T})}.$$

Algorithme de filtrage complet

Les équations du lisseur IMMHO-RTS sont données dans l'algorithme 6. Les équations des deux interactions sont résumées dans les algorithmes 7 et 8 qui suivent.

3.3. LISSAGE MULTI-MODÈLE HÉTÉROGÈNE

Algorithme 6 UN CYCLE DE LISSEUR IMMHO À INTERVALLE FIXE

$$\left[\left\{ \hat{x}_{k|T}^j, P_{k|T}^j, \{\mu_{k|T,pq}^j\} \right\}_{j \in \mathcal{M}} \right] = \text{LISSEUR-IMMHO} \left(\left\{ \hat{x}_{k+1|T}^i, P_{k+1|T}^i, \{\mu_{k+1|T,pq}^i\} \right\}_{i \in \mathcal{M}}, \left\{ \hat{x}_{k|k}^j, P_{k|k}^j, \{\mu_{k|k,pq}^j\} \right\}_{j \in \mathcal{M}}, \left\{ \bar{x}_{k|k}^i, \bar{P}_{k|k}^i \right\}_{i \in \mathcal{M}}, \left\{ \hat{x}_{k+1|k}^i, P_{k+1|k}^i \right\}_{i \in \mathcal{M}}, \left\{ \{\bar{\mu}_{k|k,pq}^{ji}\}_{(i,j) \in \mathcal{M}^2} \right\} \right)$$

- 1: **POUR** $i \in \mathcal{M}$ {1. Lissage conditionné sur le mode}
 - 2: Gain du lisseur : $G_k^i = C_{k,k+1}^i (P_{k+1|k}^i)^{-1}$ où $C_{k,k+1}^i = \int_{\mathcal{X}_k^i} (x_k^i - \bar{x}_{k|k}^i) (f_{k+1}^i(x_k^i) - \hat{x}_{k+1|k}^i)^T \mathcal{N}(x_k^i; \bar{x}_{k|k}^i, \bar{P}_{k|k}^i) dx_k^i$.
 - 3: Moyenne lissée d'interaction : $\bar{x}_{k|T}^i = \bar{x}_{k|k}^i + G_k^i (\hat{x}_{k+1|T}^i - \hat{x}_{k+1|k}^i)$.
 - 4: Covariance lissée d'interaction : $\bar{P}_{k|T}^i = \bar{P}_{k|k}^i + G_k^i (P_{k+1|T}^i - P_{k+1|k}^i) (G_k^i)^T$.
 - 5: **FIN POUR**
 - 6: **POUR** $i \in \mathcal{M}$ {2. Interaction}
 - 7: Matrice d'information prédite arrière : $(P_{k|k+1}^{b,i})^{-1} = (\bar{P}_{k|T}^i)^{-1} - (\bar{P}_{k|k}^i)^{-1}$.
 - 8: Vecteur d'information prédit arrière : $(P_{k|k+1}^{b,i})^{-1} \hat{x}_{k|k+1}^{b,i} = (\bar{P}_{k|T}^i)^{-1} \bar{x}_{k|T}^i - (\bar{P}_{k|k}^i)^{-1} \bar{x}_{k|k}^i$.
 - 9: **POUR** $j \in \mathcal{M}$
 - 10: Probabilité d'interaction lissée pour (x_p, x_q) : $\bar{\mu}_{k+1|T,pq}^{ij} = \frac{\bar{\mu}_{k|k,pq}^{ji} \mu_{k+1|T,pq}^i}{\sum_{l \in \mathcal{M}} \bar{\mu}_{k|k,pq}^{jl} \mu_{k+1|T,pq}^l}$.
 - 11: **FIN POUR**
 - 12: **FIN POUR**
 - 13: SI l'interaction M2 est choisie OU $(P_{k|k+1}^{b,i})^{-1}$ est singulière ALORS
 - 14: Construire $\forall (i, j) \in \mathcal{M}^2$ $\hat{\mathbf{x}}_{k|k}^{j \cup i}$ et $\mathbf{P}_{k|k}^{j \cup i}$ à partir de $\hat{x}_{k|k}^j$ et $P_{k|k}^j$ ainsi que $\hat{\mathbf{x}}_{k|k+1}^{b,i \cup j}$ et $\mathbf{P}_{k|k+1}^{b,i \cup j}$ à partir de $\hat{x}_{k|k+1}^{b,i}$ et $(P_{k|k+1}^{b,i})^{-1}$
 - 15: $\left[\left\{ \hat{x}_{k|T}^j, P_{k|T}^j \right\}_{j \in \mathcal{M}} \right]$
 $= \text{INTERACTION-M2} \left(\left\{ \left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} \hat{\mathbf{x}}_{k|k}^{j \cup i}, \left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} \right\}_{j \in \mathcal{M}}, \left\{ \left(\mathbf{P}_{k|k+1}^{b,i \cup j} \right)^{-1} \hat{\mathbf{x}}_{k|k+1}^{b,i \cup j}, \left(\mathbf{P}_{k|k+1}^{b,i \cup j} \right)^{-1} \right\}_{i \in \mathcal{M}}, \left\{ \{\bar{\mu}_{k+1|T,pq}^{ij}\}_{(i,j) \in \mathcal{M}^2} \right\} \right)$
 - 16: **SINON**
 - 17: $\left[\left\{ \hat{x}_{k|T}^j, P_{k|T}^j \right\}_{j \in \mathcal{M}} \right]$
 $= \text{INTERACTION-M} \left(\left\{ (P_{k|k}^j)^{-1} \hat{x}_{k|k}^j, (P_{k|k}^j)^{-1} \right\}_{j \in \mathcal{M}}, \left\{ \hat{x}_{k|k+1}^{b,i}, P_{k|k+1}^{b,i} \right\}_{i \in \mathcal{M}}, \left\{ \{\bar{\mu}_{k+1|T,pq}^{ij}\}_{(i,j) \in \mathcal{M}^2} \right\} \right)$
 - 18: **POUR** $j \in \mathcal{M}$
 - 19: Probabilité lissée des modes pour (x_p, x_q) : $\mu_{k|T,pq}^j = \sum_{i \in \mathcal{M}_{pq}} \bar{\mu}_{k|k,pq}^{ji} \mu_{k+1|T,pq}^i$
 - 20: **FIN POUR**
 - 21: Moyenne lissée globale $\hat{x}_{k|T} : \hat{x}_{k|T,p} = \sum_{j \in \mathcal{M}_p} \mu_{k|T,p}^j \hat{x}_{k|T,p}^j$
 - 22: Covariance globale de l'erreur de lissage $P_{k|T} : P_{k|T,pq} = \sum_{j \in \mathcal{M}_{pq}} \mu_{k|T,pq}^j \left[P_{k|T,pq}^j + (\hat{x}_{k|T,p}^j - \hat{x}_{k|T,p})(\hat{x}_{k|T,q}^j - \hat{x}_{k|T,q}) \right]$
-

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

Algorithme 7 INTERACTION HÉTÉROGÈNE AVEC M^2 COMBINAISONS

$\{\hat{x}_{k|T}^j, P_{k|T}^j\}_{j \in \mathcal{M}}$
 $= \text{INTERACTION-M2} \left(\left\{ \left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} \hat{\mathbf{x}}_{k|k}^{j \cup i}, \left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} \right\}_{j \in \mathcal{M}}, \left\{ \left(\mathbf{P}_{k|k+1}^{b, i \cup j} \right)^{-1} \hat{\mathbf{x}}_{k|k+1}^{b, i \cup j}, \left(\mathbf{P}_{k|k+1}^{b, i \cup j} \right)^{-1} \right\}_{i \in \mathcal{M}}, \left\{ \{\bar{\mu}_{k+1|T, pq}^{i|j}\}_{(i,j) \in \mathcal{M}^2} \right\} \right)$

1: **POUR** $i \in \mathcal{M}$ {1. Combinaison hétérogène}
2: **POUR** $j \in \mathcal{M}$
3: Moyenne lissée conditionnée sur deux modes : $\hat{\mathbf{x}}_{k|T}^{j \cup i} = \mathbf{P}_{k|T}^{j \cup i} \left[\left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} \hat{\mathbf{x}}_{k|k}^{j \cup i} + \left(\mathbf{P}_{k|k+1}^{b, i \cup j} \right)^{-1} \hat{\mathbf{x}}_{k|k+1}^{b, i \cup j} \right]$.
4: Covariance lissée conditionnée sur deux modes : $\mathbf{P}_{k|T}^{j \cup i} = \left[\left(\mathbf{P}_{k|k}^{j \cup i} \right)^{-1} + \left(\mathbf{P}_{k|k+1}^{b, i \cup j} \right)^{-1} \right]^{-1}$
5: **FIN POUR**
6: **FIN POUR**
7: **POUR** $j \in \mathcal{M}$ {2. Interaction hétérogène (équations classiques)}
8: Moyenne lissée conditionnée sur un mode $\hat{x}_{k|T}^j : \hat{x}_{k|T, p}^j = \sum_{i \in \mathcal{M}_p} \hat{\mathbf{x}}_{k|T, p}^{j \cup i} \bar{\mu}_{k+1|T, p}^{i|j}$.
9: Covariance lissée conditionnée sur un mode $P_{k|T}^j$:

$$P_{k|T, pq}^j = \sum_{i \in \mathcal{M}_{pq}} \bar{\mu}_{k+1|T, pq}^{i|j} \left[\mathbf{P}_{k|T, pq}^{j \cup i} + (\hat{\mathbf{x}}_{k|T, p}^{j \cup i} - \hat{x}_{k|T, p}^j)(\hat{\mathbf{x}}_{k|T, q}^{j \cup i} - \hat{x}_{k|T, q}^j) \right]$$
.
10: **FIN POUR**

Algorithme 8 INTERACTION HÉTÉROGÈNE AVEC M COMBINAISONS

$\{\hat{x}_{k|T}^j, P_{k|T}^j\}_{j \in \mathcal{M}}$
 $= \text{INTERACTION-M} \left(\left\{ \left(P_{k|k}^j \right)^{-1} \hat{x}_{k|k}^j, \left(P_{k|k}^j \right)^{-1} \right\}_{j \in \mathcal{M}}, \left\{ \hat{x}_{k|k+1}^{b, i}, P_{k|k+1}^{b, i} \right\}_{i \in \mathcal{M}}, \left\{ \{\bar{\mu}_{k+1|T, pq}^{i|j}\}_{(i,j) \in \mathcal{M}^2} \right\} \right)$

1: **POUR** $j \in \mathcal{M}$ {1. Interaction IMMHO arrière}
2: Moyenne arrière d'interaction $\bar{x}_{k|k+1}^{b, j} : \bar{x}_{k|k+1, p}^{b, j} = \sum_{i \in \mathcal{M}_p} \bar{\mu}_{k+1|T, p}^{i|j} \hat{x}_{k|k+1, p}^{b, i}$.
3: Covariance arrière d'interaction $\bar{P}_{k|k+1}^{b, j}$:

$$\bar{P}_{k|k+1, pq}^{b, j} = \sum_{i \in \mathcal{M}_{pq}} \bar{\mu}_{k+1|T, pq}^{i|j} \left[P_{k|k+1, pq}^{b, i} + (\hat{x}_{k|k+1, p}^{b, i} - \bar{x}_{k|k+1, p}^{b, j})(\hat{x}_{k|k+1, q}^{b, i} - \bar{x}_{k|k+1, q}^{b, j}) \right]$$
.
4: **FIN POUR**
5: **POUR** $j \in \mathcal{M}$ {2. Combinaison (équations classiques)}
6: Moyenne lissée conditionnée sur un mode : $\hat{x}_{k|T}^j = P_{k|T}^j \left[\left(\bar{P}_{k|k+1}^{b, j} \right)^{-1} \bar{x}_{k|k+1}^{b, j} + \left(P_{k|k}^j \right)^{-1} \hat{x}_{k|k}^j \right]$.
7: Covariance lissée conditionnée sur un mode : $P_{k|T}^j = \left[\left(\bar{P}_{k|k+1}^{b, j} \right)^{-1} + \left(P_{k|k}^j \right)^{-1} \right]^{-1}$.
8: **FIN POUR**

3.3.2. Lisseur IMMHO à double filtre

Alternativement à la solution basée sur une passe de lissage à rebours, il est possible de fusionner les estimées calculées par deux filtres IMMHO se propageant en sens contraire et initialisés de manière indépendante. Sans surprise, la récursion de filtrage arrière se réécrit :

1. $\forall i \in \mathcal{M}, \{P(m_{k+1}^i | z_{k+1:T})\}_{i \in \mathcal{M}} \xrightarrow{\text{Prédiction arrière}} P(m_k^i | z_{k+1:T})$
2. $\forall i \in \mathcal{M}, p(x_{k+1}^i | m_{k+1}^i, z_{k+1:T}) \xrightarrow{\text{Prédiction arrière}} p(x_k^i | m_{k+1}^i, z_{k+1:T})$
3. $\forall j \in \mathcal{M}, \{p(x_k^j | m_{k+1}^j, z_{k+1:T})\}_{j \in \mathcal{M}} \xrightarrow{\text{Interaction}} p(x_k^j | m_k^j, z_{k+1:T})$
4. $\forall j \in \mathcal{M}, p(x_k^j | m_k^j, z_{k+1:T}) \xrightarrow{\text{Correction}} p(x_k^j | m_k^j, z_{k:T})$
5. $\forall j \in \mathcal{M}, P(m_k^j | z_{k+1:T}) \xrightarrow{\text{Correction}} P(m_k^j | z_{k:T})$

On peut dériver des équations exactes d'un IMMHO à rebours avec les mêmes méthodes que l'IMMHO avant. À cet égard, nous ne donnerons pas plus de détails. Pour la construction de l'estimée lissée, souvenons-nous que la fusion des estimées des filtres avant et arrière peut se réaliser de deux façons, identiques à l'étape d'interaction du lisseur IMM-RTS.

Fusion hétérogène avec M^2 modes pour le lissage

On suppose les densités $p(x_k^j | m_k^j, z_{1:k})$ et $p(x_k^i | m_{k+1}^i, z_{k+1:T})$ données par les filtres avant et arrière respectivement. Les deux premiers moments de la densité lissée $p(x_k^j | m_k^j, z_{1:T})$ sont obtenus par application de l'algorithme 7.

Fusion homogène avec M modes pour le lissage

On fait l'hypothèse que l'on dispose des densités $p(x_k^j | m_k^j, z_{1:k})$ et $p(x_k^j | m_k^j, z_{k+1:T})$ calculées par les filtres avant et arrière. Les deux premiers moments de la densité lissée $p(x_k^j | m_k^j, z_{1:T})$ sont alors déduits de l'algorithme 8.

3.4. Application sur des données simulées

Cette dernière partie est consacrée à la mise en œuvre des algorithmes développés tout au long de ce chapitre sur trois exemples facilement reproductibles. On considère des processus dont on connaît les statistiques a priori. Le but est de caractériser les estimateurs au travers de l'analyse des quantités produites sur plusieurs réalisations. Les deux premiers exemples mettent à l'épreuve le lisseur IMM-RTS sur des problèmes de lissage sur intervalle fixe et à retard fixe. Le dernier exemple s'intéresse au filtrage et au lissage hétérogènes.

3.4.1. Lissage sur intervalle fixe

Un exemple simulé de mobile se déplaçant sur un plan est présenté afin d'analyser la qualité des estimées produites par le lisseur IMM-RTS. Dans le but de le comparer avec la solution de [Helmick et al., 1995], une dynamique inversible est considérée.

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

L'état du système est défini comme $x = (x, y, \dot{x}, \dot{y})^T$ où (x, y) sont les coordonnées cartésiennes du mobile et $(\dot{x} = \frac{dx}{dt}, \dot{y} = \frac{dy}{dt})$ son vecteur vitesse. Les modes admissibles du système sont deux marches aléatoires corrélées : une première avec une diffusion sur la vitesse $D_1 = 5^2 \text{ m}^2 \cdot \text{s}^{-3}$ (modèle 1 de *manœuvre*) et une autre avec une diffusion plus faible $D_2 = 0.5^2 \text{ m}^2 \cdot \text{s}^{-3}$ (modèle 2 de *vitesse quasi-constante*). Les équations d'état sont

$$x_k = \begin{pmatrix} 1 & 0 & \Delta t_k & 0 \\ 0 & 1 & 0 & \Delta t_k \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + q_{k-1}^j \quad (3.64)$$

avec la covariance du bruit de transition

$$Q_{k-1}^j = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2D_j\Delta t_k & 0 \\ 0 & 0 & 0 & 2D_j\Delta t_k \end{pmatrix} \text{ pour } j = 1, 2 \quad (3.65)$$

et $\Delta t_k = t_k - t_{k-1}$. Le vecteur de mesure z_k est la position bruitée du mobile relevée en coordonnées cartésiennes au temps t_k pour $k = 1, \dots, T$ avec une période de $\Delta t_k = 5\text{s}$. Ainsi, l'équation de mesure est commune à tous les modes et s'écrit :

$$z_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} x_k + r_k \text{ avec la covariance du bruit de mesure } R_k = 150^2 \mathbb{I}_2 \text{ m}^2. \quad (3.66)$$

La matrice des probabilités a priori des transitions entre modes est fixée à

$$\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \begin{pmatrix} 0,97 & 0,03 \\ 0,03 & 0,97 \end{pmatrix}. \quad (3.67)$$

Le mobile est suivi pendant 90 instants (soit 450 s) sur une trajectoire générée aléatoirement. L'algorithme a été évalué sur $s = 50$ trajectoires dont une réalisation est disponible dans la figure 3.1. Le mobile évolue d'abord selon le modèle de manœuvre 1, puis selon le modèle 2 de vitesse quasi-constante et revient dans le modèle 1 de manœuvre pour finir. Les instants de saut entre les modes se produisent de manière déterministe à $k = 30$ et $k = 60$. À l'instant initial $k = 0$, les modes sont supposés équiprobables et les positions et vitesses initiales de l'état continu x_0 sont arbitrairement fixées à 0 avec une covariance $P_{0|0} = \text{diag}([1 \ 1 \ 100 \ 100])$ pour tous les modes.

Le lisseur IMM est comparé aux filtres IMM et GPB₂ de [Blom and Bar-Shalom, 1988, Li and Jilkov, 2005], au GPB₂-RTS de [Mesot and Barber, 2009], au lisseur mi-IMM, mi-GPB₂ de [Koch, 2000] et au lisseur IMM à double filtre de [Helmick et al., 1995]. Ce dernier nécessite que le filtre IMM arrière soit initialisé sans connaissance a priori. Comme le propose [Helmick et al., 1995], l'estimée initiale de la position dans le filtre arrière à l'instant terminal T est affectée avec la mesure finale z_T , et la covariance de l'erreur associée est initialisée avec la covariance du bruit de mesure sur z_T . L'estimée de la vitesse finale $[\hat{x}_{T|T}^b, \hat{y}_{T|T}^b]^T$ est initialisée à 0 avec une matrice de covariance associée arbitrairement grande $10^6 \mathbb{I}_2 \text{ m}^2 \cdot \text{s}^{-2}$. Enfin, les modes sont supposés équiprobables au temps terminal.

On s'intéresse d'une part à la racine carrée des erreurs quadratiques moyennes (en anglais *Root*

3.4. APPLICATION SUR DES DONNÉES SIMULÉES

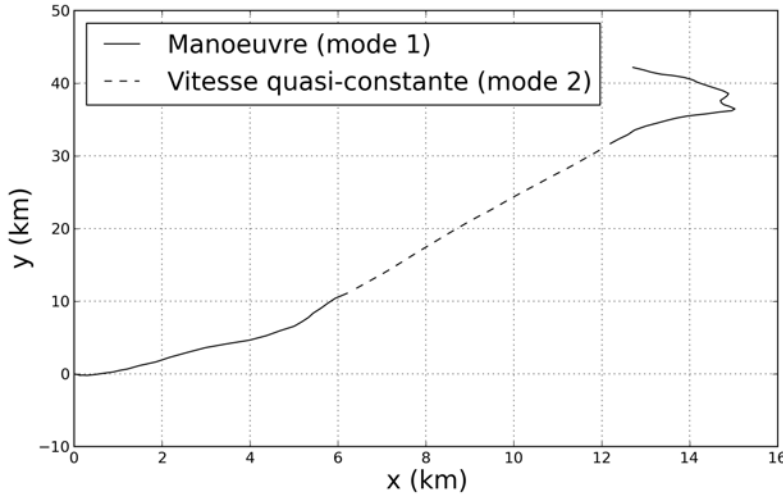


FIGURE 3.1.: Exemple de trajectoire générée aléatoirement avec des marches aléatoires corrélées.

Mean Square Error ou RMSE) entre les positions/vitesses réelles et leurs estimées :

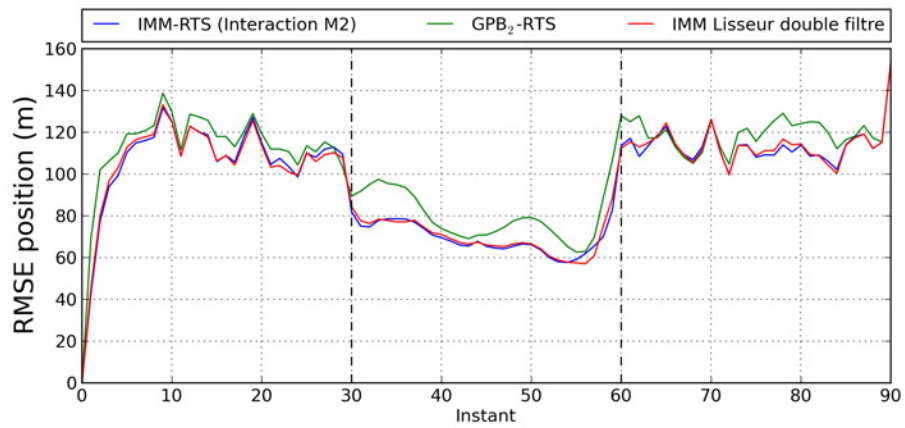
$$RMSE_{x,y}(k) = \sqrt{\frac{\sum_{i=1}^s \left\| \begin{bmatrix} (x_k^{(i)} - \hat{x}_k^{(i)})^T, (y_k^{(i)} - \hat{y}_k^{(i)})^T \end{bmatrix}^T \right\|^2}{s}} \quad (3.68)$$

$$RMSE_{\dot{x},\dot{y}}(k) = \sqrt{\frac{\sum_{i=1}^s \left\| \begin{bmatrix} (\dot{x}_k^{(i)} - \hat{\dot{x}}_k^{(i)})^T, (\dot{y}_k^{(i)} - \hat{\dot{y}}_k^{(i)})^T \end{bmatrix}^T \right\|^2}{s}} \quad (3.69)$$

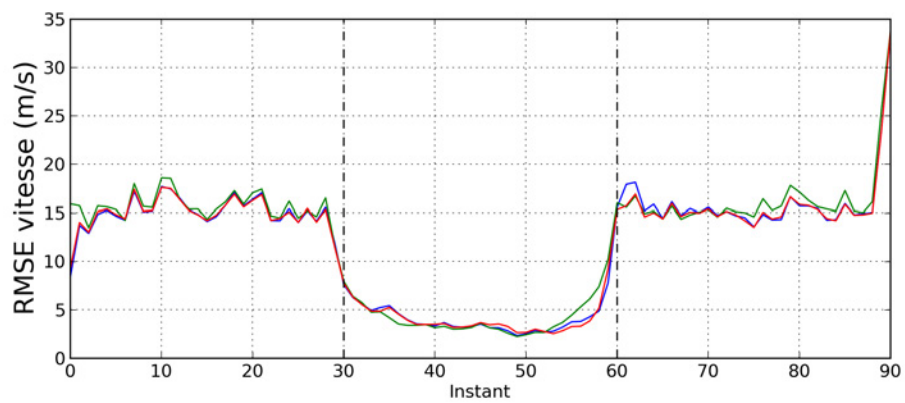
où l'exposant $^{(i)}$ référence les quantités relatives à la réalisation numéro i . D'autre part, on calcule la probabilité empirique de mauvaise détection $\tilde{P}(m_k \neq \hat{j}_{k|T})$, *i.e.* la probabilité de ne pas détecter le mode réellement actif au moyen de (3.32). Dans le tableau 3.5, les RMSE de la position et de la vitesse moyennées sur le temps sont présentées avec les probabilités de mauvaise détection moyennées sur le temps. Ces quantités sont aussi tracées en fonction du temps dans Fig. 3.2(a), Fig. 3.2(b) et Fig. 3.2(c) (seulement pour le lisseur IMM-RTS avec l'interaction M^2 , le GPB₂-RTS et le lisseur IMM double filtre).

Les résultats montrent que le lisseur IMM-RTS se comporte bien, quel que soit le type d'interaction choisi, avec une réduction significative des RMSE de l'état continu en comparaison des solutions à base de filtrage. La valeur moyenne des RMSE du lisseur double filtre et celle du lisseur IMM-RTS se confondent. Pour l'état discret, la détection du mode actif est aussi bien plus efficace. L'IMM-RTS est un peu meilleur avec l'utilisation de la formule (2.42) exploitant les propriétés markoviennes de la séquence de modes pour le calcul des probabilités lissées. Les lisseurs de [Barber, 2005] et [Koch, 2000] affichent en revanche des performances nettement plus faibles à la fois pour les RMSE et les probabilités de mauvaise détection. Les raisons sont à rechercher du côté des approximations sur l'état continu pour établir les équations de lissage et sur le calcul des probabilités de l'état discret. Le fait d'approximer certaines densités par un Dirac coûte cher en termes de performances !

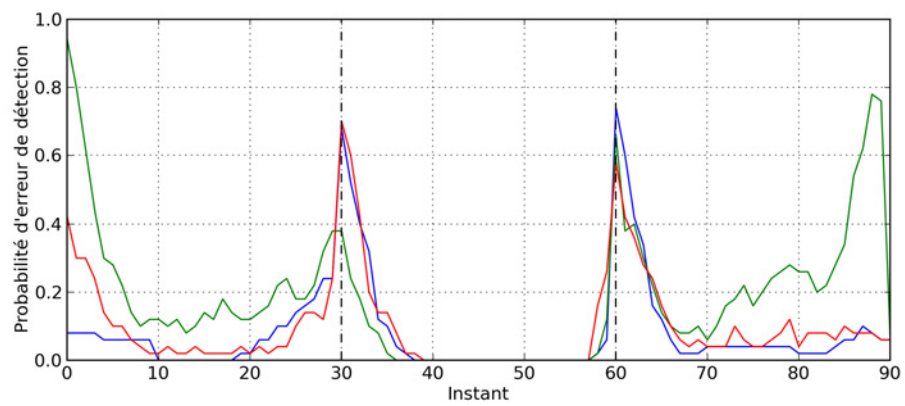
CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES



(a)



(b)



(c)

FIGURE 3.2.: Comparaison du filtre IMM et des lisseurs (de haut en bas) : (a) RMSE sur la position, (b) RMSE sur la vitesse et (c) probabilité de mauvaise détection du mode actif.

3.4. APPLICATION SUR DES DONNÉES SIMULÉES

TABLE 3.3.: Valeurs moyennées sur le temps de la RMSE sur la position (m), la RMSE sur la vitesse (m.s⁻¹) et la probabilité de mauvaise détection.

Méthode	Position	Vitesse	Mauvaise détection
IMM	156,2	24,7	0,23
GPB ₂	155,7	24,7	0,22
IMM-RTS (Interaction M^2)	96,1	11,8	0,08
IMM-RTS (Interaction M)	96,4	11,8	0,08
IMM Lisseur double filtre	96,4	11,7	0,10
GPB ₂ -RTS	104,0	12,2	0,19
Koch-RTS	104,7	12,2	0,18

3.4.2. Lissage à retard fixe

L'évaluation du lisseur multi-modèle est complétée par un exemple de lissage à retard fixe sur un système dynamique non linéaire. Nous présentons une comparaison avec l'IMM à état augmenté de [Morelande and Ristic, 2008].

Deux modèles de mouvement coordonné (en anglais *Coordinated Turn* ou CT) sont considérés avec des vitesses angulaires ω_j inconnues (pour $j = 1, 2$). Le mobile évolue toujours sur le plan et le vecteur d'état est maintenant $x = (x, \dot{x}, y, \dot{y}, \omega)^T$. La dynamique du modèle CT [Rong Li and Bar-Shalom, 1993, Li and Jilkov, 2003] est donnée par

$$x_k = \begin{pmatrix} 1 & \frac{\sin(\omega_{k-1}\Delta t_k)}{\omega_{k-1}} & 0 & -\frac{1-\cos(\omega_{k-1}\Delta t_k)}{\omega_{k-1}} & 0 \\ 0 & \cos(\omega_{k-1}\Delta t_k) & 0 & -\sin(\omega_{k-1}\Delta t_k) & 0 \\ 0 & \frac{1-\cos(\omega_{k-1}\Delta t_k)}{\omega_{k-1}} & 1 & \frac{\sin(\omega_{k-1}\Delta t_k)}{\omega_{k-1}} & 0 \\ 0 & \sin(\omega_{k-1}\Delta t_k) & 0 & \cos(\omega_{k-1}\Delta t_k) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} x_{k-1} + q_{k-1}^j \quad (3.70)$$

avec la covariance du bruit de transition $Q_{k-1}^j = \text{diag}([\sigma_{x,j}^2 M, \sigma_{y,j}^2 M, \sigma_{\omega,j}^2])$ et $M = \begin{pmatrix} \Delta t_k^4/4 & \Delta t_k^3/2 \\ \Delta t_k^3/2 & \Delta t_k^2 \end{pmatrix}$. Pour le premier mode, on fixe $\sigma_{x,1}^2 = 1$, $\sigma_{y,1}^2 = 1$, $\sigma_{\omega,1}^2 = 0$ (modèle 1 de *vitesse angulaire constante et vitesse linéaire quasi-constante*) et, pour le deuxième mode, $\sigma_{x,2}^2 = 100$, $\sigma_{y,2}^2 = 100$, $\sigma_{\omega,2}^2 = 1.75 \times 10^{-3}$ (modèle 2 de *manœuvre*). L'équation de mesure est commune à tous les modes et s'écrit :

$$z_k = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} x_k + r_k \text{ avec la covariance du bruit de mesure } R_k = 10^2 \mathbb{I}_2 \text{ m}^2. \quad (3.71)$$

La matrice des probabilités de transition entre les modes est la même que (3.67).

Comme précédemment, la trajectoire d'état cachée et les mesures sont générées aléatoirement. Les comparaisons sont menées sur 50 trajectoires dont la Fig. 3.3 en présente une réalisation. Le mobile est observé pendant 90 instants avec $\Delta t_k = 1 \text{ s } \forall k$. La valeur moyenne de l'état initial est pour tous les modes $\hat{x}_{0|0} = [0 \ 30 \ 0 \ 30 \ 0,05]^T$ avec la covariance $P_{0|0} = \text{diag}([10 \ 1 \ 10 \ 1 \ 0,001])$. Les instants des sauts sont $k = 30$ et $k = 60$. Le modèle de manœuvre est actif entre $k = 30$ et $k = 60$. Le modèle de vitesse quasi-constante et de vitesse angulaire constante est actif avant $k = 30$ et après $k = 60$. Les RMSE moyennées sur la durée de simulation sont tracées en fonction

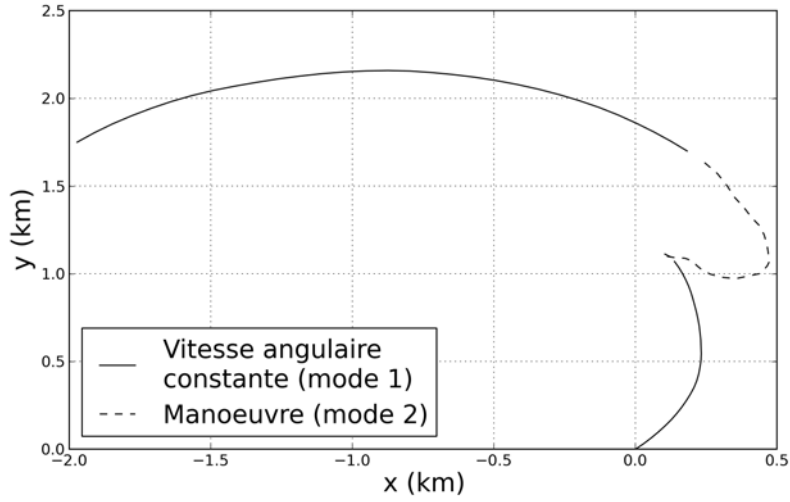


FIGURE 3.3.: Exemple de trajectoire générée aléatoirement avec des mouvements coordonnés.

du retard de lissage pour la position, la vitesse et la vitesse angulaire sur Fig. 3.4(a), Fig. 3.4(b) et Fig. 3.4(c) respectivement. La probabilité empirique de mauvaise détection moyennée sur le temps est affichée sur Fig. 3.4(d).

Au regard des valeurs des RMSE, les lisseurs IMM-RTS et IMM augmenté présentent des performances similaires quelle que soit la valeur du retard. Néanmoins, les probabilités de mauvaise détection sont légèrement plus grandes avec le lisseur IMM augmenté. Comme dans l'exemple précédent, le calcul direct des probabilités lissées (2.42) est plus fiable. L'IMM-RTS se comporte donc aussi bien que la solution alternative avec état augmenté.

3.4.3. Filtrage et lissage hétérogènes

L'objectif de ce paragraphe est d'illustrer le comportement du filtre et du lisseur IMMHO sur un système évoluant selon plusieurs variantes de marches aléatoires. Le mobile est supposé suivre trois modes $\mathcal{M} = \{1, 2, 3\}$ dont une marche aléatoire corrélée (MAC) et deux marches aléatoires (MA) de diffusions D différentes. Les paramètres sont donnés dans le tableau 3.4. Les vecteurs d'état continus correspondants sont x^1, x^2, x^3 , avec $x^1 = (x_1, x_2)^T$, $x^2 = (x_1, x_2)^T$, $x^3 = (x_1, x_2, \dot{x}_1, \dot{x}_2)^T$. Ainsi, le vecteur d'état continu complet est $x = (x_1, x_2, x_3, x_4)^T$. Avec les notations du paragraphe 3.2, il satisfait à $x_1 = x_1^1 = x_1^2 = x_1^3$, $x_2 = x_2^1 = x_2^2 = x_2^3$, $x_3 = x_3^3 = \dot{x}_1$ et $x_4 = x_4^3 = \dot{x}_2$. L'équation de mesure pour la marche aléatoire corrélée s'écrit

$$z_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} x_k^3 + r_k, \quad (3.72)$$

et pour les marches aléatoires simples,

$$z_k = x_k^j + r_k, \quad j = 1, 2. \quad (3.73)$$

Pour tous les modes, la covariance du bruit de mesure est $R_k = 150^2 \mathbb{I}_2 \text{ m}^2$.

3.4. APPLICATION SUR DES DONNÉES SIMULÉES

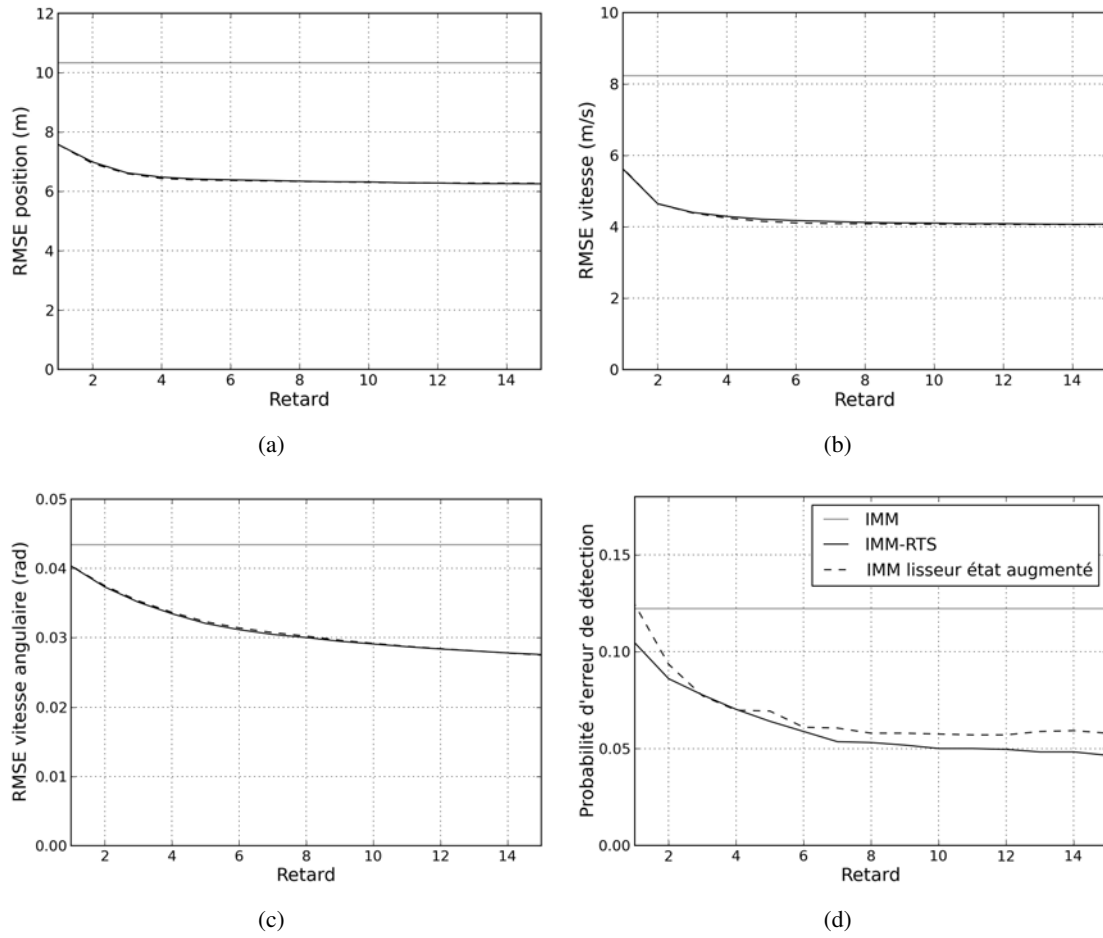


FIGURE 3.4.: Comparaison des RMSE moyennées sur le temps du filtre IMM et des lisseurs à retard fixe : (a) position, (b) vitesse, (c) vitesse angulaire et (d) probabilité de mauvaise détection du mode actif.

TABLE 3.4.: Paramètres de modèles

#	Modèle	D
1	Marche aléatoire 1 (MA1)	1
2	Marche aléatoire 2 (MA2)	64
3	Marche aléatoire corrélée (MAC)	0.09

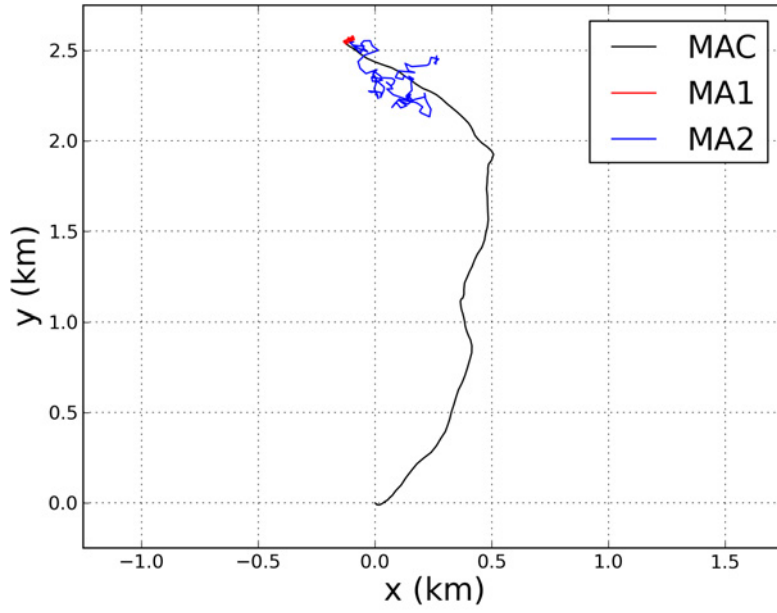


FIGURE 3.5.: Exemple de trajectoire aléatoirement générée avec des mouvements coordonnés

Les valeurs moyennes de l'état initial sont fixées à 0 pour chaque mode et les modes sont équiprobables à l'instant initial, *i.e.*

$$\mu_{0|0,pq}^i = \frac{1}{M_{pq}} \text{ si } i \in \mathcal{M}_{pq}, 0 \text{ sinon.} \quad (3.74)$$

Le nombre d'observations est fixé à 300 avec $\Delta t_k = 5s, \forall k$. Les instants des sauts sont fixés à $k = 100$ et $k = 200$. Les probabilités de saut sont fixées à $\pi_{ii} = 0,97$ et $\pi_{ji} = 0,015$ si $i \neq j$ ce qui est consistant avec le nombre de changements de mode. La probabilité de transition $\pi_{ji,pq}$ du modèle $j \in \mathcal{M}_{pq}$ au modèle $i \in \mathcal{M}_{pq}$ est obtenue avec l'heuristique suivante :

$$\pi_{ji,pq} = \frac{\pi_{ji}}{\sum_{l \in \mathcal{M}_{pq}} \pi_{li}} \text{ si } (i, j) \in \mathcal{M}_{pq} \times \mathcal{M}_{pq}, \pi_{ji,pq} = 0 \text{ sinon.} \quad (3.75)$$

Un exemple de trajectoire est donné en Fig. 3.5. Sur un total de 50 simulations, on analyse une nouvelle fois les RMSE de la position et de la vitesse (respectivement dans les figures 3.6(a) et 3.6(b)). Les probabilités empiriques de mauvaise détection du modèle actif sont fournies dans la Fig. 3.6(c). Lorsque les marches aléatoires sont actives à l'instant k , la vitesse n'est pas estimée par le modèle et on procède alors à une approximation par différences finies entre les deux instants $k - 1$ et k .

La détection du mode actif est globalement efficace, hormis autour des sauts de modes. Ce comportement est tempéré après une passe de lissage double filtre (les probabilités empiriques de mauvaise détection passent de 0,31 à 0,17 en moyenne). On observe par contre des probabilités d'erreur de détection plus élevées lorsque la marche aléatoire de grande diffusion est active. Dans ces conditions, le filtre IMMHO ne fait pas efficacement la distinction entre la MAC et la MA de grande diffusion : les modèles ne sont pas suffisamment différents pour être discriminés et leurs

3.4. APPLICATION SUR DES DONNÉES SIMULÉES

TABLE 3.5.: Valeurs moyennées sur le temps de la RMSE sur la position (m), la RMSE sur la vitesse (m.s⁻¹) et la probabilité de mauvaise détection.

Méthode	Position	Vitesse	Mauvaise détection
IMMHO	80,1	4,8	0,31
IMMHO-RTS (Interaction M2)	?	?	0,17
IMMHO-RTS (Interaction M)	?	?	0,17
IMMHO Lisseur double filtre (Interaction M2)	50,2	3,5	0,17
IMMHO Lisseur double filtre (Interaction M)	50,2	3,5	0,17

probabilités a posteriori respectives sont proches. Ces notions de distance entre les modèles sont approfondies dans [Li and Jilkov, 2003]. Du côté des erreurs sur la position, leurs valeurs sont drastiquement réduites par les lisseurs à double filtre (les erreurs moyennées passent de 80,1 à 50,2 m). Sur la vitesse, on retrouve une amélioration importante liée au lissage (de 4,8 à 3,5 m/s d'erreur moyenne). Enfin, lorsque la marche aléatoire de grande diffusion est active, le lissage améliore peu l'estimation de la vitesse. Pour les versions RTS, l'estimation est instable et diverge après quelques pas de la passe à rebours. L'explication se situe sans doute dans l'évaluation des probabilités d'interaction. N'ayant pas pu mettre en évidence d'autres modes de calcul, la version hétérogène applique directement (3.27) pour l'évaluation des probabilités d'interaction alors que la version homogène du lisseur propose une alternative utilisant l'état continu. La bonne évaluation de ces probabilités d'interaction reste donc un sujet ouvert dans le cas hétérogène pour parvenir à une mise en œuvre stable de l'IMMHO-RTS.

CHAPITRE 3. NOUVEAUX OUTILS POUR LE FILTRAGE ET LE LISSAGE MULTI-MODÈLES

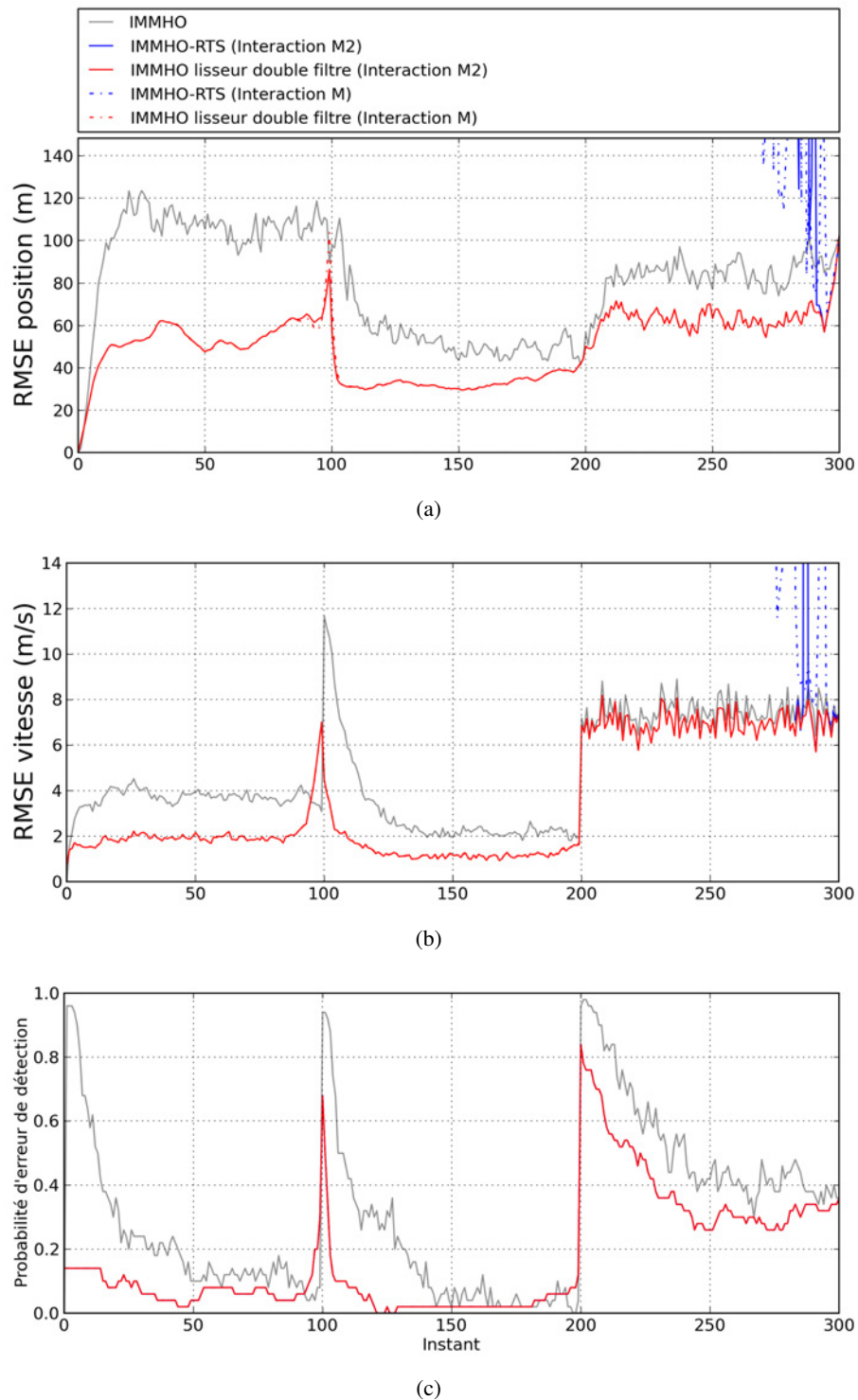


FIGURE 3.6.: Comparaison du filtre IMMHO et des lisseurs (de haut en bas) : (a) RMSE sur la position, (b) RMSE sur la vitesse et (c) probabilité empirique de mauvaise détection du mode actif (toutes les courbes des lisseurs sont confondues).

Conclusions

Les contributions de ce chapitre sont d'ordre méthodologique avec, en premier lieu, l'introduction d'un nouvel outil de lissage multi-modèle. L'algorithme recombine dans une passe à rebours les quantités intermédiaires et finales d'un IMM pour former une approximation de la densité lissée. Cette stratégie fait intervenir M lisseurs pour M modèles impliqués et la passe à rebours s'amorce avec les statistiques produites par un filtre IMM à l'instant terminal. Au delà de l'application à ARGOS, la méthode supporte aussi les dynamiques non inversibles et s'applique au lissage à retard fixe. Une fois comparée aux stratégies existantes, elle s'est montrée au moins aussi performante.

Dans un second temps, nous avons étendu le cadre multi-modèle initial aux ensembles de modèles hétérogènes. Le filtre IMMHO qui en résulte fonctionne bien sur des simulations et est plus général que la proposition de [Yuan et al., 2012] qui se limite en effet à deux modèles prédéfinis (un modèle balistique et un modèle de poussée). Nous avons par ailleurs montré qu'il est aussi possible de construire des lisseurs hétérogènes (double filtre et RTS). Cependant, la formulation basée sur une passe à rebours RTS n'est pas stable en pratique à l'opposé de l'implémentation avec deux filtres. Cela réduit donc le champ d'application de l'IMMHO à des cas ARGOS où l'initialisation d'un filtre IMMHO à rebours n'est pas problématique.

MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS



Marabout et oies.

Sommaire

4.1. Choix techniques	96
4.1.1. Algorithmes et ensemble de modèles	96
4.1.2. Particularités d'implémentation	97
4.1.3. Aspects logiciels	98
4.2. Service temps-réel avec un filtre multi-modèle	98
4.2.1. Un plus grand nombre de positions	98
4.2.2. De meilleures performances dans les cas difficiles	100
4.2.3. Caractérisation de l'erreur	101
4.3. Service à valeur ajoutée sur les données archives	103
4.3.1. Un effet régularisant sur l'erreur	103
4.3.2. Caractérisation de l'erreur	104
Conclusions	105

Cette dernière partie est dédiée à la mise en œuvre des techniques multi-modèles pour le suivi de mobiles ARGOS. L'objectif est de proposer un traitement uniforme pour toutes les balises du système aussi bien sur le plan des algorithmes que des modèles sélectionnés.

4.1. Choix techniques

4.1.1. Algorithmes et ensemble de modèles

Pour l'application à ARGOS, nous faisons le choix d'utiliser un filtre IMM classique et un lisseur IMM-RTS. Nous avons jugé que les techniques hétérogènes n'étaient pas suffisamment abouties pour être proposées dans le nouveau traitement de localisation. Comme la récursion de type RTS est instable, le lissage hétérogène impose d'utiliser une formulation double filtre. Elle nécessite d'initialiser le filtre arrière avec un ajustement par moindres carrés sur la base de quatre mesures d'un passage satellite. Pour des balises évoluant en conditions difficiles (milieu marin, forêt dense...) ou en fin de vie (baisse de la puissance d'émission), le nombre de messages collectés par passage est en général faible et le moment d'initialisation du filtre arrière est incertain.

L'ensemble de modèles se compose d'une marche aléatoire et d'une marche aléatoire biaisée partageant le vecteur d'état $x = (\lambda, \phi, f_t)^T$. Les marches sont isotropes en longitude et latitude. Les équations d'état de la marche aléatoire (cf. paragraphe 2.1.3) sont

$$x_k = x_{k-1} + q_{k-1},$$

$$\text{où } \bar{q}_{k-1} = 0, Q_{k-1} = \begin{pmatrix} 2D_\lambda \Delta t_k & 0 & 0 \\ 0 & 2D_\phi \Delta t_k & 0 \\ 0 & 0 & V_f \end{pmatrix}$$

et les équations de la marche aléatoire biaisée s'écrivent

$$x_k = x_{k-1} + \begin{pmatrix} \frac{\Delta t_k}{d_e \cos(\phi_{k-1})} & 0 \\ 0 & \frac{\Delta t_k}{d_e} \\ 0 & 0 \end{pmatrix} \hat{v}_{k|k-1} + q_{k-1},$$

$$\text{où } \bar{q}_{k-1} = 0, Q_{k-1} = \begin{pmatrix} 2D_\lambda \Delta t_k & 0 & 0 \\ 0 & 2D_\phi \Delta t_k & 0 \\ 0 & 0 & V_f \end{pmatrix}.$$

Pour ce dernier modèle, la vitesse est approximée via une moyenne mobile exponentielle de paramètre $\alpha \in [0, 1]$:

$$\hat{v}_{k+1|k} = \alpha \tilde{v}_k + (1 - \alpha) \hat{v}_{k|k-1} \quad (4.1)$$

avec $\tilde{v}_k$ la vitesse évaluée par différence finie

$$\tilde{v}_k = \frac{(\lambda_k, \phi_k)^T - (\lambda_{k-1}, \phi_{k-1})^T}{\Delta t_k}. \quad (4.2)$$

Ce type d'approximation permet de tempérer le comportement erratique d'un calcul approché par différences finies généralement observé sur les données ARGOS brutes. Le paramètre α est fixé à 0,3, de sorte que la pondération des cinq plus récentes observations de vitesse dans la moyenne

représente 86% de la pondération totale. Le paramètre de diffusion D du bruit de dynamique est affiné par recherche incrémentale et sa valeur est identique pour les deux variantes de marche aléatoire. La covariance du bruit de dynamique sur la fréquence d'émission V_f est fixée à 100 Hz² pour tenir compte des variations de température les plus communes. Pour tous les modes, l'équation de mesure Doppler est applicable :

$$z_k = g_k(x_k) + r_k.$$

Le bruit de mesure est non biaisé et sa covariance associée est fixée à 0,4 Hz (dispersion de la mesure de fréquence au niveau du satellite). Les transitions entre les modes sont supposées équiprobables car nous n'avons aucune certitude sur les comportements sous-jacents des plateformes en termes de fréquences de transition. Dans l'IMM, cela se traduit par des probabilités prédites égales les unes aux autres de sorte que les modèles sont pondérés en fin de cycle par les vraisemblances de leurs filtres respectifs. Enfin, nous avons retenu comme filtre une implémentation en racine carrée du filtre de Kalman Unscented pour prévenir les instabilités numériques [Van Der Merwe and Wan, 2001].

4.1.2. Particularités d'implémentation

Initialisation

Le filtre IMM est un procédé récursif initialisé par des estimées initiales $\hat{x}_{0|0}^i$ et leurs covariances $P_{0|0}^i$ que l'on choisit identiques quel que soit le mode $i \in \mathcal{M}$. Pour les calculer, un ajustement par moindres carrés est mené avec au moins quatre mesures et produit toujours une localisation image et nominale qu'il est impossible de distinguer à l'initialisation. En effet, il n'y a pas de localisation connue pour appliquer les tests de vraisemblance et choisir une des deux solutions. L'ambiguïté est résolue en appliquant au passage suivant une passe de filtrage sur les deux positions prises séparément. La position nominale est celle maximisant la vraisemblance globale L_k du filtre IMM lors de l'assimilation des mesures du second passage avec $L_k = \sum_{j \in \mathcal{M}} p(z_k | m_k^j, z_{1:k-1}) P(m_k^j | z_{1:k-1})$ [Bar-Shalom et al., 2001, Bar-Shalom and Chen, 2005].

Tests de vraisemblance

En sortie du filtre, les positions sont testées avant leur distribution en reprenant les tests de vraisemblance liés à la vitesse maximale du mobile et à la continuité en fréquence de l'oscillateur. Si les deux tests sont corrects alors la localisation à base de filtrage est distribuée. Si l'un des deux tests échoue alors on tente une réinitialisation avec un ajustement par moindres carrés. On délivre les deux solutions calculées si l'une d'entre elles réussit au moins deux tests de vraisemblance des moindres carrés. Ces solutions serviront de point de départ pour le filtre au prochain passage. Enfin, lorsque les tests de vraisemblance des moindres carrés échouent, la localisation associée au passage est marquée "Z" et mise de côté.

Outre le fait de se protéger des mesures aberrantes¹ dans le passage, ce mécanisme permet de détecter des cas limites en réinitialisant le processus à la volée par moindres carrés : changement de

1. Elles sont parfois provoquées par des balises "fantômes" qui émettent avec le même numéro qu'une balise du système et génèrent des fréquences anormales sur la courbe Doppler.

fréquence de l'oscillateur (à cause d'une reprogrammation de la balise), erreur dans la résolution d'ambiguïté ou encore redéploiement de la plateforme dans une autre région du monde.

Classes de localisation

Le système de classes de localisation sur lequel s'appuie le système de distribution des données est légèrement modifié. La classe B inclut les localisations calculées avec un ou deux messages. Les estimations d'erreur des classes A et B sont maintenant distribuées sous la forme d'une ellipse d'erreur et d'un rayon équivalent. La classe Z signifie maintenant que les mesures du passage n'ont pas donné de localisation jugée fiable que ce soit avec le filtre multi-modèle ou l'ajustement moindres carrés.

4.1.3. Aspects logiciels

Tous les programmes (cœur de calcul, outils de retraitement et de validation) sont développés en Java, un langage orienté objet à typage statique fort. Son principal avantage réside dans sa capacité à être exécuté sur n'importe quelle plateforme (Windows, Unix...) équipée d'une machine virtuelle Java. En comparaison du FORTRAN dans lequel est écrit le traitement moindres carrés, Java offre une gestion avancée des exceptions et de la mémoire. Il bénéficie par ailleurs d'un écosystème riche et actif : bibliothèques nombreuses (dont Apache Math pour les calculs), facilité dans la gestion des dépendances logicielles et dans le déploiement des applications (outils Maven et Ant). À ce titre, la direction technique de CLS favorise l'utilisation de Java qui est le langage principal de la chaîne de traitement ARGOS. On peut cependant reprocher au Java d'être moins adapté aux traitements scientifiques que le FORTRAN et de se révéler plus prolixe que Python (avec ses puissantes bibliothèques mathématiques Numpy/Scipy/Matplotlib).

La réécriture du logiciel en Java a en outre permis d'améliorer sa robustesse avec par exemple la possibilité de continuer les traitements même en cas d'erreur de calcul sur un passage satellite et de notifier plus précisément les opérateurs sur la nature de l'erreur. Le traitement de la localisation est en effet lancé sur un fichier de données contenant souvent plusieurs milliers de passages et il est dommageable pour les utilisateurs de ne pouvoir fournir aucune localisation à cause d'un problème sur un passage du fichier. La chaîne de localisation fonctionne en continu et, idéalement, ne doit pas se bloquer plus que quelques heures en cas d'anomalie.

4.2. Service temps-réel avec un filtre multi-modèle

On se propose en premier lieu, d'analyser les performances du traitement temps réel basé sur un filtre IMM.

4.2.1. Un plus grand nombre de positions

Les tableaux 4.1 et 4.2 présentent respectivement le nombre de positions estimées avec succès par le filtre multi-modèle ainsi que le nombre de positions retenues pour le calcul des statistiques d'er-

CHAPITRE 4. MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS

TABLE 4.1.: Nombre de localisations calculées par l'ajustement moindres carrés et le filtre multi-modèle

Type de mobile	Moindres carrés	Filtre
Marabout	3 472	3 759
Oie	16 097	21 220
Gnou	2 029	2 180
Mouflon	1 819	2 159
Tortue à dos plat	16 879	24 205
Tortue verte	8 690	15 959
Lion de mer des Galapagos	1 224	1 680
Éléphant de mer	6 305	13 120
Bateau	20 160	23 404
Bouée dérivante	48 733	54 817
Total	125 408	162 503

TABLE 4.2.: Nombre de localisations du filtre multi-modèle sélectionnées pour le calcul d'erreur

Type de mobile	4 mess. ou plus	2 ou 3 mess.	1 mess.	Total	% de localisations de Tab. 4.1
Marabout	381	53	22	456	12%
Oie	3 155	1 331	550	5 036	24%
Gnou	415	38	10	463	21%
Mouflon	323	94	23	440	20%
Tortue à dos plat	1 025	2 557	1 316	4 898	20%
Tortue verte	841	1 584	1 145	3 570	22%
Lion de mer	212	182	101	495	29%
Éléphant de mer	482	3 336	3 420	7 238	55%
Bateau	12 381	5 314	1 924	19 619	84%
Bouée	30 381	11 779	4 170	46 330	85%
Total	49 596	26 268	12 681	88 545	54%

reur empirique. On utilise les mêmes critères que l'étude menée dans le tout premier chapitre pour la sélection des localisations. Le tableau 4.1 montre que le filtre multi-modèle produit environ 30% de localisations supplémentaires par rapport aux moindres carrés. Cependant, cette augmentation n'est pas distribuée de manière homogène sur l'ensemble des jeux de données. La figure 4.1 (en fin de chapitre) illustre la répartition des localisations en fonction du nombre de messages reçus.

Le nombre de positions calculées pour les éléphants de mer est plus que doublé avec le filtre. Ce nombre s'est aussi largement accru pour les tortues vertes (+83%), les tortues à dos plat (+43%), les lions de mer (+37%) et les oies (+32%). Le gain est plus limité (en deçà de 20%) pour les autres types de mobiles. La plupart (88 %) des nouvelles positions du filtre correspondent à des passages à un message, le reste étant d'anciennes positions moindres carrés classées "Z" (non

convergence) et qui sont maintenant estimées correctement par le filtre. La capacité d'estimer une position avec un unique message bénéficie aux animaux marins pour lesquels le nombre de messages transmis est généralement faible pendant les périodes où ils font surface. Cette capacité est également bienvenue pour les oies qui migrent entre la Mongolie et l'Inde, une région où le bruit électromagnétique est important dans la bande de fréquences ARGOS et où la quantité de messages reçus est donc moindre. Reste maintenant à évaluer leur précision.

4.2.2. De meilleures performances dans les cas difficiles

La figure 4.2 compare la moyenne et l'écart-type de l'erreur empirique de localisation entre les moindres carrés et le filtre pour un minimum de quatre messages reçus. La réduction est assez marquée pour les éléphants de mer, les tortues à dos plat et les bateaux (jusqu'à 32% de gain). Pour les autres mobiles, l'amélioration est modeste. Ce comportement était attendu du fait du grand nombre de messages moyen et de la grande généralité des modèles impliqués. L'information portée par les mesures de fréquence domine largement dans la phase de correction du filtre de Kalman. Le filtre a par contre un impact plus fort sur l'écart-type des erreurs avec jusqu'à 47 % de réduction.

Pour les localisations à deux ou trois messages, les résultats sont donnés dans la figure 4.3. Les bénéfices d'un filtrage sont évidents. L'erreur moyenne sur la position se contracte entre 30% (bouées) et 78% (éléphants de mer). Les écarts-types sont réduits jusqu'à 90%. De manière générale, une amélioration est attendue lorsque le nombre de mesures du passage est limité (géométrie d'observation dégradée) et la dynamique de mouvement apporte une information déterminante.

Pour compléter cette analyse, nous avons tracé dans la figure 4.4 les erreurs de localisation empiriques moyennes en fonction de la distance à la trace pour les passages comportant au moins deux mesures. Nous nous sommes restreints aux bouées pour lesquelles le nombre de localisations est important et permet de calculer des statistiques plus fiables. On observe qu'entre 5 et 15 degrés, le gain sur l'erreur moyenne est globalement constant. C'est la zone de creux des erreurs théoriques tracées dans Fig. 1.6. En dehors, là où les erreurs théoriques s'accroissent plus rapidement, le gain du filtre est nettement plus marqué : il est plus résistant aux passages sous trace et à l'horizon où les géométries d'observation sont défavorables à cause d'une perte d'observabilité sur la longitude et la latitude respectivement.

Pour terminer, les localisations issues des passages à un message sont comparées avec celles des moindres carrés issues des passages avec deux ou trois messages dans la figure 4.5. Notons que les statistiques des balises de gnous, mouflons et marabouts sont évaluées avec très peu de localisations (cf. Tab. 4.2). Avec un seul message, les erreurs moyennes et les écarts-types sont inférieurs à 10 km ce qui est globalement meilleur que les moindres carrés avec un nombre plus important de messages. De par leur quantité et leur qualité, eu égard au peu d'information qu'une seule mesure de fréquence véhicule, ces localisations ont toute leur place pour une utilisation destinée au suivi d'animaux.

Reprenons maintenant l'éléphant de mer du premier chapitre dont la trajectoire complète est tracée dans Fig. 4.6 et par extraits dans Fig. 4.7. Ce type de plateforme tire grandement parti du nouveau traitement : l'erreur globale est réduite de 12 km à 4,5 km (-63%), les localisations à un message ajoutent de nombreux points à la trajectoire avec une précision inférieure à 10 km et une

CHAPITRE 4. MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS

caractérisation de l'erreur est toujours disponible. Par ailleurs, les localisations images ont disparu.

4.2.3. Caractérisation de l'erreur

Erreurs tangentielles

L'idée est ici de détecter un possible biais (ou retard) le long de la trajectoire causé par l'emploi d'un modèle de marche aléatoire. Nous proposons de calculer les statistiques de l'erreur empirique projetée le long de la trajectoire. Dans la figure 1.9 du chapitre 1, ces statistiques sont déduites de l'évaluation du module de $\vec{T}$ qui constitue la projection de $\vec{D}$ sur l'axe tangent à la trajectoire GPS. Le signe de l'erreur tangentielle est positif si la direction de $\vec{T}$ est en accord avec la direction de la trajectoire, et négatif sinon. Les valeurs d'erreurs tangentielles en deçà du 25^{ème} permille et au dessus du 975^{ème} permille sont retirées. Le rapport de la moyenne des erreurs tangentielles sur la moyenne du module de l'erreur est donné dans Tab. 4.3. Une valeur négative peut indiquer un retard systématique et une valeur positive, une avance.

TABLE 4.3.: Rapport de l'erreur tangentielle moyenne sur le module moyen de l'erreur.

Type de mobile	4 mess. ou plus		2 et 3 mess.		1 mess.
	MC	Filtre	MC	Filtre	Filtre
Marabout	5%	1%	5%	12%	38%
Oie	2%	4%	0%	-2%	-8%
Gnou	-1%	3%	-45%	-16%	-25%
Mouflon	-1%	-2%	19%	7%	-19%
Tortue à dos plat	-5%	-13%	1%	-9%	-14%
Tortue verte	10%	5%	8%	-15%	-22%
Lion de mer	9%	11%	13%	-7%	-2%
Éléphant de mer	7%	-2%	-3%	-8%	-18%
Bateau	1%	-1%	1%	-10%	-10%
Bouée	0%	-14%	2%	-38%	-60%

Avec un, deux ou trois messages, le rapport de l'erreur tangentielle sur l'erreur totale est négativement signé et globalement assez limité avec le filtre multi-modèle. Le signe négatif s'explique sachant qu'avec peu de messages, la phase de correction du filtre de Kalman est moins prépondérante dans l'estimation. Dans cette situation et quand la marche aléatoire est la dynamique la plus probable a posteriori, la localisation estimée est "attirée" par la dernière calculée. Le cas extrême serait un calcul sans aucun message où le filtre se base uniquement sur la prédiction. Avec les bouées, le phénomène est plus accentué étant donné que la dernière position affiche très peu d'erreur (quelques centaines de mètres). Notons que la taille des échantillons des passages à un message pour les gnous, marabouts et mouflons est trop faible pour en tirer des conclusions fiables (cf. Tab. 4.2).

Pour les moindres carrés, on retrouve aussi des phénomènes de biais pour les localisations calculées avec deux messages (cf. annexe A) car l'ajustement est mené avec la dernière fréquence

4.2. SERVICE TEMPS-RÉEL AVEC UN FILTRE MULTI-MODÈLE

d'émission connue. L'impact est une erreur de localisation, orientée principalement en latitude (environ 300 m d'erreur par Hz de biais sur la fréquence d'émission).

Ellipses d'erreur

L'adéquation entre les erreurs estimées et les erreurs réelles (consistance du filtre) est un autre indicateur de qualité. Sous la condition que l'erreur est distribuée selon une gaussienne bi-dimensionnelle, on considère les ellipses d'erreurs à 1σ , $\sqrt{2}\sigma$ et 3σ — contenant théoriquement 39,3%, 63,2% et 98,9% des localisations réelles respectivement (se reporter au paragraphe 1.2.4 pour de plus amples détails). Les ellipses à $\sqrt{2}\sigma$ sont celles fournies par ARGOS. Le tableau Tab. 4.4 donne les probabilités empiriques avec lesquelles les localisations réelles se situent dans les ellipses d'erreur centrées sur les localisations estimées associées. Elles sont comparées aux probabilités théoriques.

TABLE 4.4.: Probabilités empiriques que la localisation réelle se situe dans l'ellipse d'erreur centrée sur la localisation estimée associée.

Type de mobile	1σ (39.3%)		$\sqrt{2}\sigma$ (63.2%)		3σ (98.9%)	
	MC	Filtre	MC	Filtre	MC	Filtre
Marabout	14%	14%	27%	21%	61%	51%
Oie	17%	20%	26%	27%	51%	49%
Gnou	26%	20%	43%	34%	82%	71%
Mouflon	31%	33%	46%	43%	71%	69%
Tortue à dos plat	17%	25%	24%	37%	49%	65%
Tortue verte	29%	28%	44%	40%	73%	67%
Lion de mer	31%	22%	43%	33%	73%	60%
Éléphant de mer	23%	18%	33%	26%	56%	50%
Bateau	23%	22%	36%	34%	65%	67%
Bouée	38%	30%	54%	45%	85%	80%

Une valeur plus faible que la probabilité théorique signifie que le traitement sous-estime l'erreur réelle (il est trop optimiste). Les deux méthodes sont dans ce cas et, en moyenne, le filtre affiche des ellipses d'erreur légèrement plus resserrées que les moindres carrés (de quelques pourcents). Nous n'avons volontairement pas cherché à améliorer la consistance du filtre. De cette façon, les utilisateurs retrouvent une relative équivalence entre les erreurs —et donc les classes de localisation— des deux algorithmes afin de garantir une compatibilité ascendante dans le service de distribution des données ARGOS. On rappelle que pour certains utilisateurs, seules les localisations de classes 1, 2 ou 3 sont fournies (balises à forte puissance). Pour ces derniers, calculer des estimations d'erreurs de localisation plus grandes aurait pour effet de diminuer le nombre de positions distribuées. D'autres utilisateurs évaluent aussi la qualité de la localisation à partir de la distribution en classes. Les études utilisateurs [Kuhn et al., 2009, Witt et al., 2010, Costa et al., 2010] notent bien cette sous-estimation récurrente de l'erreur ARGOS mais en se basant sur les rayons d'erreur équivalents.

La raison de ces ellipses optimistes peut être liée à l'omission de certains effets dans le modèle d'observation comme le déplacement et la dérive en fréquence à moyen terme (*i.e.* sur la durée d'un passage). On note que les bouées, équipées des oscillateurs les plus stables et des vitesses de déplacement faibles, présentent les ellipses d'erreur les plus fiables.

4.3. Service à valeur ajoutée sur les données archives

Dans la base de données ARGOS, les mesures de fréquences reçues pour chaque balise et les éphémérides satellites sont archivées depuis 2008. C'est l'opportunité de valoriser ces données en les retraitant avec un lisseur. Nous avons implémenté le lisseur IMM-RTS dans le module de calcul ARGOS. Les erreurs de localisation du lisseur ont été comparées avec celles du filtre multi-modèle et de l'ajustement par moindres carrés.

4.3.1. Un effet régularisant sur l'erreur

Sur les passages avec un minimum de quatre messages (cf. Fig. 4.8), l'amélioration apportée par le lisseur multi-modèle est modérée avec une réduction de l'erreur empirique moyenne située entre 3% et 20%. Les progressions les plus importantes concernent les éléphants de mer (20%), les tortues à dos plat (15%) et les bouées (14%). Lorsqu'on se compare avec les moindres carrés, la réduction oscille entre 4% (oies) et 42% pour les éléphants de mer. Les deux jeux de tortues, les bateaux et les bouées affichent une progression de plus de 20%.

Avec deux ou trois messages, l'effet est plus fort puisque l'erreur est réduite globalement d'un tiers (cf. Fig. 4.9) entre le lisseur et le filtre, et de deux tiers entre le lisseur et l'ajustement par moindres carrés. Les éléphants de mer et les tortues à dos plat présentent d'ailleurs jusqu'à 76% et 77% de réduction de l'erreur respectivement. Notons que l'amplitude des erreurs se situe avec le lisseur entre 460 m pour les bouées et 6700 m pour les oies contre 1100 m et 17 000 m avec les moindres carrés.

Les passages à un message voient leurs erreurs presque divisées par deux entre le filtre et le lisseur (cf. Fig. 4.10). Elles se situent entre 556 m pour les bouées et 6400 m pour les oies. Globalement, on peut s'attendre maintenant à quelques kilomètres d'erreur en moyenne avec le lisseur sur ces passages avec une observation unique.

Comme pour le filtre, le lisseur améliore les passages où la géométrie est peu favorable. Cela est vrai si on s'intéresse au nombre de messages reçus mais aussi à la distance à la trace. Nous avons repris pour cela les données des bouées pour compléter le graphique 4.11 des erreurs moyennes en fonction de la distance à la trace. Le lisseur continue d'aplanir la courbe d'erreur puisque les effets liés à la perte d'observabilité sous la trace et à l'horizon sont presque atténués. À ce titre, on peut parler d'un effet régularisant du lissage sur l'erreur de localisation. Cela s'exprime aussi au travers des écarts-types des erreurs empiriques qui sont largement diminués entre le filtre et le lisseur : de 6% à 25% avec au moins quatre messages, de 30% à 50% avec deux ou trois mesures et de 32% à 60% avec une seule observation.

L'effet du lissage est visible sur la trajectoire Fig. 4.12 de l'éléphant de mer. On observe la disparition des pics liés à des passages à un message et la trajectoire globale se superpose assez fidèlement

4.3. SERVICE À VALEUR AJOUTÉE SUR LES DONNÉES ARCHIVES

à la vérité terrain GPS. Sur des échelles plus petites (cf. les extraits Fig. 4.13), la trajectoire est aussi plus cohérente avec la vérité terrain.

4.3.2. Caractérisation de l'erreur

Erreurs tangentielles

Sur les erreurs tangentielles, nous avons observé un effet de trainée notamment lorsque le nombre de messages est faible. L'application du lisseur a permis de supprimer ce défaut. Le tableau 4.5 montre notamment que les bateaux et bouées qui étaient les plus concernés ont un ratio bien plus faible entre l'erreur tangentielle et le module de l'erreur. Avec un message reçu, les ratios étaient globalement négatifs avec le filtre ce qui n'est plus le cas avec le lisseur. On rappelle que la solution du lisseur est équivalente à celle fournie par la fusion des estimées issues de deux filtres se propageant en sens inverses. L'effet de trainée du filtre avant est compensé par l'effet de trainée du filtre arrière. L'absence de cet effet de retard avec peu de mesures a sans doute aussi contribué à la diminution de l'erreur de localisation.

TABLE 4.5.: Rapport de l'erreur tangentielle sur le module de l'erreur

Type de mobile	4 mess. ou plus		2 et 3 mess.		1 mess.	
	Filtre	Lisseur	Filtre	Lisseur	Filtre	Lisseur
Marabout	1%	1%	10%	10%	37%	71%
Oie	4%	4%	-1%	-2%	-8%	-1%
Gnou	3%	2%	-16%	-11%	-25%	-36%
Mouflon	-2%	-1%	7%	12%	-19%	9%
Tortue à dos plat	-13%	-5%	-9%	10%	-14%	11%
Tortue verte	5%	16%	-15%	15%	-22%	9%
Lion de mer	15%	17%	-4%	18%	-2%	6%
Éléphant de mer	-2%	13%	-8%	17%	-18%	13%
Bateau	-1%	1%	-18%	0%	-24%	-6%
Bouée	-14%	1%	-38%	8%	-59%	9%

Ellipses d'erreur

Concernant les ellipses d'erreur produites par le lisseur, on retrouve un niveau comparable avec le filtre dans le tableau 4.6. Néanmoins, on note que les probabilités observées ont progressé de quelques pourcents.

CHAPITRE 4. MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS

TABLE 4.6.: Probabilités empiriques que la localisation réelle se situe dans l'ellipse d'erreur centrée sur la localisation estimée associée.

Type de mobile	1σ (39.3%)		$\sqrt{2}\sigma$ (63.2%)		3σ (98.9%)	
	Filtre	Lisseur	Filtre	Lisseur	Filtre	Lisseur
Marabout	15%	14%	22%	23%	51%	51%
Oie	20%	20%	27%	28%	49%	50%
Gnou	21%	21%	34%	34%	70%	70%
Mouflon	33%	34%	45%	47%	70%	72%
Tortue à dos plat	25%	28%	37%	40%	65%	68%
Tortue verte	28%	32%	41%	45%	67%	71%
Lion de mer	22%	24%	34%	35%	60%	61%
Éléphant de mer	18%	20%	26%	29%	50%	54%
Bateau	21%	22%	31%	33%	59%	61%
Bouée	30%	34%	45%	50%	80%	84%

Conclusions

Le filtre IMM a été appliqué sur des données ARGOS avec un ensemble de modèles à la fois réduit et généraliste. Avec cette approche, malgré la grande diversité des mobiles suivis, les erreurs moyennes et les écarts-types associés ont été significativement réduits par rapport au traitement moindres carrés. La modélisation d'état qui soutient le filtrage permet de fournir une estimation de l'erreur de positionnement dans tous les cas et d'augmenter très largement le nombre de localisations en présence d'un faible nombre de messages reçus. Le filtrage est pertinent quand la géométrie d'observation est mauvaise car celle-ci est alors indirectement complétée par l'information issue de la dernière localisation connue via le modèle de déplacement a priori.

Le lisseur IMM-RTS a en outre été implémenté. L'évaluation des performances sur le panel de balises ARGOS a montré qu'il est encore possible de réduire significativement l'erreur de localisation par rapport au filtre lorsque la balise est suivie dans des conditions difficiles (moins de 4 messages). L'effet est particulièrement important sur les localisations à un message dont l'erreur moyenne atteint quelques kilomètres. En outre, le lisseur élimine dans ce cas les effets de trainée.

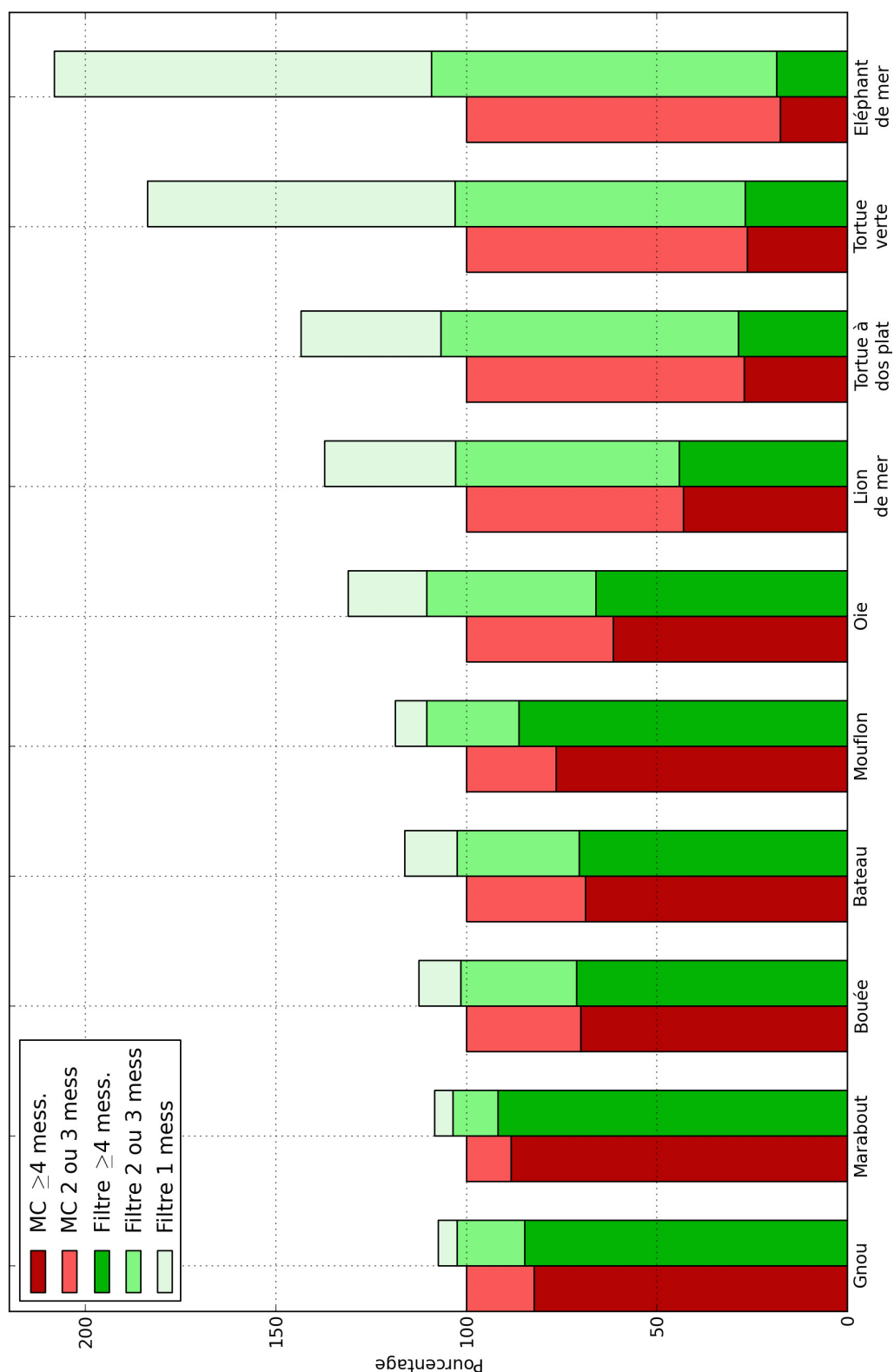


FIGURE 4.1.: Répartition des localisations en fonction du nombre de messages reçus (normalisé par rapport au nombre de localisations calculées par moindres carrés (MC))

CHAPITRE 4. MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS

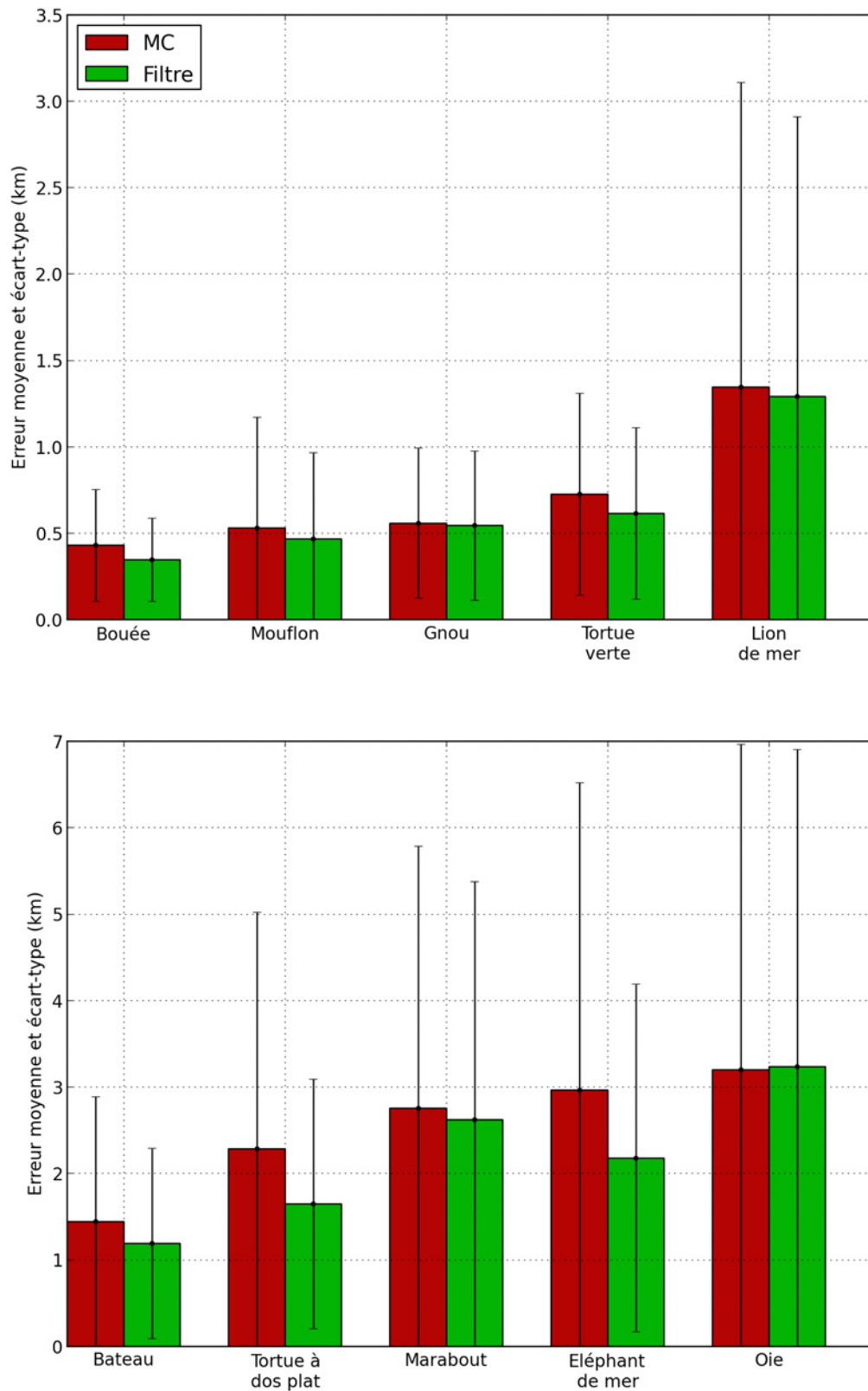


FIGURE 4.2.: Comparaison des erreurs empiriques de localisation des passages à plus de 4 messages entre l'ajustement moindres carrés et le filtre multi-modèle.

4.3. SERVICE À VALEUR AJOUTÉE SUR LES DONNÉES ARCHIVES

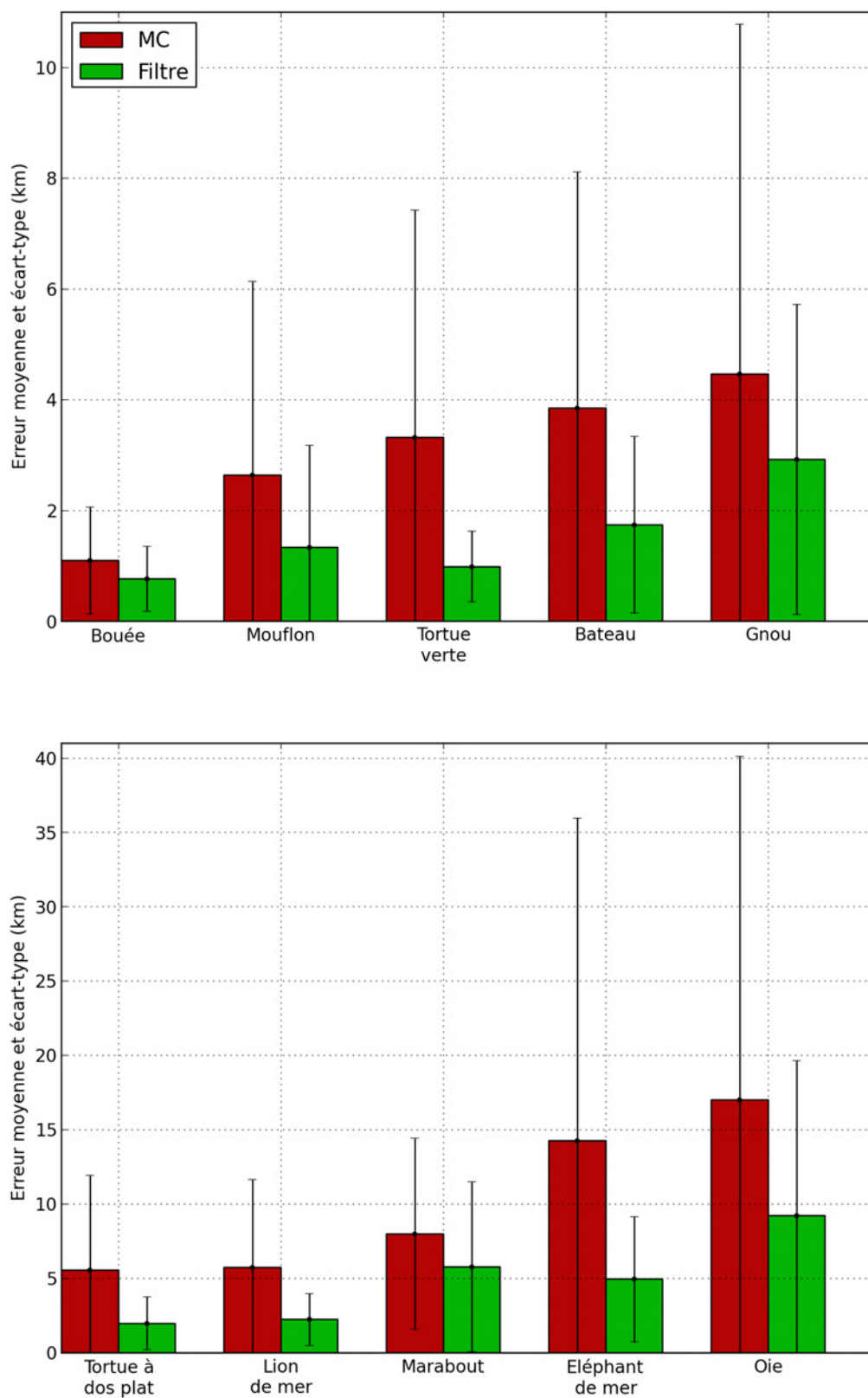


FIGURE 4.3.: Comparaison des erreurs empiriques de localisation des passages à 2 ou 3 messages entre l'ajustement moindres carrés et le filtre multi-modèle.

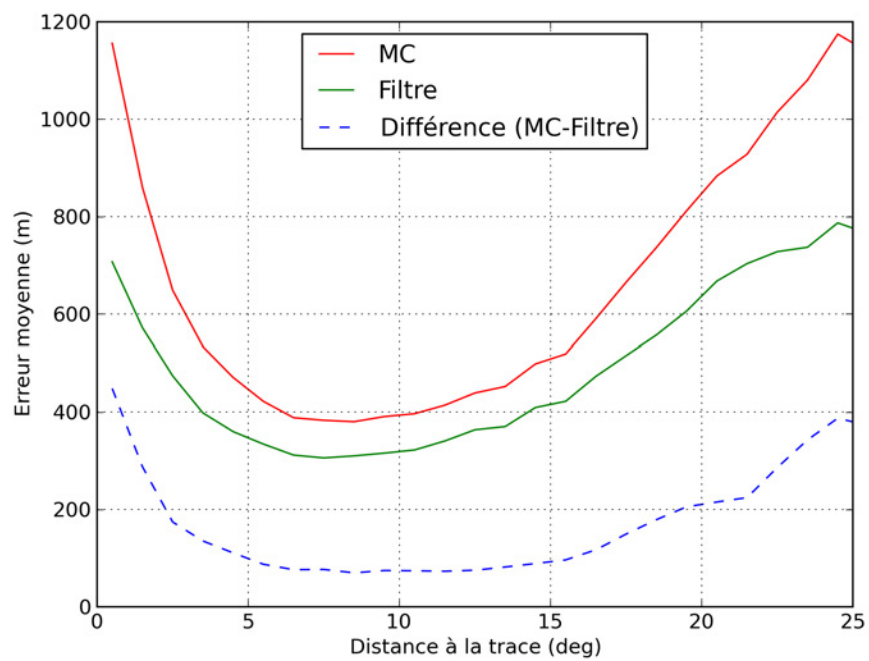


FIGURE 4.4.: Erreurs empiriques de localisation moyennes en fonction de la distance à la trace.

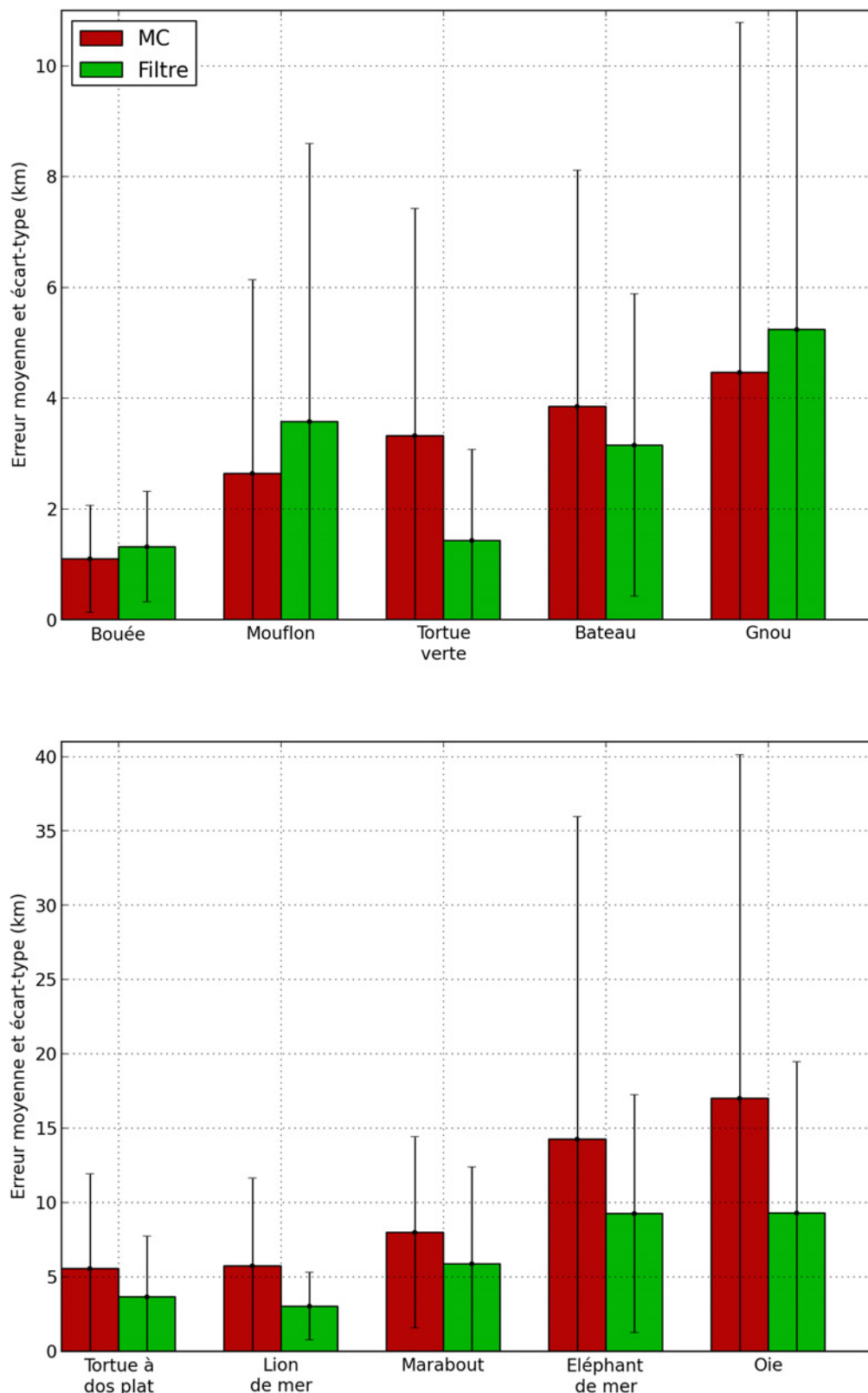
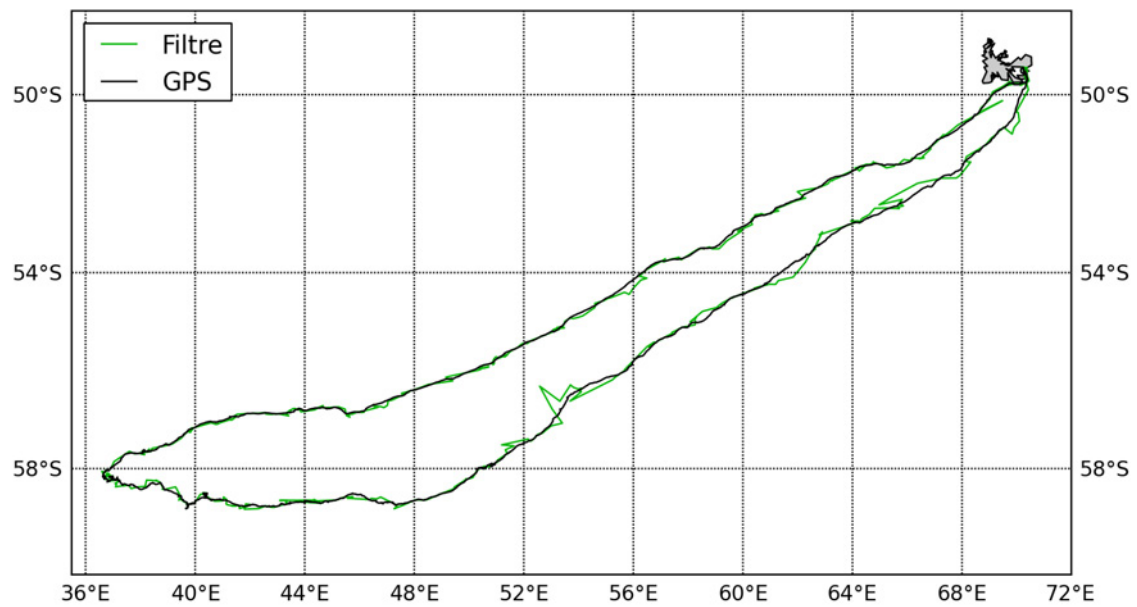
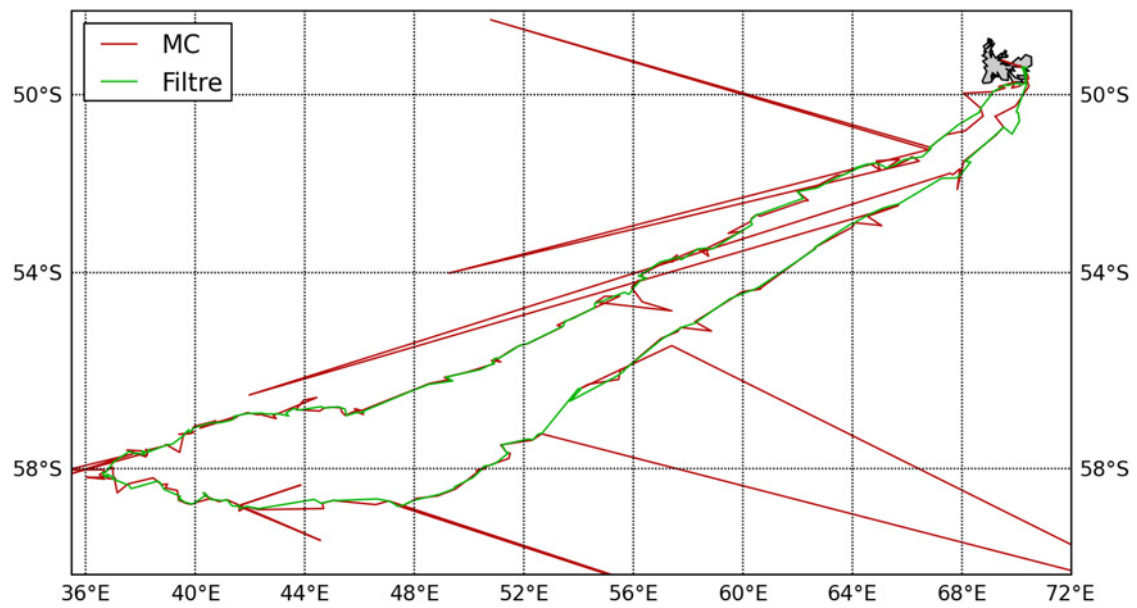


FIGURE 4.5.: Comparaison des erreurs empiriques de localisation obtenues avec le filtre multi-modèle pour des passages à un message, et des erreurs empiriques obtenues par moindres carrés avec deux ou trois messages.



(a) GPS et filtre multi-modèle.



(b) Moindres carrés et filtre multi-modèle (localisations avec **deux messages ou plus** seulement).

FIGURE 4.6.: Trajectoire d'éléphant de mer pour le traitement moindres carrés et le filtre multi-modèle.

4.3. SERVICE À VALEUR AJOUTÉE SUR LES DONNÉES ARCHIVES

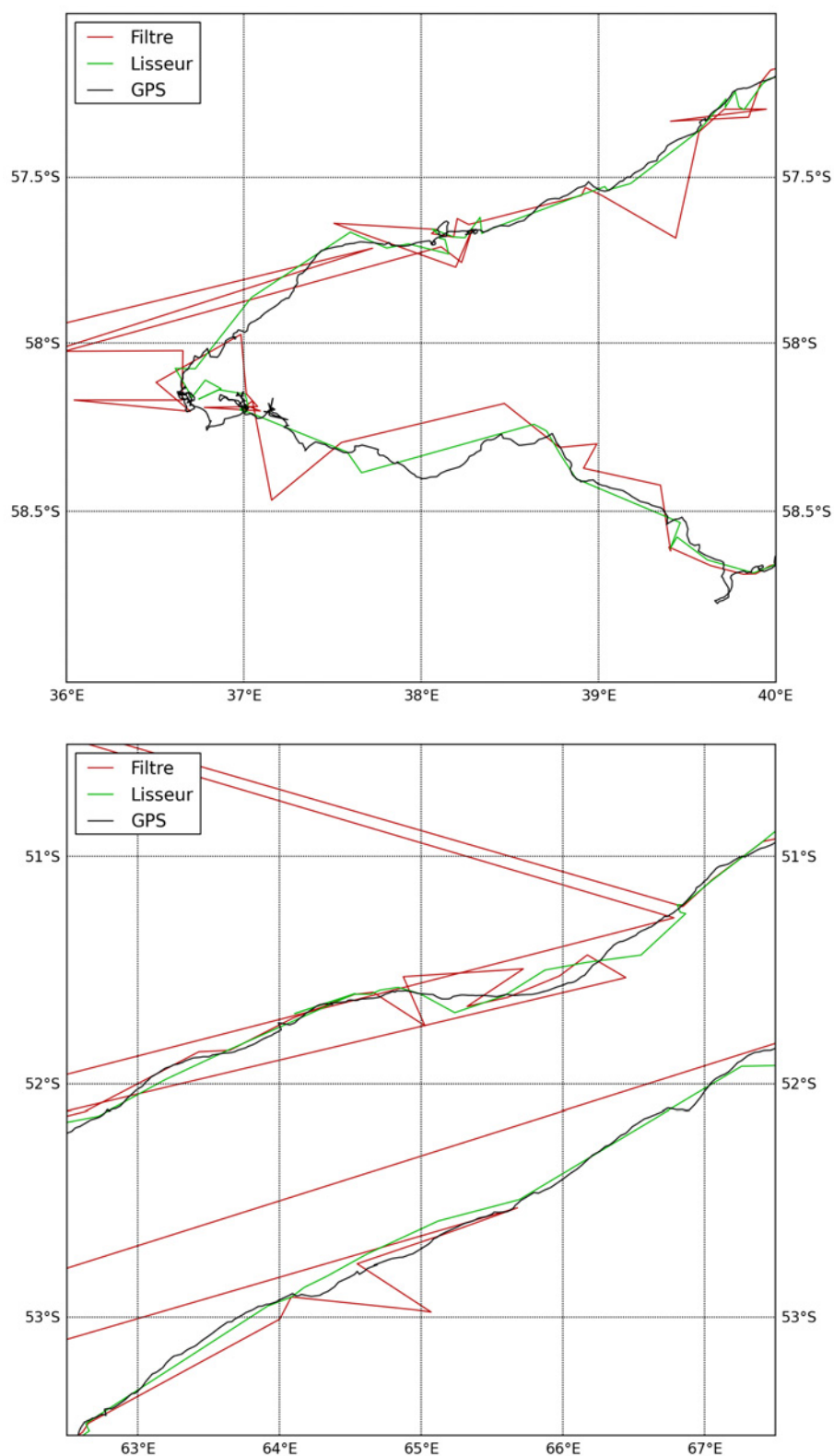


FIGURE 4.7.: Extraits de trajectoire d'éléphant de mer pour le traitement moindres carrés et le filtre multi-modèle (localisations avec **deux messages ou plus** seulement).

CHAPITRE 4. MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS

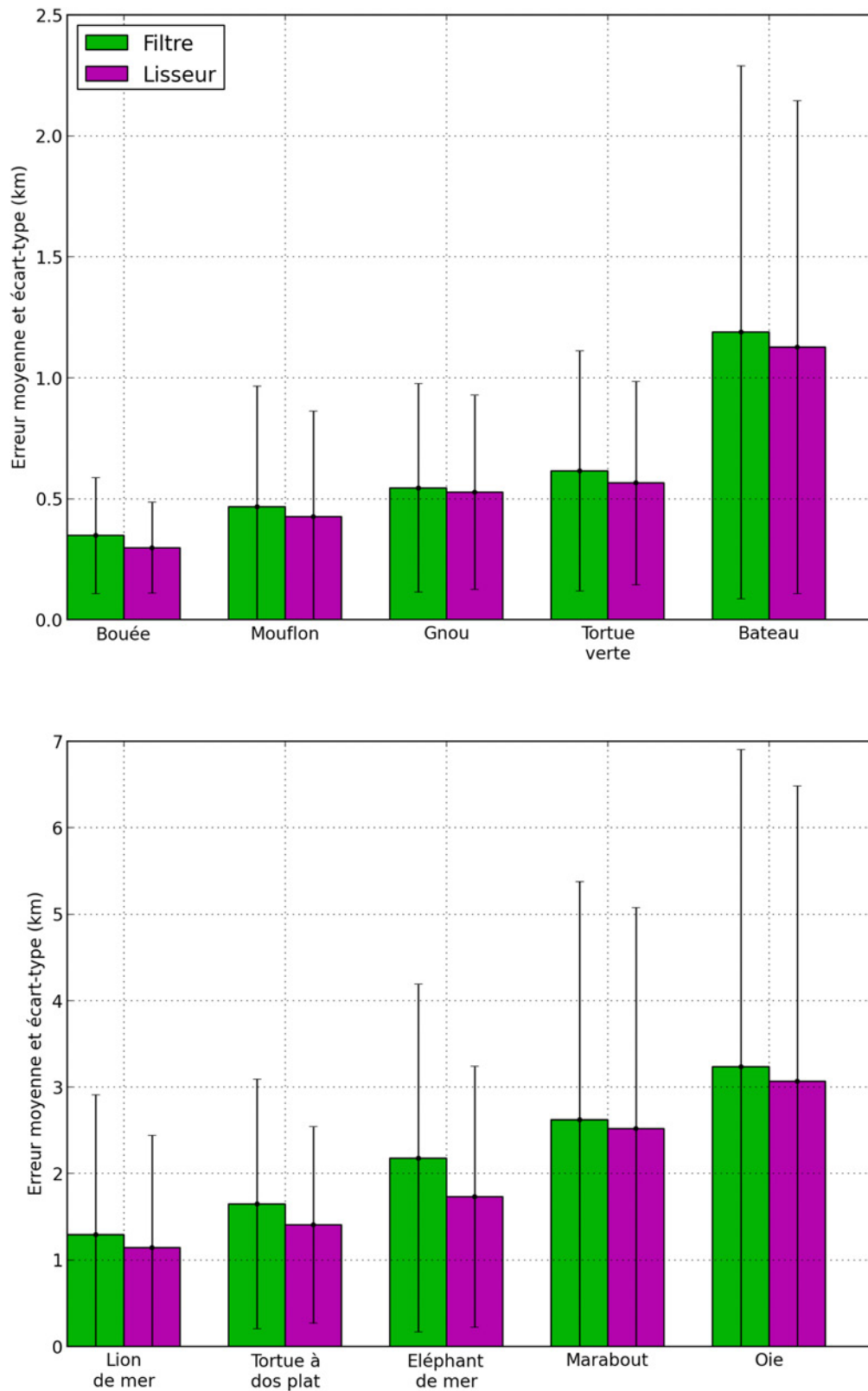


FIGURE 4.8.: Comparaison des erreurs empiriques de localisation des passages à 4 messages ou plus entre le filtre et le lisseur multi-modèles.

4.3. SERVICE À VALEUR AJOUTÉE SUR LES DONNÉES ARCHIVES

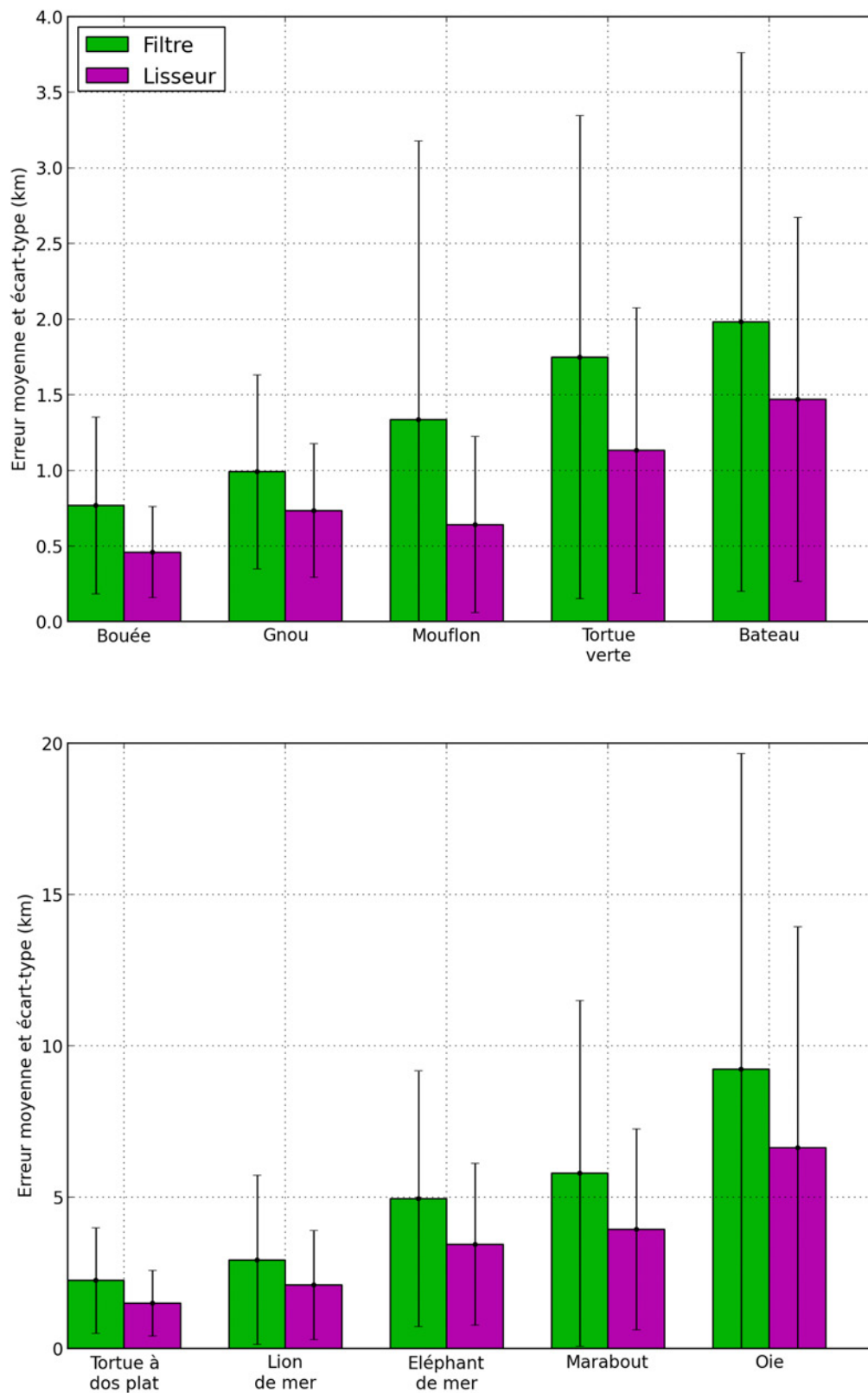


FIGURE 4.9.: Comparaison des erreurs de localisation empiriques des passages à 2 ou 3 messages entre le filtre et le lisseur multi-modèles.

CHAPITRE 4. MISE EN ŒUVRE DE NOUVEAUX ALGORITHMES DE LOCALISATION POUR ARGOS

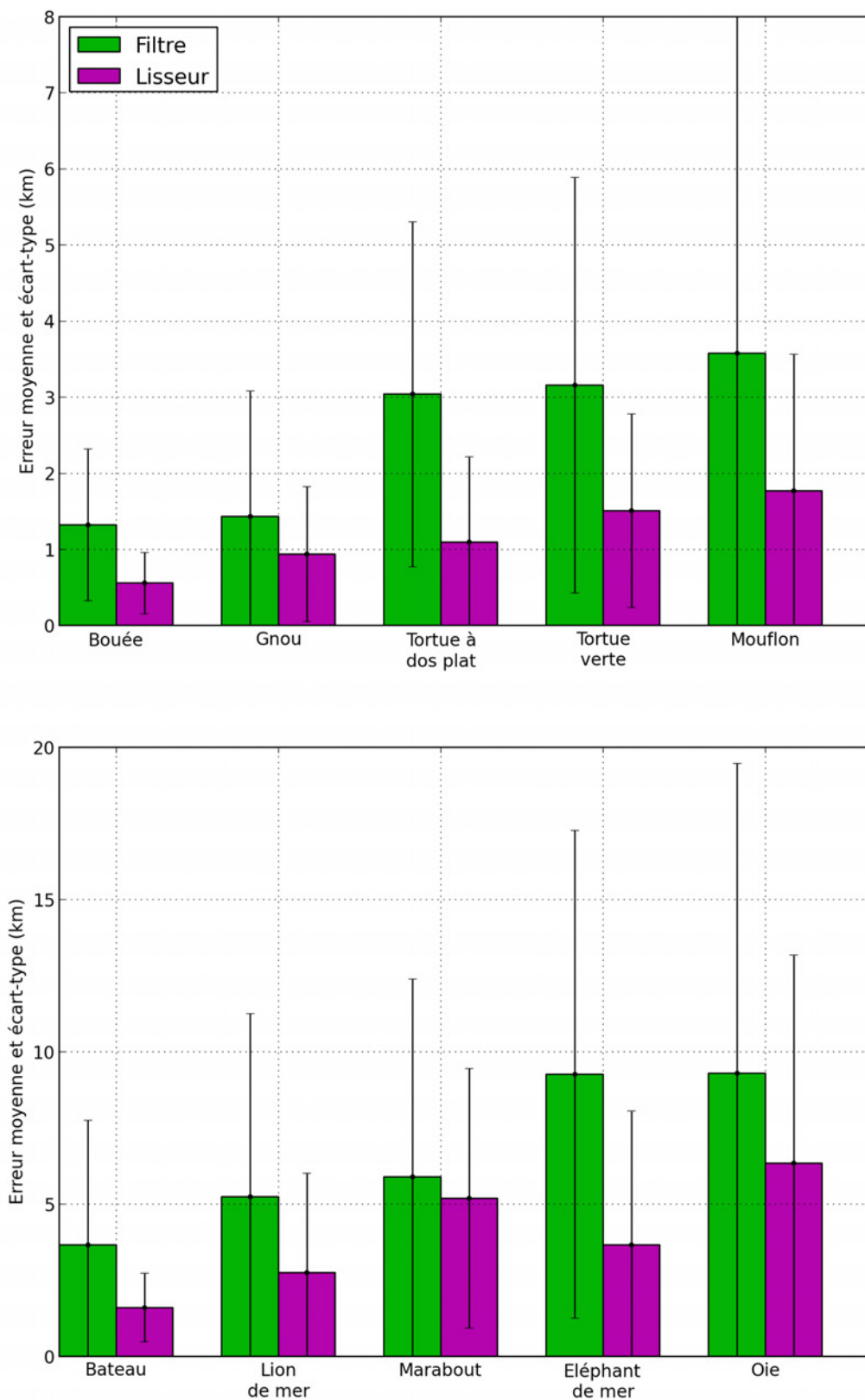


FIGURE 4.10.: Comparaison des erreurs de localisation des passages à 1 message entre le filtre et le lisseur multi-modèles.

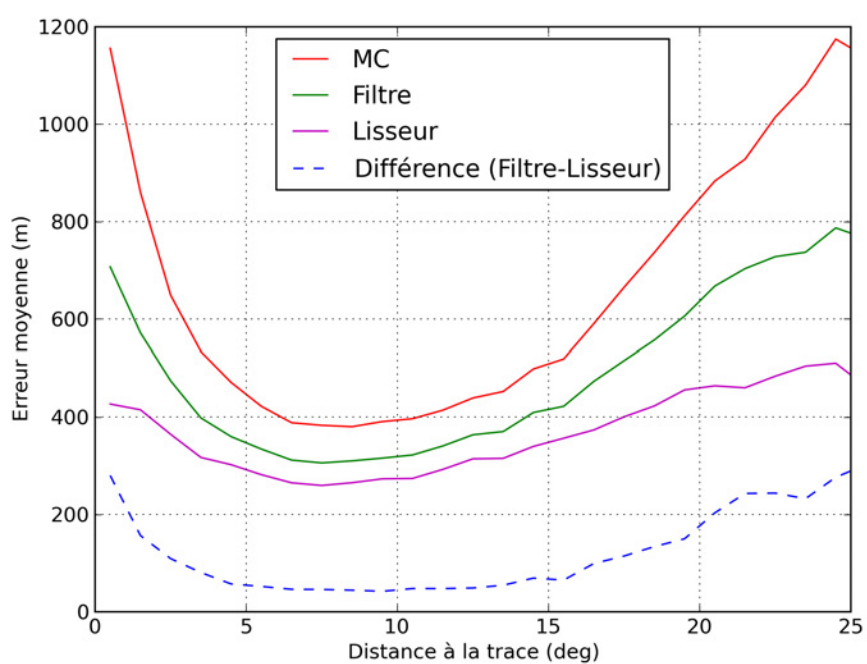
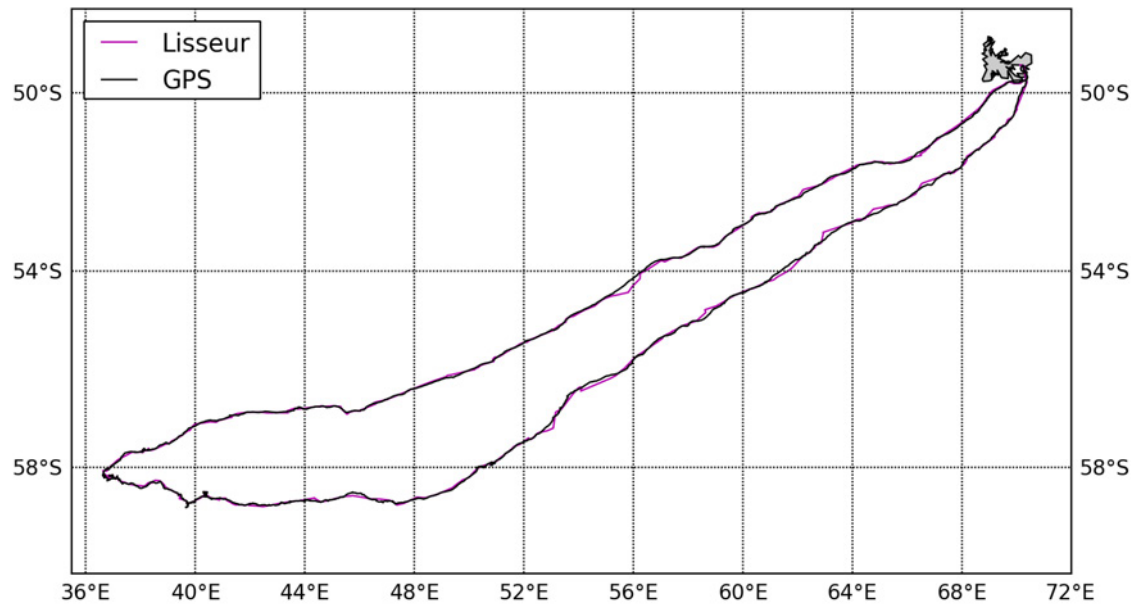
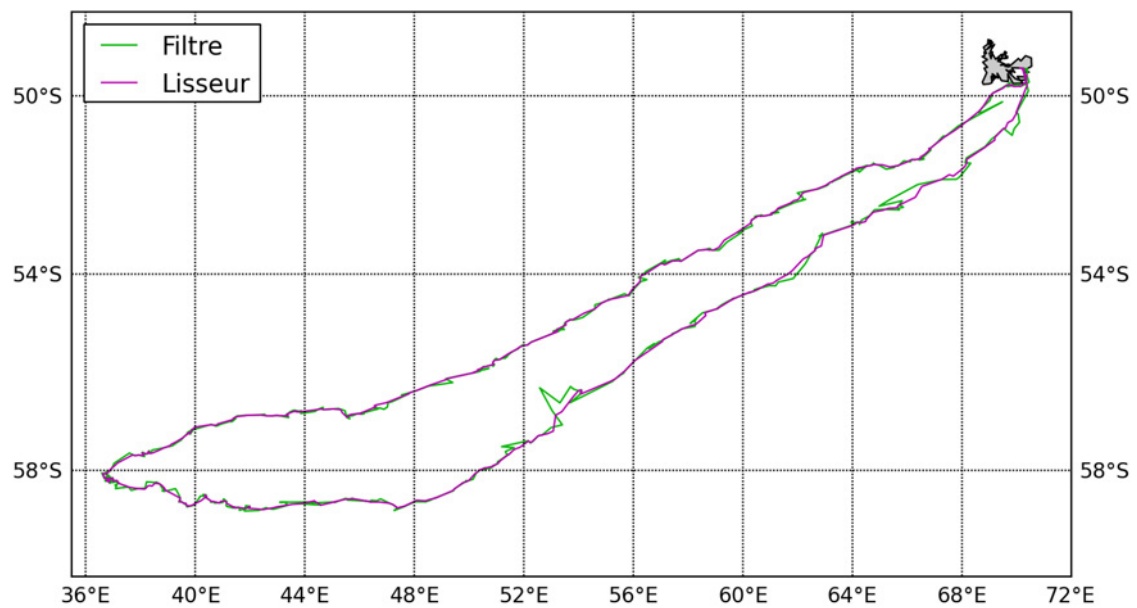


FIGURE 4.11.: Erreurs de localisation moyennes en fonction de la distance à la trace.



(a) GPS et lisseur multi-modèle.



(b) Filtre et lisseur multi-modèles.

FIGURE 4.12.: Trajectoire d'éléphant de mer pour le filtre et le lisseur multi-modèles.

4.3. SERVICE À VALEUR AJOUTÉE SUR LES DONNÉES ARCHIVES

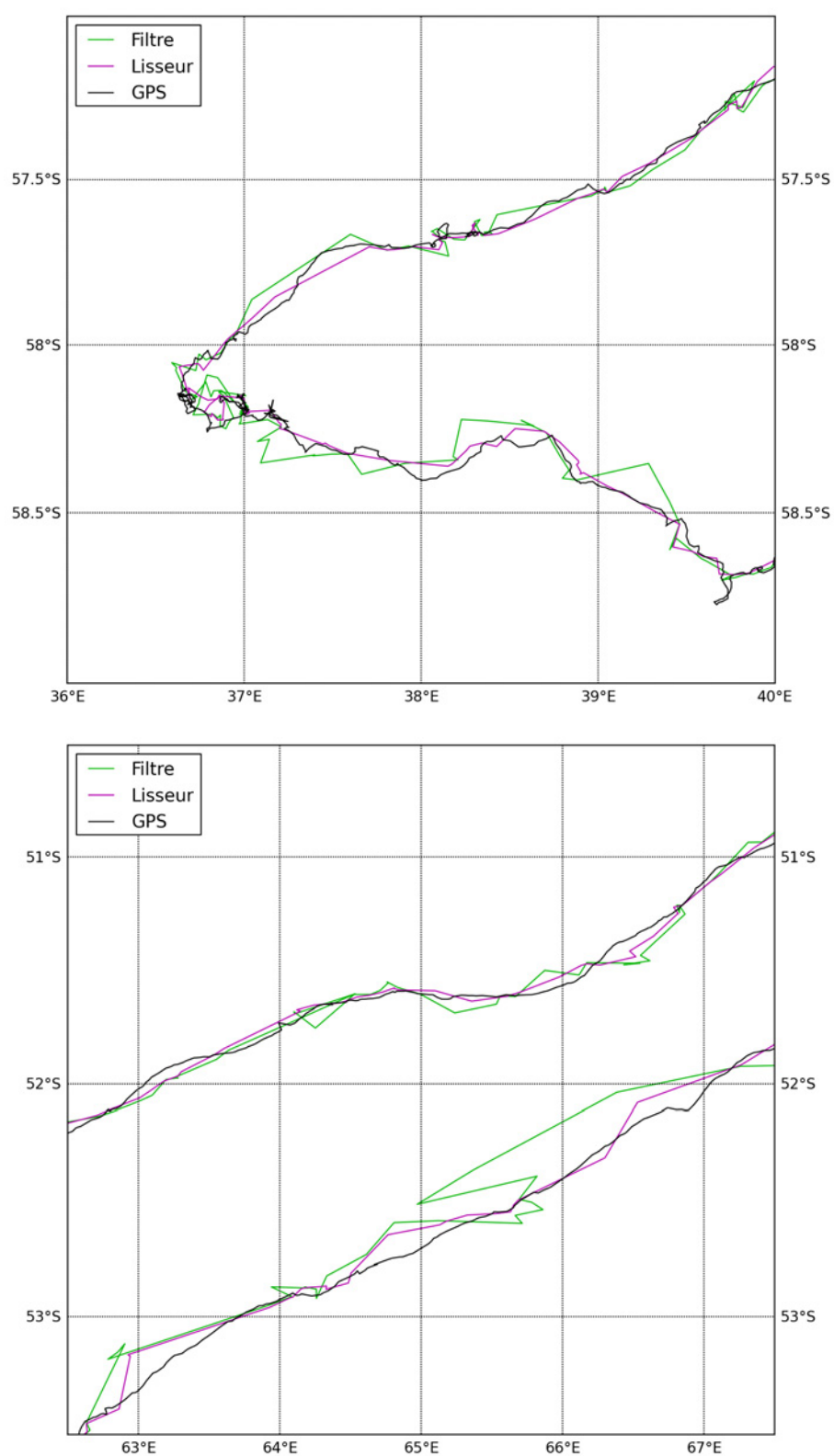


FIGURE 4.13.: Extraits de trajectoire d'éléphant de mer pour le filtre et le lisseur multi-modèles.

SYNTHÈSE ET PERSPECTIVES

Au travers de ce projet, nous avons conçu et industrialisé de nouveaux traitements de localisation pour le système ARGOS.

Initialement résolu avec un ajustement par moindres carrés, le problème de localisation a été reformulé dans le cadre plus général des systèmes dynamiques à sauts markoviens. Celui-ci a ouvert la voie à des méthodes de filtrage et de lissage multi-modèles qui éliminent deux limitations de l'algorithme historique, à savoir l'incapacité de caractériser l'erreur de positionnement et de déterminer une localisation avec un seul message reçu. La modélisation d'état sous-jacente, alliée avec la possibilité de basculer entre plusieurs dynamiques de déplacement, constitue en outre une formulation flexible pour tenir compte des comportements possibles des balises. Néanmoins, la détermination de la loi a posteriori de l'état se heurte à des limitations pratiques puisqu'elle implique une complexité exponentiellement croissante avec le temps. La littérature est riche de solutions sous-optimales qui limitent la taille de l'arbre des séquences des modèles possibles, soit en supprimant judicieusement certaines branches, soit en fusionnant celles dont les sous-séquences sont identiques. La fusion est plus apte au suivi des mobiles ARGOS dont les modèles de comportement ne sont pas parfaitement connus. Dans ce contexte, l'Interacting Multiple Model offre un rapport coût/performance avantageux pour calculer la densité a posteriori de l'état-localisation. Il n'utilise que M filtres pour M modes lors de l'assimilation de la mesure grâce au choix judicieux de l'instant d'application de la fusion dans le cycle de filtrage.

Dans le domaine du lissage, de nombreuses solutions ont aussi été établies, mais l'analyse du cycle exact de récursion a permis de lever le voile sur un autre chemin jusqu'alors inexploré. Nous avons montré que les quantités issues de l'IMM suffisent à construire récursivement l'état lissé via une passe à rebours de type RTS et que seuls M lisseurs pour M modes sont nécessaires. En ce sens, cette approche est plus économique en matière d'implémentation informatique que les méthodes équivalentes. La passe arrière de lissage s'initialise simplement avec la dernière estimée filtrée, condition indispensable pour une application à tous les cas ARGOS. Enfin, il est maintenant possible de faire jeu égal avec les performances d'estimation du lisseur double filtre qui nécessite une dynamique inversible et d'initialiser le filtre arrière sans information a priori sur l'état.

Le besoin de faire coopérer dans l'IMM des dynamiques dont les vecteurs d'état sont de dimensions différentes (variantes de marches aléatoires) nous a incités à proposer une extension de la formulation initiale. Cette problématique est peu traitée dans la littérature voire souvent éludée en réadaptant les vecteurs d'état lors de la phase d'interaction de l'IMM. Nous avons montré qu'il est possible d'établir les équations d'un IMM hétérogène dans le cas gaussien en réinterprétant la phase d'interaction par couple de variables. Plutôt que de manipuler les vecteurs d'état complets, les moments de la densité d'interaction sont restitués en appliquant l'interaction sur les ensembles équivalents de densités de couples de variables. Il est aussi possible de déduire des formules de lissage. L'application sur des données simulées montre que l'approche est viable. Le seul bémol tient dans l'instabilité de la solution de lissage RTS qui empêche son application à ARGOS. Ce dernier point constitue un sujet ouvert.

La mise en œuvre de ces méthodes pour ARGOS a été l'opportunité de réécrire le module de localisation en se basant sur un filtre et un lisseur IMM homogènes. Le choix a été fait de considérer un ensemble de modèles génériques. Nous avons testé ces nouvelles techniques sur un panel large de balises avec la création d'outils dédiés d'une part au retraitement des données de la base archive ARGOS et d'autre part à l'évaluation des performances des différents algorithmes. Les nouvelles approches de filtrage et de lissage augmentent le nombre de localisations distribuées et améliorent leur précision pour des plateformes émettant dans des situations défavorables : peu de messages reçus, géométries d'observation dégradées, localisations images régulières... Sur les passages à quatre observations ou plus, l'erreur est réduite d'environ 25% entre les moindres carrés et le lisseur. Sur des passages avec seulement deux ou trois messages, l'erreur du lissage a été ramenée globalement au tiers de l'erreur des moindres carrés. La précision de localisation avec une unique observation de fréquence est de l'ordre de quelques kilomètres en moyenne. Le filtre et le lisseur présentent des performances plus régulières avec la réduction des écarts-types et de l'étendue des erreurs quelle que soit la géométrie d'observation.

L'algorithme exploitant le filtre multi-modèle a été proposé à la communauté ARGOS après plus d'un an de développement, de promotion et de validation pour atteindre les standards de qualité de CLS. Les utilisateurs peuvent, depuis le 15 mars 2011, choisir entre l'algorithme historique et le filtre pour suivre leurs balises. Au 1^{er} mai 2013, 88% des balises liées aux applications océanographiques ont choisi de basculer et 63% des plateformes animaux. Autre chiffre encourageant, près de 90% des plateformes nouvellement déclarées sont localisées avec le filtre multi-modèle. De son côté, le service à valeur ajoutée de lissage a été proposé dans le Plan à Moyen Terme ARGOS et doit faire l'objet d'une présentation à notre direction commerciale et au CNES. Il a été utilisé de manière ponctuelle par [Domeier and Nasby-Lucas, 2013] pour étudier sur deux ans la migration de requins blancs femelles dans le Pacifique. Ces évolutions sont essentielles pour entretenir l'intérêt d'ARGOS et faire progresser la qualité de nos services.

À court et moyen terme, nous envisageons plusieurs améliorations de la chaîne de localisation :

- le remplacement du modèle numérique de terrain basé sur GTOPO30 (1 km de résolution à l'équateur) par SRTM3 (100 m de résolution [Jarvis et al., 2008]) couplé avec la donnée du niveau moyen des mers ;
- l'exploitation du rapport signal à bruit, fourni par certains satellites à la réception d'un message, afin de mieux quantifier le bruit de mesure ;
- la mise au point d'un processus d'initialisation qui amorce l'algorithme sans attendre un passage à quatre messages (filtre d'information par exemple) ;
- la refonte du module d'orbitographie pour fournir des éphémérides de meilleure qualité (à la dizaine de mètres contre cent mètres actuellement).

Sur le long terme, l'intégration de contraintes et de nouvelles sources de données dans le traitement constitue un premier axe de travail pour rendre les trajectoires plus réalistes. Cela s'adresse plus particulièrement aux animaux marins dont les trajets sont limités par les côtes et dont les capteurs fournissent parfois des données de profondeur de plongée et de température de surface. La prise en compte de sources supplémentaires d'information nous incite, comme second axe de réflexion, à faire évoluer nos approches de filtrage. Les filtres de Kalman ont bien été adaptés aux traitements sous contraintes [Simon, 2010] mais il semble plus pertinent de s'affranchir de l'hypothèse gaussienne lorsque les contraintes sont très fortement non linéaires (côte très découpée). La

conformation des densités est souvent loin d'être gaussienne avec la présence possible de plusieurs modes. Il existe des outils de filtrage qui n'imposent pas la forme des distributions de probabilité comme les méthodes séquentielles de Monte Carlo [Gordon et al., 1993, Doucet et al., 2001] et les filtres sur grille [Pedersen, 2007, Thygesen et al., 2009] utilisés par exemple chez CLS pour la localisation de poissons à partir de mesures d'intensité lumineuse. Comparés à la donnée d'une trajectoire et de ses ellipses d'erreur, ils proposent une vision plus riche et complète de la localisation possible d'un animal sous la forme d'une approximation non paramétrique de la densité de probabilité de ses positions successives.

Annexe A

COMPLÉMENTS SUR L'ANALYSE D'ERREUR A PRIORI

Cette annexe prolonge l'analyse d'erreur a priori en s'intéressant d'abord à l'impact d'un biais sur l'altitude et la fréquence d'émission. Dans un deuxième temps, on détaille les effets de la dérive en fréquence de la balise et de son déplacement pendant le passage.

A.1. Biais sur l'altitude

Lorsque l'altitude est une variable estimée, l'erreur globale $\sqrt{\sigma_\lambda^2 + \sigma_\phi^2 + \sigma_h^2}$ est donnée dans Fig. A.1. L'erreur atteint son minimum vers 6 degrés de distance et sa valeur est multipliée par dix par rapport au cas où l'altitude est donnée. Sous la trace, on observe toujours une divergence de l'erreur à cause d'une perte d'observabilité sur la longitude. Vers l'horizon, l'erreur diverge à cause de la très mauvaise observabilité de l'altitude h .

ARGOS suppose donc que l'altitude est connue lors du calcul. Les animaux marins et les bouées utilisent une valeur nulle alors que celle des animaux terrestres et des oiseaux est fournie par un modèle de terrain. Pour les oiseaux migrateurs qui passent du temps en vol, leur altitude réelle est donc biaisée par rapport à celle du modèle de terrain. Pour une distance à la trace donnée, les erreurs sur la longitude et la latitude sont proportionnelles à l'altitude. Le coefficient de proportionnalité est donné dans la figure 1.2 en fonction de la distance à la trace. Ces chiffres sont obtenus avec un passage de quatre messages parfaitement répartis (sans perturbations) et en appliquant un ajustement moindres carrés avec des biais de plus en plus grands sur l'altitude. La longitude présente des coefficients de -0,5 km d'erreur par km de biais sur l'altitude entre 15 et 25 deg de distance à la trace, et de -1 km/km entre 5 et 15 deg. Sous la trace, la valeur du coefficient diverge. En revanche, la latitude n'est pas impactée par un biais sur l'altitude avec une droite de pente nulle.

A.1. BIAIS SUR L'ALTITUDE

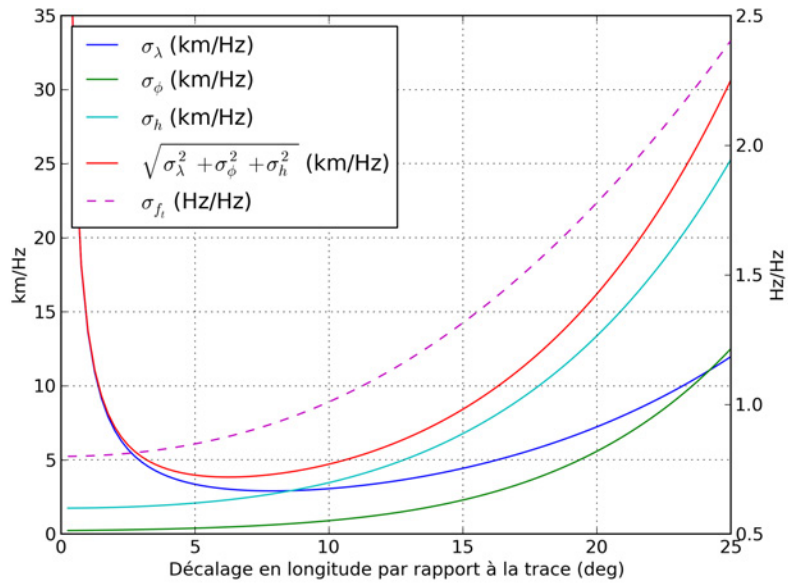


FIGURE 1.1.: Écarts-types des erreurs de localisation en fonction de la distance à la trace (par Hertz de bruit à 1 sigma sur la fréquence de réception) quand l'altitude est estimée.

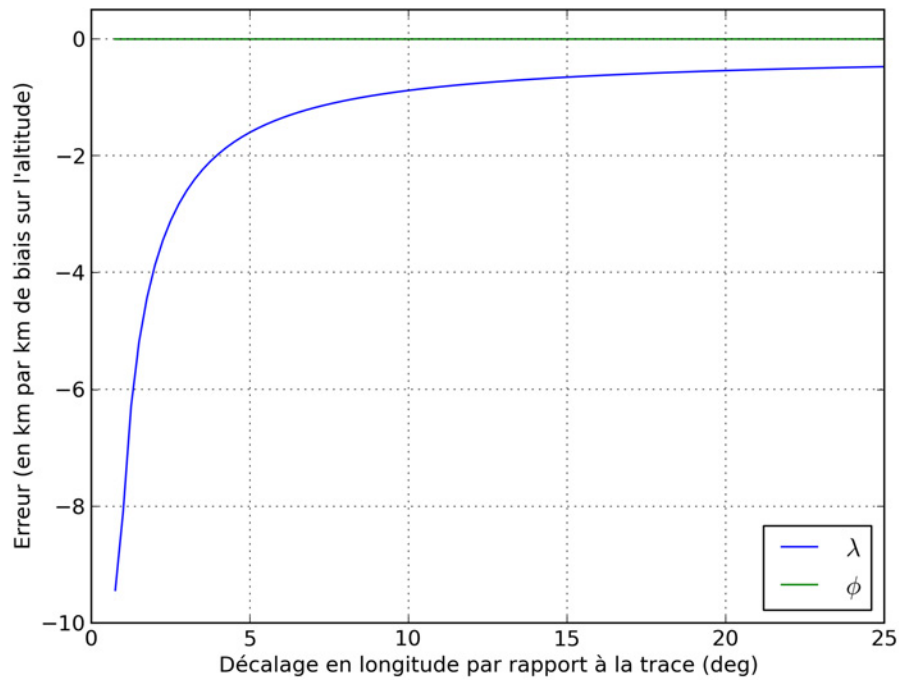


FIGURE 1.2.: Pente de la droite (km/km) donnant l'erreur en longitude λ et latitude ϕ en fonction du biais sur l'altitude

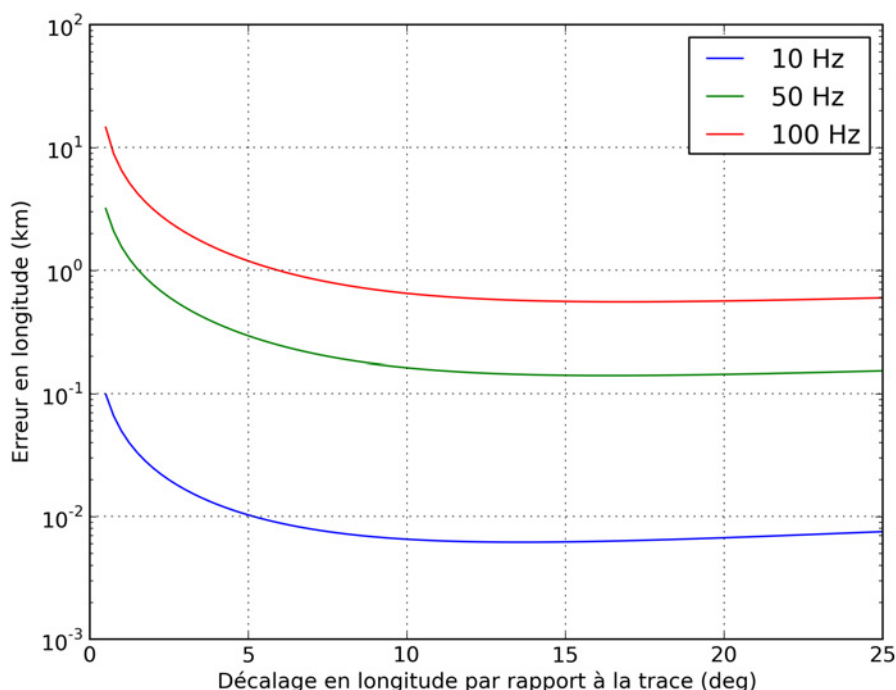


FIGURE 1.3.: Erreur en longitude (km) pour des biais sur la fréquence d'émission de 10 Hz, 50 Hz et 100 Hz. Le comportement est quadratique.

A.2. Biais sur la fréquence d'émission

L'ajustement par moindres carrés fait l'hypothèse que la fréquence d'émission est égale à la dernière connue lorsque seulement deux messages sont reçus. Pour les animaux qui subissent des gradients de température importants et pour lesquels peu de messages sont généralement reçus, cette supposition a un coût sur l'erreur de localisation. Pour la longitude, l'erreur résultante est quadratique avec le biais sur la fréquence. La figure 1.3 affiche sa valeur pour plusieurs biais sur la fréquence en fonction de la distance à la trace. Pour 10 Hz, on a environ 10 m d'erreur sur la longitude. Pour 100 Hz, on atteint presque 1 km. Sous la trace cette erreur diverge. Quant à l'erreur en latitude, elle est proportionnelle au biais de fréquence. Le coefficient de proportionnalité est donné dans la figure 1.4 en fonction de la distance à la trace. Avec 100 Hz, l'erreur présente une moyenne de 30 km, ce qui est bien plus qu'en longitude. Par ailleurs, on n'observe pas de divergence en s'approchant de la trace au sol du satellite. Ces données ont été calculées avec un passage de deux messages parfaitement répartis et un ajustement par moindres carrés réduit à l'estimation de la longitude et de la latitude.

A.3. Dérive sur la fréquence d'émission

Dans ce cas-ci, la fréquence est estimée mais elle subit une dérive pendant le passage à cause d'une instabilité de l'oscillateur. L'ajustement par moindres carrés compense la dérive en fréquence par

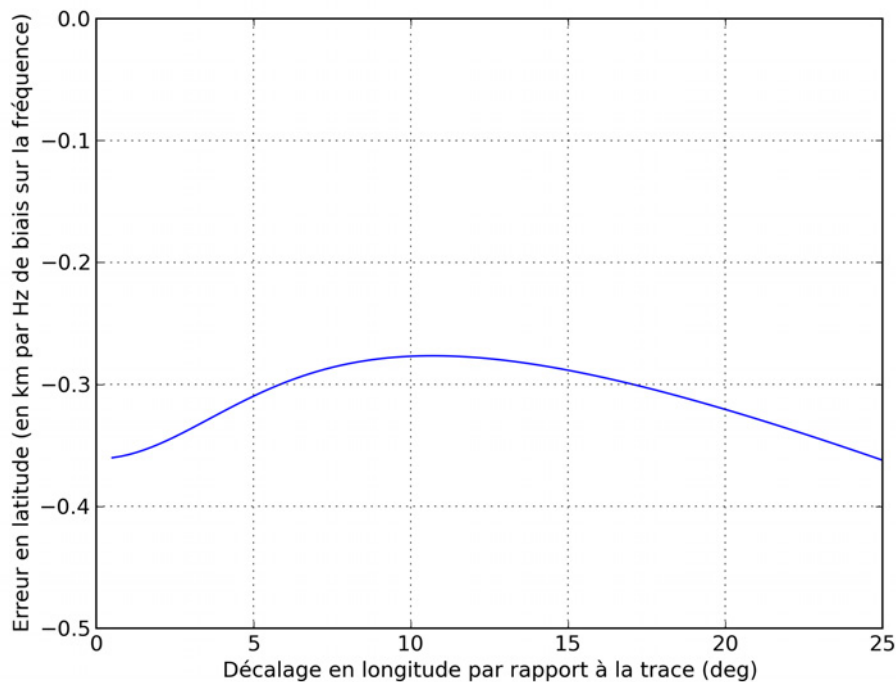
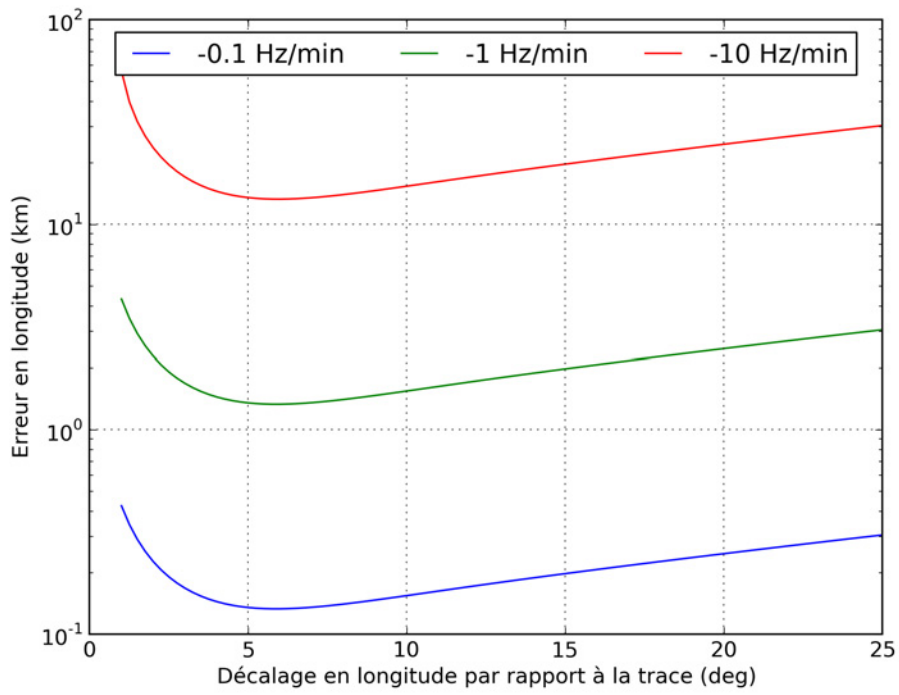


FIGURE 1.4.: Pente de la droite (km/Hz) donnant l'erreur en latitude en fonction du biais sur la fréquence d'émission.

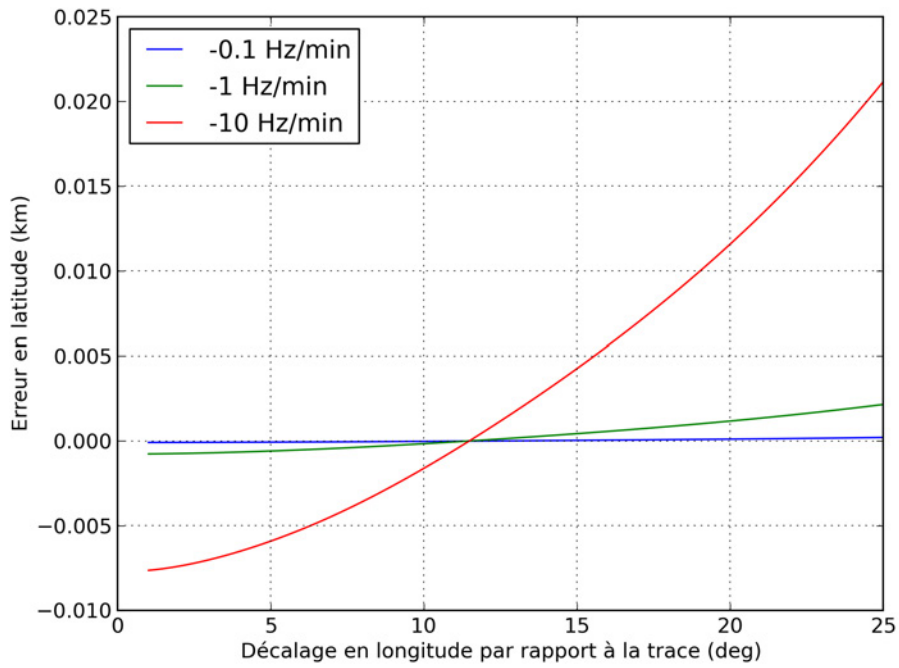
une erreur sur la longitude proportionnelle à la valeur de la dérive (Fig. A.5(a)) et elle atteint un peu plus de 10 km pour une pente de -10 Hz/min. En comparaison, l'erreur en latitude est négligeable (Fig. A.5(b)).

A.4. Déplacement pendant le passage

Pour terminer cette analyse complémentaire, nous considérons que la plateforme se déplace pendant le passage vers le nord ou vers l'est. Dans la direction du déplacement (en longitude vers l'est pour la Fig. A.6(a), en latitude vers le nord Fig. A.7(b)), l'erreur de localisation ne dépend pas de la distance à la trace et elle augmente proportionnellement à la vitesse (avec une valeur de 2,7 km pour 10 m/s). Lorsque la balise bouge vers l'est, l'erreur en latitude, *i.e.* perpendiculaire à la direction de la balise, diverge quand la distance à la trace augmente (cf. Fig. A.6(b)). En direction du nord, l'erreur en longitude est proportionnelle à la vitesse de la plateforme et varie selon la distance à la trace (cf. Fig. A.7(a)). Pour une vitesse de 10 m/s nord, l'ajustement affiche une erreur autour de 5 km en moyenne. Fait remarquable, une dérive en fréquence et un déplacement de la balise vers le nord ont exactement le même impact sur l'erreur (comparer les figures A.5(a) et A.7(a)). En d'autres termes, si l'on estime la vitesse et/ou la dérive en fréquence du mobile, l'ajustement n'est pas en mesure de différencier ces deux quantités.



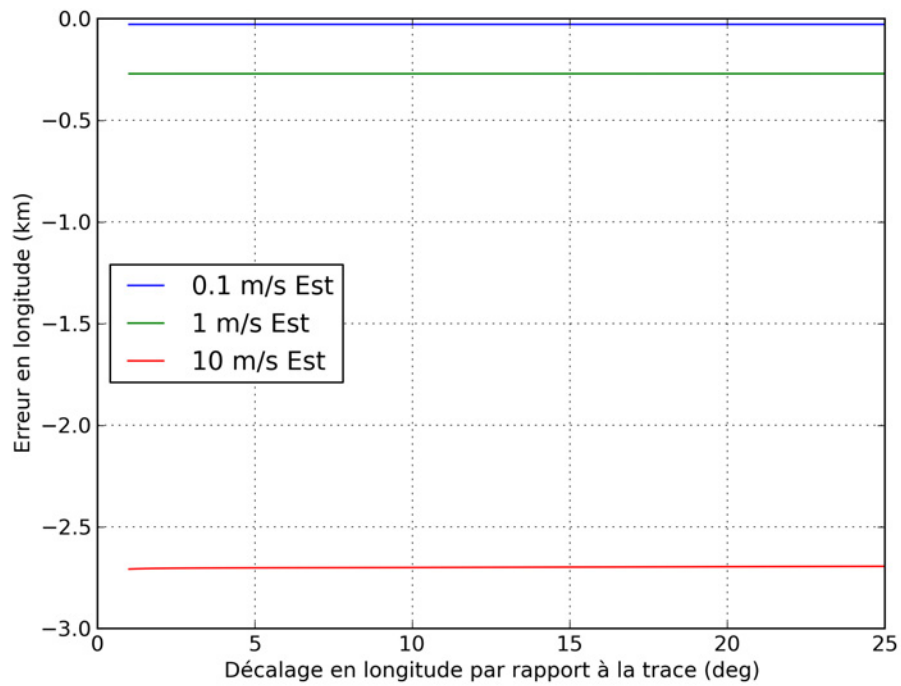
(a)



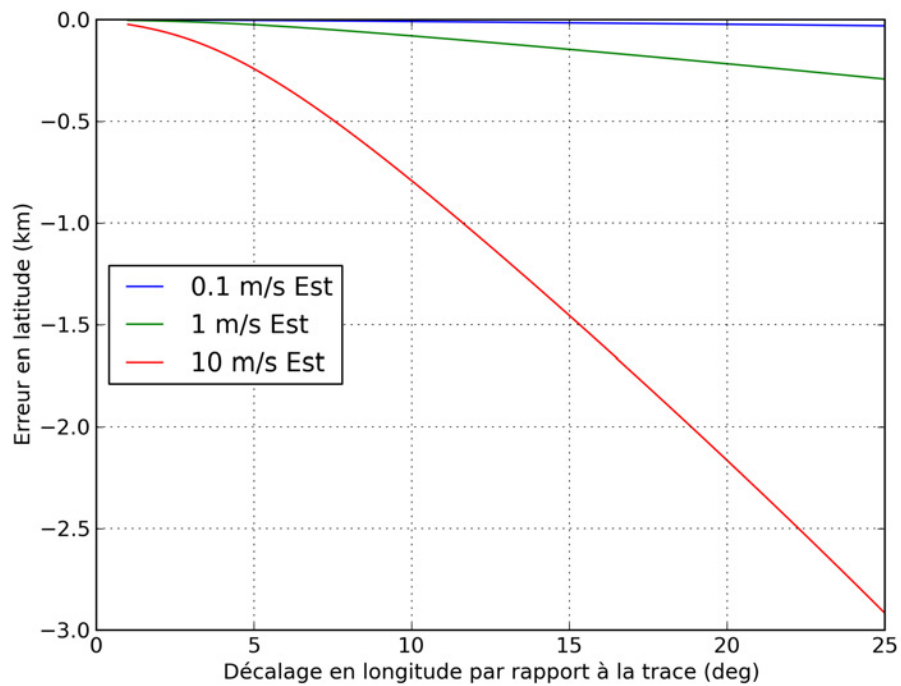
(b)

FIGURE 1.5.: Erreurs sur la longitude (a) et la latitude (b) pour une dérive de la fréquence d'émission de -0,1 Hz/min, -1 Hz/min, -10 Hz/min.

A.4. DÉPLACEMENT PENDANT LE PASSAGE

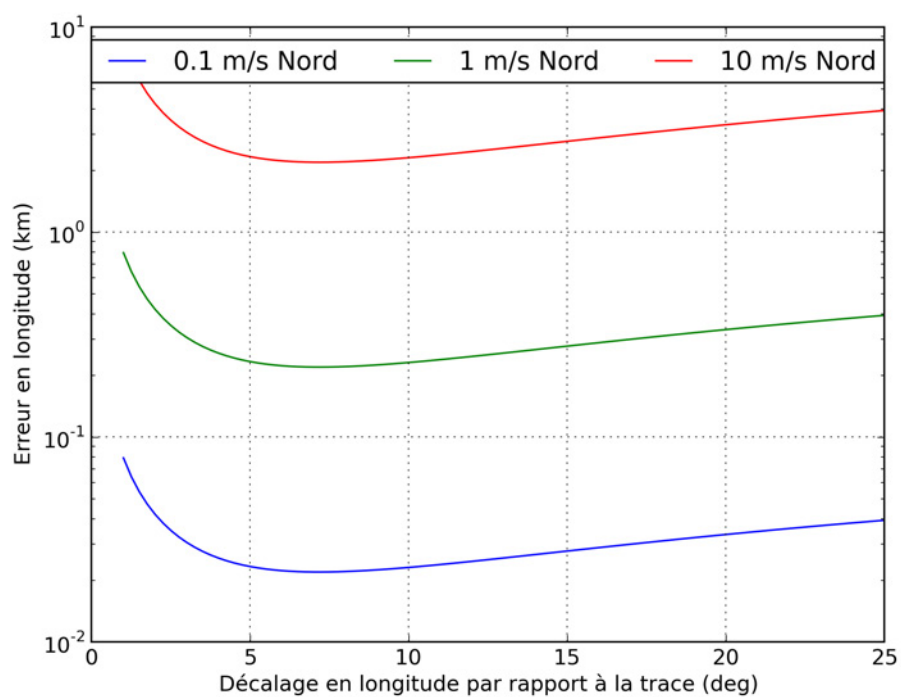


(a)

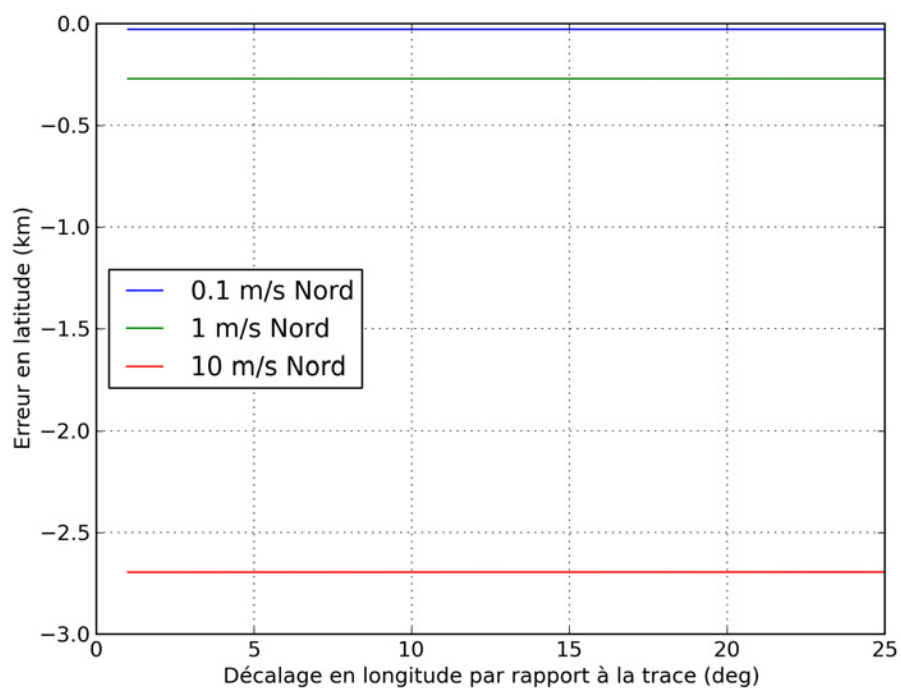


(b)

FIGURE 1.6.: Erreurs sur la longitude (a) et la latitude (b) avec un déplacement vers l'est de 0,1 m/s, 1 m/s et 10 m/s.



(a)



(b)

FIGURE 1.7.: Erreurs sur la longitude (a) et la latitude (b) avec un déplacement vers le nord de 0,1 m/s, 1 m/s et 10 m/s.

Annexe B

OUTILS DE POST-TRAITEMENT DES TRAJECTOIRES ARGOS

Améliorer la précision de localisation ARGOS est une problématique majeure des utilisateurs et plus particulièrement des suiveurs d'animaux. Cette communauté est très au fait du fonctionnement du système et des phénomènes grevant la précision de localisation ARGOS [Douglas et al., 2012] : faible nombre de messages par passage, mauvaise stabilité de l'oscillateur en cas de gradients de température, déplacement significatif de la balise pendant le passage pour les animaux migrants, courbe Doppler tronquée à cause de masques naturels (végétation ou eau)... Toutes ces connaissances ont été acquises par des études spécifiques sur la qualité de localisation ARGOS dont certaines des plus récentes sont [Kuhn et al., 2009, Witt et al., 2010, Costa et al., 2010].

En réponse, de nombreuses méthodes ont été proposées pour minorer l'erreur de localisation a posteriori et obtenir dans les cas difficiles des trajectoires plus lisses et réalistes. Certains utilisateurs se basent sur des techniques de décimation des points de la trajectoire. D'autres font usage de techniques de lissage et d'interpolation ou tirent partie de la formulation des modèles à espace d'état. L'ensemble de ces techniques est appliqué aux trajectoires ARGOS calculées par moindres carrés, autrement dit à des données affectées par des localisations images et où l'incertitude de localisation n'est pas toujours connue.

B.1. Filtrage par décimation

Dans cette première approche, les localisations sont filtrées en tenant compte des distances parcourues, des vitesses de déplacement ou des angles formés par trois positions successives. Ces critères sont souvent utilisés de conserve et leur emploi s'est complexifié au fil des années [McConnell et al., 1992, Keating, 1994, Austin et al., 2003, Freitas et al., 2008, Douglas et al., 2012].

Les utilisateurs ont bien compris que rejeter les positions en se basant uniquement sur la vitesse

moyenne calculée par différences finies entre deux localisations successives pousse à éliminer beaucoup de données de bonne qualité [Freitas et al., 2008]. Des vitesses aberrantes (*i.e.* supérieures à la vitesse de déplacement connue de l'animal) se produisent souvent pour deux localisations consécutives proches dans le temps et dont les erreurs respectives sont élevées. Par exemple, [Austin et al., 2003] ne supprime une position que si les quatre vitesses moyennes calculées avec les deux positions suivantes et précédentes sont aberrantes. Une position est également invalidée si la distance parcourue est statistiquement trop grande par rapport aux observations des sept derniers jours (au delà du 99^{ème} centile). Enfin, les localisations qui forment des pics sont mises de côté (angle aberrant).

Dans le même esprit, le *Douglas ARGOS Filter* (DAF) se veut une boîte à outil de référence pour traiter des trajectoires ARGOS par décimation. Elle est notamment accessible via la plateforme Movebank (www.movebank.org) qui offre la possibilité de stocker, partager et traiter avec le DAF des données ARGOS [Douglas et al., 2012]. Outre le fait d'utiliser des critères sur la distance, l'angle et la vitesse de l'animal, le DAF applique des filtres plus fins pour des espèces qui effectuent des migrations rapides sur de longues distances (les oiseaux en l'occurrence).

Au final, ces méthodes se révèlent particulièrement efficaces en matière de réduction d'erreur [Douglas et al., 2012]. Cependant, elles ne sont pas capables de corriger les points aberrants de la trajectoire et ne tiennent pas compte des erreurs d'observation. Par ailleurs, elles tendent à appauvrir l'échantillon de données et à augmenter l'intervalle de temps entre deux positions [Tremblay et al., 2009].

B.2. Lissage des trajectoires

L'algorithme de [Freeman et al., 1997] recalcule chaque position en la pondérant par la précédente et la suivante. Les poids sont fonction des classes de localisation des positions considérées. D'autres utilisateurs exploitent des courbes de Bézier ou des splines [Ferraroli et al., 2004, Tremblay et al., 2006] pour établir une trajectoire lissée. Cela offre en outre la possibilité d'interpoler les données de manière à obtenir des localisations régulièrement datées.

Plus originale, la méthode de [Tremblay et al., 2009] appliquée aux animaux marins reconstitue un trajet lissé en moyennant des trajectoires générées aléatoirement. Chaque trajectoire est une marche aléatoire dont le tirage de la direction et de l'amplitude à chaque pas est, d'une part, influencé par les localisations ARGOS et les erreurs associées, et, d'autre part, contraint par la bathymétrie, les côtes ou la température de surface de l'eau.

Quand il s'agit de déterminer des zones d'occupation animale, [Tougaard et al., 2008] propose une approche statistique. Plutôt que de chercher à corriger ou éliminer les localisations pour délimiter les habitats naturels, l'espace est discrétisé dans une grille. La probabilité d'occupation dans chaque portion de la grille est d'autant plus grande quelle est peuplée de localisations (en les pondérant par leurs erreurs estimées respectives).

B.3. Modélisation d'état

Les approches tirant partie de la formulation stochastique des modèles d'état [Royer and Lutcavage, 2008, Jonsen et al., 2005] offrent un cadre méthodologique plus propice pour réduire les incertitudes sans pour autant supprimer de localisations. Cette formulation permet d'estimer la position et l'erreur associée en adjoignant au modèle d'observation (ici une simple identité sur les localisations moindres carrés) une dynamique de mouvement a priori dans le but d'exploiter la cohérence temporelle des mobiles. Elle s'applique de manière transparente à l'interpolation de positions [Johnson et al., 2008, Patterson et al., 2010] et permet aussi de prendre en compte des contraintes sur la trajectoire.

Une frange d'utilisateurs considère la modélisation à espace d'état sophistiquée et compliquée à comprendre et à mettre en œuvre [Douglas et al., 2012, Tremblay et al., 2009]. Elle oblige les biologistes à s'associer avec des statisticiens, alors que les premiers jugent plus simples et intuitives les méthodes de décimation ou de lissage. D'aucuns soutiennent qu'il est préférable de nettoyer la trajectoire par décimation préalablement à toute modélisation et analyse.

Annexe C

FILTRES DE KALMAN ÉTENDU ET UNSCENTED

Les notations de cette annexe sont introduites dans le paragraphe 2.1.

C.1. Filtre étendu

Soient les vecteurs aléatoires x et y tels que $y = u(x)$ où $u : \mathbb{R}^{d_x} \rightarrow \mathbb{R}^{d_y}$. On suppose que la densité de probabilité p_x de x est connue avec $\hat{x} = E[x]$ et $P_x = cov(x)$. De manière générale, la moyenne $\hat{y} = E[y]$ se calcule avec

$$\hat{y} = \int_y y p_y(y) dy = E[u(x)] = \int_x u(x) p_x(x) dx \quad (C.1)$$

et la covariance $P_y = cov(y)$ avec

$$P_y = \int_y (y - \hat{y})(y - \hat{y})^T p_y(y) dy \quad (C.2)$$

$$= \int_x (u(x) - E[u(x)])(u(x) - E[u(x)])^T p_x(x) dx. \quad (C.3)$$

En pratique, les calculs peuvent se révéler complexes. Si on pose $u(x) \approx u(\hat{x}) + U(x - \hat{x})$ —valable au premier ordre—, il est possible d'approximer la moyenne $\hat{y}$ par $u(\hat{x})$ et la covariance P_y par $UP_x U^T$ où $U = \frac{\partial u}{\partial x}|_{\hat{x}}$.

Pour propager la covariance prédite et la covariance de l'innovation, le filtre de Kalman étendu linéarise donc les fonctions de transition f_{k+1} et d'observation g_{k+1} au voisinage de $\hat{x}_{k|k}$ et $\hat{x}_{k+1|k}$ respectivement [Sorenson, 1985]. La moyenne prédite $\hat{x}_{k+1|k}$ est construite par application directe de f_{k+1} sur $\hat{x}_{k|k}$ et la mesure prédite $\hat{z}_{k+1|k}$ par application de g_{k+1} sur $\hat{x}_{k+1|k}$.

Algorithme 9 Une récursion du filtre de Kalman étendu

Prédiction

État prédit $\hat{x}_{k+1|k} = f_{k+1}(\hat{x}_{k|k})$

Covariance prédite $P_{k+1|k} = F_{k+1}P_{k|k}F_{k+1}^T + Q_k$ avec $F_{k+1} = \left. \frac{\partial f_{k+1}}{\partial x_k} \right|_{\hat{x}_{k|k}}$

Correction

Mesure prédite $\hat{z}_{k+1|k} = g_{k+1}(\hat{x}_{k+1|k})$

Innovation $\tilde{z}_{k+1} = z_{k+1} - \hat{z}_{k+1|k}$

Covariance de l'innovation $S_{k+1} = G_{k+1}P_{k+1|k}G_{k+1}^T + R_{k+1}$ avec $G_{k+1} = \left. \frac{\partial g_{k+1}}{\partial x_{k+1}} \right|_{\hat{x}_{k+1|k}}$

Gain $K_{k+1} = P_{k+1|k}G_{k+1}^T S_{k+1}^{-1}$

État filtré $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{z}_{k+1}$

Covariance filtrée $P_{k+1|k+1} = (I - K_{k+1}G_{k+1})P_{k+1|k}$

Bien que relativement simple à mettre en place, le filtre étendu présente plusieurs désavantages majeurs : l'estimée n'est plus celle du minimum d'erreur quadratique moyen, la covariance de l'erreur d'estimation n'est pas consistante et il est possible que le filtre diverge si les fonctions sont fortement non-linéaires.

C.2. Filtre Unscented

Reprenons les variables x et y du paragraphe précédent. La transformée sigma détermine un ensemble de points $\{X_i\}_{i=0\dots 2d_x}$ qui permettent d'approximer les deux premiers moments de y après l'application de u . Les points sont choisis selon :

$$X_0 = \hat{x} \quad (C.4)$$

$$X_i = \hat{x} + \gamma(\sqrt{P_x})_i \quad i = 1 \dots d_x \quad (C.5)$$

$$X_{i+d} = \hat{x} - \gamma(\sqrt{P_x})_{i+d} \quad i = 1 \dots d_x \quad (C.6)$$

avec $(\sqrt{P_x})_i$ la $i^{\text{ème}}$ colonne de la matrice $\sqrt{P_x}$. Cette dernière est définie telle que $\sqrt{P_x}\sqrt{P_x}^T = P_x$ et s'obtient par une factorisation de Cholesky. On a en outre

$$\hat{x} = \sum_{i=0}^{2d_x} W_i^{(m)} X_i \text{ et } P_x = \sum_{i=0}^{2d_x} W_i^{(c)} (X_i - \hat{x})(X_i - \hat{x})^T. \quad (C.7)$$

où $\{W_i^{(m)}\}_{i=0\dots 2d_x}$ et $\{W_i^{(c)}\}_{i=0\dots 2d_x}$ sont les poids de la transformée. Leur calcul est détaillé dans [Van Der Merwe and Wan, 2001]. En posant $Y_i = u(X_i)$, la moyenne et la covariance de y sont approximées par

$$\hat{y} \approx \sum_{i=0}^{2d_x} W_i^{(m)} Y_i \text{ et } P_y \approx \sum_{i=0}^{2d_x} W_i^{(c)} (Y_i - \hat{y})(Y_i - \hat{y})^T. \quad (C.8)$$

Le filtre de Kalman Unscented [Julier and Uhlmann, 1997, Wan and Van Der Merwe, 2000] prend le parti d'approximer statistiquement les variables aléatoires en appliquant une transformée en

points sigma lors de la traversée de la fonction de transition f_{k+1} par x_k et de la fonction d'observation g_{k+1} par x_{k+1} .

Algorithme 10 Une récursion du filtre de Kalman Unscented

Prédiction

$$\text{Points sigma } X_{k|k} = \begin{bmatrix} \hat{x}_{k|k} & \hat{x}_{k|k} + \gamma\sqrt{P_{k|k}} & \hat{x}_{k|k} - \gamma\sqrt{P_{k|k}} \end{bmatrix}$$

$$\text{Points sigma transformés } X_{i,k+1|k}^* = f_{k+1}(X_{i,k|k}), \forall i = 0 \dots 2d_x$$

$$\text{État prédit } \hat{x}_{k+1|k} = \sum_{i=0}^{2d_x} W_i^{(m)} X_{i,k+1|k}^*$$

$$\text{Covariance prédite } P_{k+1|k} = \sum_{i=0}^{2d_x} W_i^{(c)} \left[X_{i,k+1|k}^* - \hat{x}_{k+1|k} \right] \left[X_{i,k+1|k}^* - \hat{x}_{k+1|k} \right]^T + Q_k$$

Correction

$$\text{Points sigma } X_{k+1|k} = \begin{bmatrix} \hat{x}_{k+1|k} & \hat{x}_{k+1|k} + \gamma\sqrt{P_{k+1|k}} & \hat{x}_{k+1|k} - \gamma\sqrt{P_{k+1|k}} \end{bmatrix}$$

$$\text{Points sigma transformés } Z_{i,k+1|k} = g_{k+1}(X_{i,k+1|k}), \forall i = 0 \dots 2d_x$$

$$\text{Mesure prédite } \hat{z}_{k+1|k} = \sum_{i=0}^{2d_x} W_i^{(m)} Z_{i,k+1|k}$$

$$\text{Innovation } \tilde{z}_{k+1} = z_{k+1} - \hat{z}_{k+1|k}$$

$$\text{Covariance de l'innovation } S_{k+1} = \sum_{i=0}^{2d_x} W_i^{(c)} \left[Z_{i,k+1|k} - \hat{z}_{k+1|k} \right] \left[Z_{i,k+1|k} - \hat{z}_{k+1|k} \right]^T + R_{k+1}$$

$$\text{Covariance croisée } P_{x_{k+1}\tilde{z}_{k+1}} = \sum_{i=0}^{2d_x} W_i^{(c)} \left[X_{i,k+1|k} - \hat{x}_{k+1|k} \right] \left[Z_{i,k+1|k} - \hat{z}_{k+1|k} \right]^T$$

$$\text{Gain } K_{k+1} = P_{x_{k+1}\tilde{z}_{k+1}} S_{k+1}^{-1}$$

$$\text{État filtré } \hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1} \tilde{z}_{k+1}$$

$$\text{Covariance filtrée } P_{k+1|k+1} = P_{k+1|k} - K_{k+1} S_{k+1} K_{k+1}^T$$

Les approximations de la moyenne et de la covariance sont valables au second ordre de leur développement de Taylor. Elles permettent d'assurer la consistance de la covariance de l'erreur d'estimation.

TABLE DES FIGURES

1.	Exemples de terminaux ARGOS. (a) Pigeon ramier équipé d'une balise de 17 grammes, et (b) flotteur sub-surface réalisant des profils de température et salinité des océans.	4
2.	Plans d'orbite de la constellation ARGOS en avril 2013 affichés selon les deux heures locales du nœud ascendant (points de l'orbite où le satellite traverse le plan de l'écliptique du sud vers le nord). Ce sont aussi les heures locales de passage des satellites au dessus d'un même point de la terre.	5
1.1.	Pour un point M de l'espace, on affiche sa projection M' sur le plan équatorial $\vec{XO}\vec{Y}$ (c) et sur le plan $\vec{VO}\vec{Z}$ (d). Les coordonnées cartésiennes sont données par le triplet (x, y, z) et les coordonnées géographiques par (λ, ϕ, h) . L'altitude h est la longueur du segment MI . La latitude ϕ est l'angle entre la normale à l'ellipsoïde passant par M et le plan équatorial. La latitude paramétrique Θ est l'angle entre ce même plan et le segment OM	11
1.2.	Le balise $\vec{B}$ avec la vitesse $\vec{V}_B$ est observée par le satellite $\vec{S}$ avec la vitesse $\vec{V}_S$. Le vecteur radial balise-satellite est $\vec{r}$	12
1.3.	(a) Différence $f_r - f_i$ et (b) dérivée en fréquence en fonction du temps pour différentes distances de la balise à la trace sub-satellite (en degrés le long de l'équateur). L'origine du temps est prise en début de passage et on ne tient pas compte du cercle de visibilité.	14
1.4.	Cartographie des résidus (en log) pour un passage satellite d'un seul message (a) et de plus de deux messages (b) sans prise en compte du cercle de visibilité du satellite. Le résidu est la somme des écarts entre les fréquences réellement reçues et leurs valeurs théoriques. La position réelle de la plateforme est le point violet. La courbe noire représente la trace au sol du satellite (déplacement nord-sud). Les étoiles marquent les positions au sol du satellite au moment de la réception des messages.	16
1.5.	Les positions nominale et image se situent aux intersections des cônes solutions avec l'ellipsoïde terrestre.	17
1.6.	Écarts-types des erreurs de localisation en fonction de la distance à la trace (par Hertz de bruit à 1 sigma sur la fréquence de réception).	20
1.7.	Valeur de $\sqrt{\sigma_\lambda^2 + \sigma_\phi^2}$ en fonction de la distance à la trace et du nombre de messages reçus (par Hertz de bruit à 1 sigma sur la fréquence de réception).	21

Table des figures

1.8.	Valeur de $\sqrt{\sigma_\lambda^2 + \sigma_\phi^2}$ en fonction de la distance à la trace et de la distribution dans le passage (par Hertz de bruit à 1 sigma sur la fréquence de réception). On considère quatre mesures réparties uniformément selon plusieurs configurations : sur le passage entier, les deux premiers tiers, la première moitié, le premier tiers et le tiers du milieu.	21
1.9.	Vue du dessus. L'erreur est la distance entre la localisation ARGOS et la position GPS interpolée à la même date, c'est-à-dire le module du vecteur $\vec{D}$. L'erreur tangentielle est la projection orthogonale de $\vec{D}$ sur la droite joignant les deux positions GPS successives encadrant la position GPS interpolée, c'est-à-dire le module du vecteur $\vec{T}$	28
1.10.	Erreurs de localisation des passages à plus de 4 messages pour l'ajustement moindres carrés.	33
1.11.	Erreurs de localisation des passages à 2 ou 3 messages pour l'ajustement moindres carrés.	34
1.12.	Trajectoire d'éléphant de mer effectuant un aller-retour depuis les îles Kerguelen (l'aller est la partie basse du trajet).	35
1.13.	Extraits de trajectoire d'éléphant de mer pour le traitement moindres carrés. . . .	36
2.1.	Graphique de dépendance d'un système à sauts markoviens avec m l'état discret (mode), x l'état continu et z la mesure. L'état complet du système est (x, m) . Les flèches bleues sont associées aux probabilités de transition $P(m_k m_{k-1})$, les rouges aux modèles de transition $p(x_k x_{k-1}, m_k)$ et les vertes aux modèles d'observation $p(z_k x_k, m_k)$	44
2.2.	Arbre des séquences possibles entre l'instant initial et l'instant $k = 2$ pour trois modèles. Sur le graphique, les densités associées aux séquences de modes $(1, 3, 3)$ et $(3, 2, 1)$ sont données. La densité globale de l'état à $k = 2$ est la mixture des densités associées à chaque branche.	46
2.3.	Décimation des branches de l'arbre des possibles : on ne conserve qu'un sous ensemble des termes de la mixture globale sur l'état (ici trois).	47
2.4.	Fusion des branches : on approxime par une densité plus simple q l'ensemble des densités dont la sous-séquence de modes $m_{k-n+1:k}$ est identique (ici $n = 2$). On obtient ainsi un arbre plus petit que celui de Fig. 2.2.	48
2.5.	Graphique de dépendance d'un système à sauts markoviens se propageant en temps inverse. Les flèches bleues sont associées aux probabilités de transition à rebours $P(m_k m_{k+1})$, les rouges aux modèles de transition à rebours $p(x_k x_{k+1}, m_{k+1})$ et les vertes aux modèles d'observation $p(z_k x_k, m_k)$	55
3.1.	Exemple de trajectoire générée aléatoirement avec des marches aléatoires corrélées.	86
3.2.	Comparaison du filtre IMM et des lisseurs (de haut en bas) : (a) RMSE sur la position, (b) RMSE sur la vitesse et (c) probabilité de mauvaise détection du mode actif.	87
3.3.	Exemple de trajectoire générée aléatoirement avec des mouvements coordonnés.	89
3.4.	Comparaison des RMSE moyennées sur le temps du filtre IMM et des lisseurs à retard fixe : (a) position, (b) vitesse, (c) vitesse angulaire et (d) probabilité de mauvaise détection du mode actif.	90

Table des figures

3.5. Exemple de trajectoire aléatoirement générée avec des mouvements coordonnés .	91
3.6. Comparaison du filtre IMMHO et des lisseurs (de haut en bas) : (a) RMSE sur la position, (b) RMSE sur la vitesse et (c) probabilité empirique de mauvaise détection du mode actif (toutes les courbes des lisseurs sont confondues).	93
4.1. Répartition des localisations en fonction du nombre de messages reçus (normalisé par rapport au nombre de localisations calculées par moindres carrés (MC)) . . .	106
4.2. Comparaison des erreurs empiriques de localisation des passages à plus de 4 messages entre l'ajustement moindres carrés et le filtre multi-modèle.	107
4.3. Comparaison des erreurs empiriques de localisation des passages à 2 ou 3 messages entre l'ajustement moindres carrés et le filtre multi-modèle.	108
4.4. Erreurs empiriques de localisation moyennes en fonction de la distance à la trace.	109
4.5. Comparaison des erreurs empiriques de localisation obtenues avec le filtre multi-modèle pour des passages à un message, et des erreurs empiriques obtenues par moindres carrés avec deux ou trois messages.	110
4.6. Trajectoire d'éléphant de mer pour le traitement moindres carrés et le filtre multi-modèle.	111
4.7. Extraits de trajectoire d'éléphant de mer pour le traitement moindres carrés et le filtre multi-modèle (localisations avec deux messages ou plus seulement).	112
4.8. Comparaison des erreurs empiriques de localisation des passages à 4 messages ou plus entre le filtre et le lisseur multi-modèles.	113
4.9. Comparaison des erreurs de localisation empiriques des passages à 2 ou 3 messages entre le filtre et le lisseur multi-modèles.	114
4.10. Comparaison des erreurs de localisation des passages à 1 message entre le filtre et le lisseur multi-modèles.	115
4.11. Erreurs de localisation moyennes en fonction de la distance à la trace.	116
4.12. Trajectoire d'éléphant de mer pour le filtre et le lisseur multi-modèles.	117
4.13. Extraits de trajectoire d'éléphant de mer pour le filtre et le lisseur multi-modèles.	118
1.1. Écarts-types des erreurs de localisation en fonction de la distance à la trace (par Hertz de bruit à 1 sigma sur la fréquence de réception) quand l'altitude est estimée.	124
1.2. Pente de la droite (km/km) donnant l'erreur en longitude λ et latitude ϕ en fonction du biais sur l'altitude	124
1.3. Erreur en longitude (km) pour des biais sur la fréquence d'émission de 10 Hz, 50 Hz et 100 Hz. Le comportement est quadratique.	125
1.4. Pente de la droite (km/Hz) donnant l'erreur en latitude en fonction du biais sur la fréquence d'émission.	126
1.5. Erreurs sur la longitude (a) et la latitude (b) pour une dérive de la fréquence d'émission de -0,1 Hz/min, -1 Hz/min, -10 Hz/min.	127
1.6. Erreurs sur la longitude (a) et la latitude (b) avec un déplacement vers l'est de 0,1 m/s, 1 m/s et 10 m/s.	128
1.7. Erreurs sur la longitude (a) et la latitude (b) avec un déplacement vers le nord de 0,1 m/s, 1 m/s et 10 m/s.	129

LISTE DES TABLEAUX

1.1. Classes de localisation ARGOS	25
1.2. Liste et caractéristiques des plateformes du jeu de données	27
1.3. Écart de temps maximum avec la plus proche position GPS et nombre de localisations ARGOS sélectionnées pour le calcul d'erreur	29
1.4. Statistiques sur les passages satellites (valeurs moyennes)	29
1.5. Statistiques sur les trajectoires calculées (valeurs moyennes)	30
2.1. Principales quantités utilisées par l'IMM	50
3.1. Principales quantités manipulées par le lisseur IMM-RTS (en plus de celles de l'IMM de Tab. 2.1)	70
3.2. Notations spécifiques de l'IMMHO	73
3.3. Valeurs moyennées sur le temps de la RMSE sur la position (m), la RMSE sur la vitesse ($m.s^{-1}$) et la probabilité de mauvaise détection.	88
3.4. Paramètres de modèles	90
3.5. Valeurs moyennées sur le temps de la RMSE sur la position (m), la RMSE sur la vitesse ($m.s^{-1}$) et la probabilité de mauvaise détection.	92
4.1. Nombre de localisations calculées par l'ajustement moindres carrés et le filtre multi-modèle	99
4.2. Nombre de localisations du filtre multi-modèle sélectionnées pour le calcul d'erreur	99
4.3. Rapport de l'erreur tangentielle moyenne sur le module moyen de l'erreur.	101
4.4. Probabilités empiriques que la localisation réelle se situe dans l'ellipse d'erreur centrée sur la localisation estimée associée.	102
4.5. Rapport de l'erreur tangentielle sur le module de l'erreur	104
4.6. Probabilités empiriques que la localisation réelle se situe dans l'ellipse d'erreur centrée sur la localisation estimée associée.	105

BIBLIOGRAPHIE

- [Ackerson and Fu, 1970] Ackerson, G. and Fu, K. (1970). On state estimation in switching environments. *Automatic Control, IEEE Transactions on*, 15(1) :10–17.
- [Austin et al., 2003] Austin, D., McMillan, J., and Bowen, W. (2003). A three-stage algorithm for filtering erroneous argos satellite locations. *Marine Mammal Science*, 19(2) :371–383.
- [Averbuch et al., 1991] Averbuch, A., Itzikowitz, S., and Kapon, T. (1991). Radar target tracking-Viterbi versus IMM. *Aerospace and Electronic Systems, IEEE Transactions on*, 27(3) :550–563.
- [Bar-Shalom and Chen, 2005] Bar-Shalom, Y. and Chen, H. (2005). IMM estimator with out-of-sequence measurements. *IEEE Transactions on Aerospace and Electronic Systems*, 41(1) :90–98.
- [Bar-Shalom and Li, 1993] Bar-Shalom, Y. and Li, X. (1993). *Estimation and Tracking : Principles, Techniques and Software*. YBS Eds., Norwood.
- [Bar-Shalom et al., 2001] Bar-Shalom, Y., Li, X., and Kirubarajan, T. (2001). *Estimation with applications to tracking and navigation : theory algorithms and software*. Wiley-Interscience.
- [Bar-Shalom et al., 2011] Bar-Shalom, Y., Willett, P., and Tian, X. (2011). Tracking and data fusion. *A Handbook of Algorithms*. Yaakov Bar-Shalom.
- [Barber, 2005] Barber, D. (2005). Expectation correction for smoothing in switching linear gaussian state space models. *Journal of Machine Learning Research*, 7 :2515–2540.
- [Barber, 2012] Barber, D. (2012). *Bayesian Reasoning and Machine Learning*. Cambridge University Press.
- [Benhamou et al., 2011] Benhamou, S., Sudre, J., Bourjea, J., Ciccione, S., De Santis, A., and Luschi, P. (2011). The role of geomagnetic cues in green turtle open sea navigation. *PLoS ONE*, 6(10) :e26672.
- [Blom, 1984] Blom, H. (1984). An efficient filter for abruptly changing systems. In *Decision and Control, 1984. The 23rd IEEE Conference on*, volume 23, pages 656–658. IEEE.
- [Blom and Bar-Shalom, 1988] Blom, H. and Bar-Shalom, Y. (1988). The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. on Automatic Control*, 33(8) :780–783.
- [Boers and Driessen, 2003] Boers, Y. and Driessen, J. (2003). Interacting multiple model particle filter. *IEE Proceedings on Radar, Sonar and Navigation*, 150(5) :344 – 349.

- [Chen and Tugnait, 2000] Chen, B. and Tugnait, J. (2000). Interacting Multiple Model Fixed-Lag Smoothing Algorithm for Markovian Switching Systems. *IEEE Trans. on Aerospace and Electronic Systems*, 36(1) :243–250.
- [CLS, 2011] CLS (2011). *Argos User Manual*. www.argos-system.org.
- [Codling et al., 2008] Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of The Royal Society Interface*, 5(25) :813–834.
- [Cork and Walker, 2007] Cork, L. and Walker, R. (2007). Sensor fault detection for UAVs using a nonlinear dynamic model and the IMM-UKF algorithm. In *IEEE Conf. on Information, Decision and Control (IDC'07)*, pages 230–235, Adelaide, Australia.
- [Costa et al., 2010] Costa, D. P., Robinson, P. W., Arnould, J. P. Y., Harrison, A., Simmons, S. E., Hassrick, J. L., Hoskins, A. J., Kirkman, S. P., Oosthuizen, H., Villegas-Amtmann, S., and Crocker, D. E. (2010). Accuracy of ARGOS locations of pinnipeds at-Sea estimated using fastloc GPS. *PLoS ONE*, 5(1) :e8677.
- [Davis, 1975] Davis, P. J. (1975). *Interpolation and approximation*. Dover Publications.
- [Domeier and Nasby-Lucas, 2013] Domeier, M. L. and Nasby-Lucas, N. (2013). Two-year migration of adult female white sharks (*carcharodon carcharias*) reveals widely separated nursery areas and conservation concerns. *Animal Biotelemetry*, 1(1) :2.
- [Doucet et al., 2001] Doucet, A., Gordon, N. J., and Krishnamurthy, V. (2001). Particle filters for state estimation of jump markov linear systems. *IEEE Transactions on Signal Processing*, 49(3) :613–624.
- [Douglas et al., 2012] Douglas, D., Weinzierl, R., C Davidson, S., Kays, R., Wikelski, M., and Bohrer, G. (2012). Moderating argos location errors in animal tracking data. *Methods in Ecology and Evolution*.
- [Dragon et al., 2012] Dragon, A., BarHen, A., Monestiez, P., and Guinet, C. (2012). Horizontal and vertical movements as predictors of foraging success in a marine predator. *Marine Ecology Progress Series*, 447 :243–257.
- [Dragon et al., 2010] Dragon, A., Monestiez, P., Bar-Hen, A., and Guinet, C. (2010). Linking foraging behaviour to physical oceanographic structures : Southern elephant seals and mesoscale eddies east of kerguelen islands. *Progress In Oceanography*, 87(1-4) :61–71.
- [Fancy et al., 1988] Fancy, S. G., Pank, L. F., Douglas, D. C., Curby, C. H., and Garner, G. W. (1988). Satellite telemetry : A new tool for wildlife research and management,. Technical report.
- [Ferraroli et al., 2004] Ferraroli, S., Georges, J., Gaspar, P., and Maho, Y. (2004). Where leather-back turtles meet fisheries. *Nature*, 429 :521–522.
- [Forney, 1973] Forney, G.D., J. (1973). The Viterbi algorithm. *Proceedings of the IEEE*, 61(3) :268 – 278.
- [Freeman et al., 1997] Freeman, A., Nicholls, D., Wilson, K., and Bartle, J. (1997). Radio-and satellite-tracking westland petrels *procellaria westlandica*.
- [Freitas et al., 2008] Freitas, C., Lydersen, C., Fedak, M., and Kovacs, K. (2008). A simple new algorithm to filter marine mammal argos locations. *Marine Mammal Science*, 24(2) :315–325.
- [Gerdan and Deakin, 1999] Gerdan, G. and Deakin, R. (1999). Transforming cartesian coordinates x, y, z to geographical coordinates ϕ, λ, h . *Australian surveyor*, 44(1) :55–63.

- [Gordon et al., 1993] Gordon, N. J., Salmond, D. J., and Smith, A. F. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140, pages 107–113. IET.
- [Gustafsson, 2000] Gustafsson, F. (2000). *Adaptive Filtering and Change Detection*. John Wiley & Sons.
- [Harris et al., 1990] Harris, R., Fancy, S., Douglas, D., Garner, G., Amstrup, S., McCabe, T., and Pank, L. (1990). Tracking wildlife by satellite : current systems and performance.
- [Hawkes et al., 2011] Hawkes, L. A., Balachandran, S., Batbayar, N., Butler, P. J., Frappell, P. B., Milsom, W. K., Tseveenmyadag, N., Newman, S. H., Scott, G. R., Sathiyaselvam, P., Takekawa, J. Y., Wikelski, M., and Bishop, C. M. (2011). The Trans-Himalayan flights of Bar-Headed geese (*Anser indicus*). *Proceedings of the National Academy of Sciences*, 108(23) :9516–9519.
- [Helmick et al., 1995] Helmick, R., Blair, W., and Hoffman, S. (1995). Fixed-interval smoothing for Markovian switching systems. *IEEE Trans. on Information Theory*, 41(6) :1845–1855.
- [Helmick et al., 1996] Helmick, R., Blair, W., and Hoffman, S. (1996). One-step fixed-lag smoothers for Markovian switching systems. *IEEE Trans. on Automatic Control*, 41(7) :1051–1056.
- [Iridium Communications Inc., 2012] Iridium Communications Inc. (2012). www.iridium.com.
- [Jarvis et al., 2008] Jarvis, A., Reuter, H., Nelson, A., and Guevara, E. (2008). Hole-filled srtm for the globe version 4. Available from the CGIAR-SXI SRTM 90m database : <http://srtm.csi.cgiar.org>.
- [Jilkov et al., 2002] Jilkov, V., Li, X., and Lu, L. (2002). Performance enhancement of the IMM estimation by smoothing. In *IEEE Int. Conf. on Information Fusion (FUSION'2002)*, volume 1, pages 713–720, Annapolis, MD.
- [Johnson et al., 2008] Johnson, D., London, J., Lea, M., and Durban, J. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5) :1208–1215.
- [Jonsen et al., 2005] Jonsen, I., Flemming, J., and Myers, R. (2005). Robust state-space modeling of animal movement data. *Ecology*, 86(11) :2874–2880.
- [Julier, 2002] Julier, S. (2002). The scaled unscented transformation. In *American Control Conference, 2002. Proceedings of the 2002*, volume 6, pages 4555–4559. IEEE.
- [Julier and Uhlmann, 1997] Julier, S. and Uhlmann, J. (1997). A new extension of the Kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, volume 3, page 26. Citeseer.
- [Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D) :35–45.
- [Keating, 1994] Keating, K. A. (1994). An alternative index of satellite telemetry location error. *The Journal of wildlife management*, pages 414–421.
- [Kelley, 1999] Kelley, C. T. (1999). *Iterative Methods for Optimization*. Number 18. SIAM Frontiers in Applied Mathematics.
- [Kitagawa, 1994] Kitagawa, G. (1994). The two-filter formula for smoothing and an implementation of the gaussian-sum smoother. *Annals of the Institute of Statistical Mathematics*, 46(4) :605–623.

- [Koch, 2000] Koch, W. (2000). Fixed-interval retrodiction approach to Bayesian IMM-MHT for maneuvering multiple targets. *IEEE Transactions on Aerospace and Electronic Systems*, 36(1) :2–14.
- [Kuhn et al., 2009] Kuhn, C., Johnson, D., Ream, R., and Gelatt, T. (2009). Advances in the tracking of marine species : using GPS locations to evaluate satellite track data and a continuous-time movement model. *Marine Ecology Progress Series*, 393 :97–109.
- [Li and Jilkov, 2003] Li, X. and Jilkov, V. (2003). Survey of maneuvering target tracking. Part I : Dynamic models. *IEEE Trans. on Aerospace and Electronic Systems*, 39(4) :1333–1364.
- [Li and Jilkov, 2005] Li, X. and Jilkov, V. (2005). Survey of Maneuvering Target Tracking. Part V : Multiple-Model Methods. *IEEE Trans. on Aerospace and Electronic Systems*, 41(4) :1255–1321.
- [Lopez, 2009] Lopez, R. (2009). Evolution algorithmique de la localisation des balises ARGOS. Master’s thesis, Institut National des Sciences Appliquées de Toulouse.
- [Lopez and Danès, 2012] Lopez, R. and Danès, P. (2012). Exploiting Rauch-Tung-Striebel formulae for IMM-based smoothing of Markovian switching systems. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012* , pages 3953–3956. IEEE.
- [Lopez et al., 2010] Lopez, R., Danès, P., and Royer, F. (2010). Extending the imm filter to heterogeneous-order state space models. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 7369–7374. IEEE.
- [McConnell et al., 1992] McConnell, B., Chambers, C., Nicholas, K., and Fedak, M. (1992). Satellite tracking of grey seals (*Halichoerus grypus*). *Journal of Zoology*, 226(2) :271–282.
- [McGinnity and Irwin, 2000] McGinnity, S. and Irwin, G. (2000). Multiple model bootstrap filter for maneuvering target tracking. *Aerospace and Electronic Systems, IEEE Transactions on*, 36(3) :1006–1012.
- [Mesot and Barber, 2009] Mesot, B. and Barber, D. (2009). A simple alternative derivation of the expectation correction algorithm. *IEEE Signal Processing Letters*, 16(2) :121–124.
- [Morelande and Ristic, 2008] Morelande, M. and Ristic, B. (2008). Smoothed state estimation for nonlinear Markovian switching systems. *IEEE Trans. Aerospace and Electronic Systems*, 44(4) :1309–1325.
- [Nadarajah et al., 2012] Nadarajah, N., Tharmarasa, R., McDonald, M., and Kirubarajan, T. (2012). IMM forward filtering and backward smoothing for maneuvering target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 48(3) :2673–2678.
- [National Geospatial-Intelligence Agency, 1997] National Geospatial-Intelligence Agency (1997). Department of Defense World Geodetic System 1984, Its Definition and Relationships With Local Geodetic Systems, Third Edition. Technical report.
- [Patterson et al., 2010] Patterson, T. A., McConnell, B. J., Fedak, M. A., Bravington, M. V., and Hindell, M. A. (2010). Using GPS data to evaluate the accuracy of state-space methods for correction of argos satellite telemetry error. *Ecology*, 91(1) :273–285.
- [Pedersen, 2007] Pedersen, M. (2007). *Hidden Markov models for geolocation of fish*. PhD thesis, M. Sc. thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Kgs. Lyngby, Denmark. IMM Publication.

- [Pitre et al., 2005] Pitre, R., Jilkov, V., and Li, X. (2005). A comparative study of multiple-model algorithms for maneuvering target tracking. In *Proc. SPIE*, volume 5809, pages 549–560.
- [Rapoport and Oshman, 2004] Rapoport, I. and Oshman, Y. (2004). A new estimation error lower bound for interruption indicators in systems with uncertain measurements. *Information Theory, IEEE Transactions on*, 50(12) :3375–3384.
- [Rauch et al., 1965] Rauch, H., Tung, F., and Striebel, C. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8) :1445–1450.
- [Robinson et al., 2012] Robinson, P. W., Costa, D. P., Crocker, D. E., Gallo-Reynoso, J. P., Champagne, C. D., Fowler, M. A., Goetsch, C., Goetz, K. T., Hassrick, J. L., Hückstädt, L. A., Kuhn, C. E., Maresh, J. L., Maxwell, S. M., McDonald, B. I., Peterson, S. H., Simmons, S. E., Teutschel, N. M., Villegas-Amtmann, S., and Yoda, K. (2012). Foraging behavior and success of a mesopelagic predator in the northeast pacific ocean : Insights from a Data-Rich species, the northern elephant seal. *PLoS ONE*, 7(5) :e36728.
- [Rong Li and Bar-Shalom, 1993] Rong Li, X. and Bar-Shalom, Y. (1993). Design of an interacting multiple model algorithm for air traffic control tracking. *IEEE Transactions on Control Systems Technology*, 1(3) :186–194.
- [Royer and Lutcavage, 2008] Royer, F. and Lutcavage, M. (2008). Filtering and interpreting location errors in satellite telemetry of marine animals. *Journal of Experimental Marine Biology and Ecology*, 359(1) :1–10.
- [Särkkä and Hartikainen, 2010] Särkkä, S. and Hartikainen, J. (2010). On Gaussian optimal smoothing of non-linear state space models. *IEEE Trans. on Automatic Control*, 55(8) :1938–1941.
- [Seegar et al., 1996] Seegar, W., Cutchis, P., Fuller, M., Surer, J., Bhatnagar, V., and Wall, J. (1996). Fifteen years of satellite tracking development and application to wildlife research and conservation. *Johns Hopkins APL Technical Digest*, 17(4) :401–411.
- [Simon, 2010] Simon, D. (2010). Kalman filtering with state constraints : a survey of linear and nonlinear algorithms. *IET Control Theory Applications*, 4(8) :1303 –1318.
- [Sorenson, 1985] Sorenson, H. W. (1985). *Kalman filtering : theory and application*. IEEE Press.
- [Thygesen et al., 2009] Thygesen, U. H., Pedersen, M. W., and Madsen, H. (2009). Geolocating fish using hidden markov models and data storage tags. *Tagging and Tracking of Marine Animals with Electronic Devices*, page 277–293.
- [Toledo-Moreo et al., 2006] Toledo-Moreo, R., Zamora-Izquierdo, M., and Gomez-Skarmeta, A. (2006). IMM-EKF based road vehicle navigation with low cost GPS/INS. In *IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems (MFI'2006)*, pages 433–438, Heidelberg, Germany.
- [Tougaard et al., 2008] Tougaard, J., Teilmann, J., and Tougaard, S. (2008). Harbour seal spatial distribution estimated from argos satellite telemetry : overcoming positioning errors. *Endangered species research*, 4(1-2) :113—122.
- [Tremblay et al., 2009] Tremblay, Y., Robinson, P. W., and Costa, D. P. (2009). A parsimonious approach to modeling animal movement data. *PLoS ONE*, 4(3) :e4711.
- [Tremblay et al., 2006] Tremblay, Y., Shaffer, S., Fowler, S., Kuhn, C., McDonald, B., Weise, M., Bost, C., Weimerskirch, H., Crocker, D., Goebel, M., et al. (2006). Interpolation of animal tracking data in a fluid environment. *Journal of Experimental Biology*, 209(1) :128.

- [Tugnait, 1982] Tugnait, J. (1982). Detection and estimation for abruptly changing systems. *Automatica*, 18(5) :607–615.
- [Tugnait, 1983] Tugnait, J. (1983). Detection and identification of abrupt changes in linear systems. In *American Control Conference, 1983*, pages 960–965. IEEE.
- [Tugnait and Haddad, 1979] Tugnait, J. and Haddad, A. (1979). A detection-estimation scheme for state estimation in switching environments. *Automatica*, 15(4) :477–481.
- [USGS, 2010] USGS (cited 2010). GTOPO30. [Available online at <http://eros.usgs.gov/>].
- [Van Der Merwe and Wan, 2001] Van Der Merwe, R. and Wan, E. (2001). The square-root unscented kalman filter for state and parameter-estimation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, volume 6, pages 3461–3464. IEEE.
- [Wan and Van Der Merwe, 2000] Wan, E. and Van Der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158. IEEE.
- [Williams and Rasmussen, 2006] Williams, C. and Rasmussen, C. (2006). Gaussian processes for machine learning.
- [Witt et al., 2010] Witt, M., Åkesson, S., Broderick, A., Coyne, M., Ellick, J., Formia, A., Haysf, G., Luschi, P., Stroud, S., and Godleya, B. (2010). Assessing accuracy and utility of satellite-tracking data using Argos-linked Fastloc-GPS. *Animal Behaviour*, 80 :571e581.
- [Yuan et al., 2012] Yuan, T., Bar-Shalom, Y., Willet, P., Mozeson, E., Pollak, S., and Hardiman, D. (2012). A multiple IMM estimation approach with unbiased mixing for thrusting projectiles. *IEEE Trans. on Aerospace and Electronic Systems*, 48(4) :3250–3267.
- [Zarrouati, 1987] Zarrouati, O. (1987). *Trajectoires spatiales*. Centre National d'Études Spatiales.

[2013-07-17 10:28:39 +0200] [5bf8026]

