Bayesian Approaches to Trajectory Estimation in Maritime Surveillance

Melita Hadzagic



Department of Electrical & Computer Engineering McGill University Montreal, Canada

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

 \bigodot 2010 Melita Hadzagic

Abstract

In maritime surveillance, multisensor data differ to a great extent in their temporal resolution. Additionally, due to multi-level security and information management processing, many contact reports arrive hours after observations. This makes the contact report data usually available for batch processing. The dissimilar multi-source information environment results in contact reports with heteroscedastic and correlated errors (i.e. measurement errors characterized by normal probability distributions with non-constant and nondiagonal covariance matrices), while the obtained measurement errors may be relatively large. Hence, the appropriate choice of a trajectory estimation algorithm, which addresses the aforementioned issues of the surveillance data, will significantly contribute to increased awareness in the maritime domain.

This thesis presents two novel batch single ship trajectory estimation algorithms employing Bayesian approaches to estimation: (1) a stochastic linear filtering algorithm and (2) a curve fitting algorithm which employs Bayesian statistical inference for nonparametric regression. The stochastic linear filtering algorithm employs a combination of two stochastic processes, namely the Integrated Ornstein-Uhlenbeck process (IOU) and the random walk (RW), process to describe the ship's motion. The assumptions on linear modeling and bivariate Gaussian distribution of measurement errors allow for the use of Kalman filtering and Rauch-Tung-Striebel optimal smoothing. In the curve fitting algorithm, the trajectory is considered to be in the form of a cubic spline with an unknown number of knots in two-dimensional Euclidean plane of longitude and latitude. The function estimate is determined from the data which are assumed Gaussian distributed. A fully Bayesian approach is adopted by defining the prior distributions on all unknown parameters: the spline coefficients, the number and the locations of knots. The calculation of the posterior distributions is performed using Markov Chain Monte Carlo (MCMC) and reversible jump Markov sampling due to the varying dimensions of subspaces where the searches are performed. Both algorithms assume no knowledge about the ship motion model, however assuming standard ship maneuvers.

The quality of the estimated trajectories obtained by both algorithms is assessed using several simulated scenarios and evaluated statistically. The positional measurements, received at irregular time intervals are assumed to have heteroscedastic and correlated errors and available in batches. The performance evaluation includes the performance comparison of both algorithms with another batch stochastic optimization algorithm for trajectory estimation, i.e. the genetic algorithm (GA). The sensitivity analysis is carried out with respect to perturbations in parameters of the algorithms. The results show similar performance between the linear stochastic filtering algorithm and the Bayesian spline regression algorithm, while both algorithms show superiority over the GA-based trajectory fitting with respect to tracking accuracy, due to complete account for uncertainty. Batch data processing approach is confirmed to be more suitable in maritime surveillance than standard recursive approaches. The thesis demonstrates that for the accurate trajectory estimation it is crucial to completely account for uncertainty of measurements, especially if the measurements are characterized by heteroscedastic and correlated errors. The results of this thesis are useful as they facilitate selecting the appropriate approach to data processing in maritime surveillance applications, hence contribute to increased maritime domain awareness. These can also serve for selecting appropriate methods for data processing in dissimilar sensor and other environments in which data have large and heteroscedastic measurement errors.

Resumé

En surveillance maritime, les données multi-senseurs diffèrent, dans une large mesure, en terme de leur résolution temporelle. De plus, en raison de la gestion du traitement d'information de sécurité multi-niveau, plusieurs rapports de contact sont reçus des heures après leur observation. Ceci rend les données des rapports de contact disponible pour traitement en lot. L'information multi-source provenant d'environnements dissimilaires, les rapports de contact ont des erreurs hétéroscédastiques et corrélées (i.e., des erreurs de mesure caractérisées par une distribution de probabilité normale et une matrice de covariance non constante), ainsi qu'une erreur de mesure pouvant être relativement large. En conséquence, le choix approprié d'un algorithme d'estimation de trajectoire adressant les problèmes susmentionnés de données de surveillance contribuera significativement à accroître la perception de situation dans le domaine maritime.

Cette thèse présente deux nouveaux algorithmes pour l'estimation en lot de trajectoire de navire simple et employant une approche d'estimation bayésienne: (1) un algorithme de filtration linéaire stochastique, et (2) un algorithme de lissage de courbe réalisant une régression non-paramétrique par inférence statistique bayésienne. L'algorithme de filtration linéaire stochastique emploi la combinaison de deux processus stochastiques, c'est-à-dire le processus d'Ornstein-Uhlenbeck intégré (IOU) et le processus de marche aléatoire (RW), pour décrire le mouvement du navire. Les suppositions de modèle linéaire et de distribution gaussienne bivariée des erreurs de mesure permettent l'utilisation du filtre de Kalman et du lissage optimal de Rauch-Tung-Striebel. Dans l'algorithme de lissage de courbe, la trajectoire est considérée représentée par un spline cubique avec un nombre de noeuds inconnu dans le plan euclidien à deux dimensions, en longitude et latitude. L'estimation de fonction est déterminée depuis les données assumées de distribution gaussienne. Une approche pleinement bayésienne est adoptée en définissant la distribution a-priori sur tous les paramétres inconnus: les coefficients du spline, et le nombre et la location des noeuds. Le calcul des distributions a-posteriori est réalisé en utilisant une chaîne de Markov Monte Carlo (MCMC) et un échantillonage de Markov à saut réversible en raison du nombre variable de dimension des sous-espaces où les recherches sont réalisées. Les deux algorithmes ne supposent aucune connaissance sur le modèle de mouvement du navire, mais assument des manoeuvres navales standard.

La qualité des trajectoires estimées obtenues par les deux algorithmes est démontrée à l'aide de plusieurs scénarios simulés et évaluée statistiquement. Les mesures de position reçues à intervalle de temps irrégulier sont assumées être disponible en lot et avoir des erreurs hétéroscédastiques et corrélées. L'évaluation des performances inclue la comparaison des performances des deux algorithmes avec un autre algorithme d'optimisation en lot pour l'estimation de trajectoire, c'est-à-dire l'algorithme génétique (GA). L'analyse de sensibilité est aussi réalisée par rapport aux perturbations dans les paramètres des algorithmes. Les résultats montrent des performances similaires des deux algorithmes et leur supériorité par rapport au lissage de trajectoire basé sur l'algorithme génétique. L'approche de traitement en lot des données est confirmée être plus appropriée en surveillance maritime que les approches récursives standards. La thèse démontre aussi que, pour l'estimation précise de trajectoire, il est crucial de prendre en compte complètement l'incertitude des mesures, et en particulier lorsque ces mesures sont caractérisées par des erreurs hétéroscédastiques et corrélées. Les résultats de cette thèse sont utiles en ce qu'ils facilitent la sélection de l'approche appropriée au traitement des données dans les applications de surveillance maritime, et ainsi contribuent à accroître la perception de situation dans le domaine maritime.

Originality and Contributions

The work in this thesis has been carried out entirely by the doctoral student alone.

The original contributions of the thesis can be listed as follows:

- A novel batch single ship trajectory estimation algorithm is introduced. It assumes no knowledge about the ship motion model, while assuming standard ship maneuvers. As a batch processing algorithm, it addresses the availability of data for processing, an issue specific to maritime surveillance applications. The algorithm developed here is based on the algorithm found in Di Matteo et al. [1] which employs Bayesian inference to spline fitting using nonparametric regression model. However, the work of Di Matteo et al. [1] fails to accommodate explicitly for heteroscedastic and correlated errors which are inherent to the problem of trajectory estimation in maritime surveillance, as well as for decreasing values of predictor variables which occurs in case of ships' backward turning maneuvers. Additionally, Bayesian inference is performed for nonparametric regression models with two different choices of response and predictor variables. A complete account for the measurement uncertainty by this algorithm guarantees improved performance over other batch stochastic optimization algorithms for ship trajectory estimation based on error approximations such as the one reported in [2].
- Assessment of the robustness of the proposed algorithm by a sensitivity analysis is carried out. This includes verifying if the results of the Bayesian analysis remain unchanged in the presence of perturbation in parameters of the algorithm.
- The first time development, implementation and assessment in the open literature of the stochastic linear filtering algorithm which employs the Integrated Ornstein-

Uhlenbeck processes in conjunction with Kalman filtering and Rauch-Tung-Striebel optimal smoothing is provided. It is presented here that such motion modeling is possible only when the IOU process is used in the specific combination with the random walk process. The assessment of the IOU-based trajectory estimation algorithm with special focus on the importance of the modeling of measurement uncertainty was reported in Hadzagic and Michalska (2010) [3], [4].

- Sensitivity analysis of the linear stochastic filtering algorithm with respect to perturbations in model parameters is performed.
- Assessment of the efficiency of both algorithms for various ground truth scenarios in terms of a chosen performance measure is carried out. A separate analysis is performed for each nonparametric regression model.
- Comparative analysis of the Bayesian nonparametric regression spline based algorithm (both models) with the batch linear stochastic algorithm for trajectory estimation is carried out.
- Respective comparative analysis of the Bayesian nonparametric regression spline algorithm (both models) and the IOU-model based algorithm with another batch stochastic optimization algorithm, namely a genetic algorithm, used to form a trajectory from heteroscedastic data [3] is provided.
- Analysis and discussion of the results of the comparisons, conclusions about the novel algorithms are provided, and possible future research avenues are outlined.
- Development of an integrated software platform for maritime surveillance which includes the following submodules:

- a stochastic ground truth scenario generator for randomized ship trajectories in the dissimilar sensor environment,
- a Bayesian spline regression algorithm module, based on a modification of open source software, found here [5]. The modification involves injection of the heteroscedastic and correlated measurement error information and the rotation of coordinate system to accommodate maneuvers in x - y plane when the position in x-direction is decreasing.
- the IOU-KF-RTSS algorithm module,
- a GA-based trajectory estimation algorithm module, conceptually based on the one described in [2].

The above software is developed in C++ and consists of 6701 lines of code.

List of Publications

- M. Hadzagic, H. Michalska, and A. Jouan, IMM-JVC and IMM-JPDAF algorithms for closely maneuvering targets, *Proc. of the IEEE 36th Asilomar Conf. on Signals*, *Systems, and Computers*, 2002.
- 2. M. Hadzagic, Track-before-detect methods for tracking low-observable targets, Technical Report, Lockheed Martin Canada, 2005.
- M. Hadzagic, H. Michalska, and E. Lefebvre, Track-before-detect methods for tracking low-observable targets, Sensors and Transducers Magazine, spec. issue: Multisensor Data and Information Processing, 2005.
- 4. E. Lefebvre, M. Hadzagic, and É. Bossé, On quality of information in multi-source fusion environments, Advances and Challenges in Multisensor Data and Information Processing, Vol.8 NATO Security through Science Book Series: Information and Communication Security, 2007.
- M. Hadzagic, H. Michalska, and E. Lefebvre, Data fusion challenges and architectures in maritime surveillance, *REPARTI Colloquium on Distributed Intelligent* Systems, McGill, Montreal, QC, 2007.
- M. Hadzagic, Multi-Target Statistical Tracker: mathematical background, tech. rep., Lockheed Martin Canada-McGill University, 2007.
- 7. M. Hadzagic SimTrack v.1.0, sw package for ship track scenario generation, 2008.
- M. Hadzagic, H. Michalska, and D. Grenier, Integrated Ornstein-Uhlenbeck Process in Application to Ship Tracking, *Proc. of the 10th IASTED Conference on Signal* and Image Processing, Kailua-Kona, HI, 2008.

- E. Lefebvre, M. Hadzagic, and É. Bossé, Information Quality Assessment in Fusion Systems, Proc. of Cognitive Systems with Interactive Sensors, 16-18 November, Paris, France, 2009.
- M. Hadzagic and H. Michalska, A linear stochastic process and a genetic algorithm for ship trajectory modeling, Signal Processing, Proc. of SPIE Defence, Security and Sensing, Conference on Sensor Fusion, and Target Recognition XIX, April 5-9, Orlando, USA, 2010.
- M. Hadzagic and H. Michalska, Batch stochastic trajectory estimation using the Ornstein-Uhlenbeck process and a genetic algorithm, *submitted to the ELSEVIER Journal on Information Fusion, in review*, January 2010.
- 12. M. Hadzagic and H. Michalska, A Bayesian Approach to Trajectory Estimation, in preparation for submission to ISIF Journal of Information Fusion, 2010.

Acknowledgments

I would like to express my sincere thanks and gratitude to people and institutions that contributed to this thesis.

Firstly, I would like to thank Lockheed Martin Canada (LMC), especially their former R&D Department with Dr. Elisa Shahbazian, and NSERC Canada for their support in this research. Specifically, I would like to thank Dr. Eric Lefebvre, formerly of LMC, who encouraged me to embark on ships when I felt as if I was on a raft in the stormy Ph.D. ocean. The motivation and valuable background he has provided me with on this topic, have helped me greatly in this research. As a part of the Defense Research and Development Canada (DRDC) Data Fusion group I appreciate also having worked with Dr. Éloi Bossé and Dr. Dominic Grenier which allowed me to explore data fusion further than the Level 1.

I would like to express my sincere gratitude to my supervisor Dr. Hannah Michalska for her generous support towards the end of my studies when I needed it the most. I am very grateful for her constructive comments and her time.

I would like to thank Dr. Alain Vandal from McGill's Department of Mathematics and Statistics for his insightful advices on issues pertaining to research in the field of Bayesian statistical inference.

My deepest gratitude to Professor Peter Caines, who was always kind and available for questions and discussions when I was in need.

I would also like to thank my external reviewer, Dr. Pierre Valin, whose remarks and comments helped me improve the content of the thesis.

Thanks to Hugo Levasseur's generous assistance, the code used for numerical simulations was significantly improved. I appreciate Ljubinko Igrutinovic's help and always finding time to answer my C++ questions despite his personal difficulties. I also thank McGill University, Karen McGrath from McGill Student Counselling Services, and particularly Centre for Intelligent Machines (CIM) for providing a pleasant work environment for which the biggest credits go to Cynthia Davidson, Marlene Gray, Jan Binder, Patrick McLean, and also staff members Michael and Eric Jean-Marie. During my time at CIM, I had a chance to meet and befriend many wonderful people: Dr. Miguel Torres-Torriti, Dr. Alessio Salerno, Dr. Dany Dionne, Dr. Casey Lambert, Dr. Shawn Arseneau, and many others. Thank you all for the mutual support and fun together. Special thanks to Dany for helping with the translation of Abstract to French, and also Dr. Vladislave Ganine for proofreading a part of the manuscript. I am particularly grateful to my friend and colleague Shahram Tabandeh for valuable research discussions, constructive comments, and the proofreading. Thanks Shahram and Vlad for making my last year at McGill the enjoyable experience.

Many thanks to my friend Vojin Jarić for his omnipresent trans-Atlantic MSN moral and music support, and for pushing me to keep working when everybody was asleep.

Finally, this thesis could not have been done without generous support and love from my family. Mom and Dad, thank you for Everything! With you behind me, I was always able to go forward despite any difficulty. Thank you Sandra for all those "vamos Rafa" through cyberspace, your wits, and for always being here/there for me. The infinite gratitude goes to my Hugo for his constant encouragement, endless patience, love and support. Thank you for sharing your life with me at my hardest times.

To my parents.

In memory of a dear friend Ljubinko Igrutinovic.

Contents

1 Introduction

	1.1	Motivating Application	1
	1.2	Trajectory Estimation Algorithms for Multisensor Data Fusion-The Back-	
		ground	7
		1.2.1 Bayesian Approach to Linear/Nonlinear Filtering	13
		1.2.2 Bayesian Statistical Analysis for Trajectory Estimation	16
	1.3	Problem and Research Goals	20
	1.4	Literature Survey	22
	1.5	Approach and Contributions	28
2	Pro	blem Statement	33
	2.1	Characterization of the Measurement Data, Measurement Errors and the	
		Surveillance Region	33
	2.2	Characterization of Model Features and Estimation Goals	37
	2.3	The Linear Stochastic Bayesian Filtering Problem	37
	2.4	The Nonparametric Regression Spline Problem	39

3 Trajectory Estimation Using Linear Stochastic Filtering with Integrated

 $\mathbf{1}$

	Ornstein Uhlenbeck Process 42			42
	3.1	1 Development of the Ship Motion Model		
		3.1.1	Random Walk	48
		3.1.2	Link between the IOU Model and the RW Model	50
	3.2	Contin	nuous-Discrete Kalman Filter Recursion for the IOU Motion Process	52
		3.2.1	The Algorithm	54
	3.3	Rauch	-Tung-Striebel Fixed Optimal Smoothing	55
4	Bay	vesian \$	Statistical Inference Approach to Trajectory Estimation	60
	4.1	Metho	d Development	61
		4.1.1	Bayesian Model	62
		4.1.2	Bayesian Simulation: Reversible-Jump Metropolis-Hastings Markov	
			Chain MC	63
		4.1.3	The Bayesian Spline Regression Algorithm	65
		4.1.4	Convergence of the Algorithm	68
		4.1.5	Metropolis-Hastings	70
	4.2	Maneu	avers in $x - y$ Plane When x is Decreasing $\ldots \ldots \ldots \ldots \ldots \ldots$	70
5	AC	Genetic	e Algorithm for Trajectory Estimation	73
	5.1	Genet	ic Algorithms	73
	5.2	Trajec	tory Estimation Using a GA	74
6	Sun	nmary	of Simulation Results for Trajectory Estimation Algorithms	77
	6.1	Scenar	rio	77
	6.2	Choice	e of Performance Measure	79
	6.3	IOU-ŀ	KF-RTSS Algorithm Performance Evaluation	79

	6.4	Sensitivity Analysis in the IOU Parameters	
	6.5	IOU-KF-RTSS Algorithm Comparison with a Genetic Algorithm (GA) for	
		Trajectory Estimation	86
	6.6	Performance Evaluation of the Bayesian Regression Spline Algorithm	90
		6.6.1 Regression on $x - y$	90
		6.6.2 Regression on $t - x$ and $t - y$	93
	6.7	Sensitivity Analysis in Parameters of the Bayesian Regression Spline Algorithm	94
	6.8	On the Computational Note	97
7	Con	cluding Remarks	98
	7.1	General Summary	98
	7.2	Conclusions	100
	7.3	Future Research Avenues	101
A	The	Relationship of the Error Ellipse Measurement to a Gaussian Bivari-	
	ate	Distribution	103
	A.1	The Flat-Earth Approximation	109
Re	efere	nces	111

List of Figures

1.1	A snapshot of the RMP [6]. \ldots	2
1.2	Concept of Level 1 data fusion for a surveillance system	3
1.3	Possible estimated trajectories from dissimilar multisensor surveillance data.	4
1.4	(a) Heteroscedastic and (b) heteroscedastic and correlated measurement errors.	6
2.1	The positional measurement with elliptical area of uncertainty	35
3.1	Random Walk process with the initial course that is chosen randomly from	
	$U[0^0, 360^0]$ and $\theta \sim exp(1/\alpha)$, i.e. average time between the course changes	
	is $1/\alpha$	49
3.2	The overview of the algorithm	57
3.3	Flowchart of function that calculates the trajectory	58
3.4	Flowchart of function that calculates a single point expansion	59
4.1	Turning maneuver.	71
4.2	The rotation of data	72
5.1	Chromosome representation of a three-segment trajectory	74
6.1	Scenario specifications.	78

6.2	IOU-KF-RTSS: The AEEs and the STDs of produced tracks for a 3-	
	segment trajectory.	80
6.3	IOU-KF-RTSS: Averaged AEEs and STDs for estimated trajectories of	
	$1000\ {\rm ground}\ {\rm truth}\ {\rm tracks}\ {\rm of}\ {\rm different}\ {\rm lengths}\ {\rm and}\ {\rm different}\ {\rm number}\ {\rm of}\ {\rm contacts}$	
	per track	81
6.4	IOU: Averaged AEEs and STDs of 1000 ground truth tracks in the vicinity	
	of turning maneuver.	82
6.5	IOU: AEEs for varying IOU parameters σ and β	83
6.6	IOU-KF-RTSS: AEEs for varying IOU parameters σ and β in terms of	
	TOL and $V.\ldots$	84
6.7	IOU-KF-RTSS: AEEs for tracking ships traveling with speed, $V = 20$ knots,	
	for $\beta = 0.25$ and $\sigma = 6$	85
6.8	IOU-KF-RTSS: STDs for tracking ships traveling with speed, $V = 20$ knots,	
	for $\beta = 0.25$ and $\sigma = 6$	85
6.9	IOU-KF-RTSS vs. GA: The difference of the AEEs and the STDs of	
	produced trajectories for a 3-segment ground truth trajectory	86
6.10	IOU-KF-RTSS vs. GA: Averaged AEEs for 1000 ground truth tracks of	
	different lengths and different number of contacts per track. \ldots	87
6.11	IOU-KF-RTSS vs. GA: Averaged STDs for 1000 ground truth tracks of	
	different lengths and different number of contacts per trajectory	88
6.12	IOU-KF-RTSS vs. GA: Averaged AEEs of a 1000 ground truth tracks	
	in the vicinity of turning maneuver	89
6.13	IOU-KF-RTSS vs. GA: Averaged STDs for a 1000 ground truth tracks	
	in the vicinity of turning maneuver	89

	٠	٠	٠
VII	1	1	1
ΛV	л	л	л

6.14	BRS , (x,y): AEE for a single 3-segment trajectory with 14 contacts versus	
	IOU-KF-RTSS and GA.	91
6.15	BRS, (x,y): STD for a single 3-segment trajectory with 14 contacts versus	
	IOU-KF-RTSS and GA.	92
6.16	BRS, (x,y): Averaged AEEs for a 50 ground truth tracks	92
6.17	BRS , (x,y) : Averaged STDs for a 50 ground truth tracks	93
6.18	BRS from (t, x) and (t, y) : Averaged AEEs and STDs for 50 ground truth	
	tracks obtained by regression with respect to t	94
6.19	BRS, (x,y): AEEs for 7-segment ground-truth trajectory with different	
	number of contacts per trajectory	95
6.20	BRS, (x,y): STDs for 7-segment ground-truth trajectory with different	
	number of contacts per trajectory	95
6.21	AEEs for the $BRS(x,y)$ vs. IOU-KF-RTSS and GA-based for a 7-segment	
	ground truth trajectory for $n = 14$ and $n = 32$	96
6.22	AEEs for the BRS(x,y) algorithm for different initial number of knots, $iknot =$	
	3 and $iknot = 7$, for ten 10-segment ground truth trajectories	97
A.1	The uncertainty ellipse with 96% of probability confinement region	104

List of Tables

6.1	Prior parameters for the Bayesian regression spline algorithm	90
A.1	Table of χ^2 distribution percentage points	109

List of Acronyms

AEE	Average Euclidean Error
AIS	Automated Identification System
DMHT	Distributed Multi-Hypothesis Tracker
EEZ	Exclusive Economic Zone
EKF	Extended Kalman Filter
ELINT	Electronic Intelligence
ESM	Electronic Support Measure
GA	Genetic Algorithm
HFSWR	High Frequency Surface Wave Radar
IOU	Integrated Ornstein-Uhlenbeck
JDL	Joint Directors of Laboratories
KF	Kalman Filter
LS	Least Square
LSDE	Linear Stochastic Differential Equation
MAP	Maximum A Posteriori
MCMC	Markov Chain Monte Carlo
MDA	Maritime Domain Awareness
ML	Maximum Likelihood

MMSE	Maximum Mean Square Error
MSDF	Multisensor Data Fusion
NM	nautical mile
RMP	Recognized Maritime Picture
RMSE	Root Mean Squares Error
RTS	Rauch-Tung-Striebel
RW	Random Walk
SIS	Sequential Importance Sampling
SMC	Sequential Monte Carlo
STD	Standard Deviation
UAV	Unmanned Aerial Vehicle

Chapter 1

Introduction

This chapter discusses a motivating application together with the most important results reported in the literature concerning multisensor data fusion for target location or trajectory estimation (or tracking) in maritime surveillance. The literature is discussed with a special focus on measurement uncertainty issues and multisensor data batch processing methods. Approaches involving multivariate statistical analysis such as piecewise regression algorithms and Bayesian data analysis are of special interest. With this as a background, the thesis research goals are stated and the contributions of this thesis are explained.

1.1 Motivating Application

Maritime surveillance is a set of activities which include search and rescue operations, detection of contraband and environmental hazards, maintaining national sovereignty in the Exclusive Economic Zone (EEZ) (e.g. detection of illegal fishing and protection of ship routes), and defence. Multiple information sources, including military, government and commercial are used to compile a picture of these activities, the so-called Recognized Maritime Picture (RMP). The sensors employed for collection of data include High Frequency Surface Wave Radars (HFSWRs), Automated Identification Systems (AISs), spaceborne and airborne sensors usually mounted on aircrafts or unmanned aerial vehicles (UAVs), electronic support measures (ESM), electronic intelligence (ELINT), etc. Figure 1.1 depicts a snapshot of a version of the RMP with contacts represented by dots which illustrate multiple ships locations and their identifications (IDs) in the surveillance area. The stand-



Fig. 1.1 A snapshot of the RMP [6].

alone dots with no ID (the so-called "ambiguities") represent contact reports which are not fused properly by a tracking algorithm.

Increasing sophistication of maritime surveillance systems has generated a great deal of interest in development of multisensor data fusion (MSDF) systems in order to produce more accurate, better integrated and automated RMP. Such improvements are necessary for reducing operator overload during periods of peak activity and for providing improved and constant resolution of ambiguous sensor inputs [7]. An important component of such MSDF systems are tracking (or trajectory estimation) algorithms. Within the most widely adopted architecture model for multisensor data fusion, the Joint Director Laboratories (JDL) as reported in Hall and Llinas [8], and its subsequent revisions [9], [10], the tracking algorithms make part of the so-called Level 1 fusion. The general concept of Level 1 fusion for the case of ship tracking with multiple sensors is shown in Fig. 1.2.



Fig. 1.2 Concept of Level 1 data fusion for a surveillance system.

The important stage prior to fusion of any data, especially in dissimilar multisensor case, is the data alignment. The data alignment includes temporal, spatial and measurement unit adjustments necessary for subsequent processing. Furthermore, the time when an observation is received at a fusion system must be distinguished from the time when the observation was actually made.

The goal of a tracking (or trajectory estimation) algorithm is to produce the best estimate of the location of the object of interest towards forming a trajectory of motion and possibly also making inferences about the object attributes (e.g. operating frequency, ID, etc.) from surveillance data.

Fig. 1.3 shows examples of estimated trajectories obtained by two different tracking algorithms. The measurements (or contacts) are assumed to be obtained from five different sensors at the time instants t_1, \ldots, t_5 , respectively, where the ellipses signify the regions of uncertainty for any given measurement. Different sizes and orientations of the measurement error ellipses indicate the dissimilar sensor environment.



Fig. 1.3 Possible estimated trajectories from dissimilar multisensor surveillance data.

Several issues must be considered when designing a trajectory estimation algorithm for a multisensor data fusion system for maritime surveillance. These can be identified as:

• Large difference in temporal resolution of data. For example, a HFSWR provides near real-time tracking with position updates every three minutes on average, while a surveillance aircraft flying along a predefined path may update a ship's position only once per day [2].

• Latency. Many contact reports arrive hours after the observation. The delay is caused by multi-level security issues and information management issues [6].

The two issues mentioned above make the contact/sensor report data mostly available in batches.

NOTE: This should not be confused with the availability of measurements in terms of a sensor performance, i.e. the sensor models under which a sensor can produce useful data and which involve signal and thermal noise powers.

- Possibly compromised information. Even a reliable and relatively certain sensor information may be compromised. For example, the Global Positioning Systems (GPS) and the AIS transponders required to be mounted on board of a ship can be switched off or duped (e.g. illegal fishing on the borders of the Exclusive Economic Zone (EEZ)). Sending out patrolling ships or a surveilling aircraft far as ~ 200 nautical miles (NM) offshore may be costly or fuel and time limiting. In such situations one has to rely on the knowledge of the motion model and the HFSWR. However, this knowledge can be limited or insufficient, especially if the motion models are considered as deterministic like in commonly used tracking algorithms. Also, the weather conditions may impair the HFSWR observations and HSWFR tracks. An alternative is to use passive sensors such as the ELINT or the ESM.
- Measurements with relatively large heteroscedastic and correlated errors. Each sensor is characterized by the measure of location of accuracy as the size of the minimum area having a specified probability of containing the measurement, i.e.

the area of uncertainty (AOU). The AOUs can be of three types: circular or ellipse, line of bearing, and bearing box. As indicated in Fig. 1.3, in maritime surveillance the most common AOU is the ellipse. The measurement vector, the so-called *ellipse contact report*, comprises of the ellipse centre coordinates (i.e. a measured location), magnitudes (or "lengths") of the semi-axes, the ellipse orientation (or heading), the associated time tag and the specified probability that the true location is within the ellipse region. Given the specified probability, while assuming that the measurement errors are normally distributed, Gaussian statistics can be used to describe the positional measurement error by transforming the error ellipses to their related location covariance matrices. With reference to the practical problem of position estimation in a dissimilar multi-sensor environment, the covariance matrices of measurement errors must generally be considered as non-diagonal and non-constant, rendering the errors heteroscedastic and correlated [11], see Fig.1.4.



Fig. 1.4 (a) Heteroscedastic and (b) heteroscedastic and correlated measurement errors.

NOTE: The term *heteroscedastic* is mostly used in statistical inference for regression techniques for fitting data with errors in variables. It designates nonconstant error covariance matrix across measurements.

Additionally, the measurement errors ellipse may be relatively large in size. For example, the magnitude of the semi-major axis of the measurement error ellipse may be measured up to 5 nautical miles [NM].

In order to contribute to increased awareness and improved decision making in the maritime domain, the aforementioned issues should be addressed when designing a trajectory estimation algorithm within a multisensor data fusion system for maritime surveillance. Special attention must be paid to the availability of data for processing, effective characterization of the uncertainty of surveillance data, its incorporation into the estimation algorithm, and the choice of a suitable technique for target trajectory estimation. This necessitates the study of both sequential and batch trajectory estimation algorithms.

The following Section presents general background information on trajectory estimation within a multisensor data fusion surveillance system. The focus is on standard mathematical tools for trajectory estimation which employ the probability theory to quantify the uncertainty and rely on the Bayesian inferencing. Within this context the problem of interest is also identified. The literature survey will be presented separately in the subsequent Section.

1.2 Trajectory Estimation Algorithms for Multisensor Data Fusion-The Background

Tracking refers to processing of measurements obtained from an object of interest in order to maintain an estimate of its current *state*, which typically consists of kinematic components

(e.g. position and velocity) and possibly ship attributes (e.g. radar operating frequency, ship ID) [12].

Measurements or *contacts* are noise corrupted observations related to the state of the object of interest, such as position, velocity, range, etc.

Estimates of the kinematic characteristics are available in the form of an estimated *track* (or *trajectory*).

A *track* is a state trajectory estimated from a set of measurements that have been associated with the same target.

Trajectory estimation algorithms from multi-sensor measurements can be classified in terms of the following characteristics:

- Single versus multiple target tracking. There exists a fundamental distinction between single-target tracking and multiple-target tracking as the latter requires a data association procedure in order to determine the origin of measurements. The data association procedure is also necessary to take into account the presence of clutter, countermeasures, or false alarms. The descriptions of various data association procedures can be found in [13].
- Sensor level versus central level fusion or hybrid fusion. These fusion approaches are determined by the extent of the data processing, data product types and fusion level. In sensor level fusion, each sensor detects and estimates target trajectories. They are passed to the *fusion processor* which combines the information from the sensors to improve the state estimate of the target. This is optimal if the sensors use independent signature-generation phenomena to derive information about the target, [14]. The information sent to the fusion processor includes detection or classification decision, target location, and the information on how well the sensor has

been able to detect the target, i.e. the area of uncertainty. In the central level fusion, sensors provide data subject to filtering and baseline estimation at the sensor level. The sensor tracks are then passed to the fusion processor. The data are processed at higher rates than in sensor level fusion. An alternative to sensor level and central level fusion is hybrid fusion which combines these two approaches [14].

- Continuous versus discrete. Measurements may be acquired continuously in time or, else at discrete time instants. The length of the time interval between discrete time instants can be either constant or varying.
- Batch versus sequential. The tracking implementation approach may either be based on batch or else sequential processing. In the batch processing approach all gathered observations over a given time horizon are analyzed and processed at once, while in a sequential approach the measurements are processed as they are received.
- Random versus non-random model. There exist two models for estimating an unknown trajectory parameter θ .

Let $\mathbf{Y} = \{y_1, y_2, \dots, y_k\}$ be a set of k observations.

- Non-random. Non-random model assumes an unknown true value of the $\theta = (\theta_1, \theta_2, \dots, \theta_k), k \ge 1$ which is non-random. The probability density which assigns the probability of observing **Y**, when the parameter value is θ , is referred to as a *likelihood function* for the measurement data conditioned on the parameter θ . It is defined by, see [15]

$$\Lambda(\theta) = p(\mathbf{Y}|\theta) \tag{1.1}$$

The likelihood function represents a measure of the evidence from the data.

In case of non-random parameter estimation the common criteria for the best fit of the estimated parameter to data are:

1. The maximum likelihood (ML) method. The ML estimate of a nonrandom parameter θ is assumed to maximize the likelihood function (1.1), i.e.

$$\hat{\theta}^{ML}(\mathbf{Y}) = \operatorname{argmax}_{\theta} \Lambda(\theta) = \operatorname{argmax}_{\theta} p(\mathbf{Y}|\theta)$$
(1.2)

While θ represents an unknown constant, the ML estimate $\hat{\theta}^{ML}(\mathbf{Y})$ represents a random variable since it is a function of the set of random observations Y.

2. The least square (LS) method. The LS estimate of a nonrandom parameter represents the value of the parameter that minimizes the square error between the measurements and the observed function of the parameter. Given the scalar linear or nonlinear measurement model

$$y_j = h_{j,\theta} + v_j \qquad \qquad j = 1, \dots, k \qquad (1.3)$$

where $h_{j,\theta}$ is known linear or nonlinear function, v_j is a random measurement error, the LS estimator of θ is defined as

$$\hat{\theta}_k^{LS} = \arg\min_{\theta} \{ \sum_{j=1}^k (y_j - h_{j,\theta})^2 \}$$
(1.4)

The non-random model approach does not incorporate any set of prior beliefs. Nevertheless, the usual claims of "objectivity" are illusory since the method still requires numerous assumptions about the model, the measurement error, etc. [16]

- **Random.** The unknown parameter $\theta = (\theta_1, \theta_2, \dots, \theta_k), k \ge 1$ is considered to be a random variable with some *a priori* belief that can be expressed in terms of the prior probability density function (p.d.f.) $p(\theta)$. For a fixed number observations $\mathbf{Y} = \{y_1, y_2, \dots, y_n\}$ the statements about θ are made considering all possible samples of θ obtained from repeated sampling from the conditional posterior density $p(\theta|\mathbf{Y})$ as derived from the Bayes' Theorem, [15],

$$p(\theta|\mathbf{Y}) = \frac{p(\mathbf{Y}|\theta)p(\theta)}{p(\mathbf{Y})} \propto p(\mathbf{Y}|\theta)p(\theta)$$
(1.5)

where $p(\mathbf{Y}|\theta)$ is a known measurement likelihood function. The approach that employs this model is therefore called *Bayesian*. If data plays a role in determining the prior distribution, $p(\theta)$, the approach is referred to as empirical Bayesian.

Estimation of a random parameter θ with a prior p.d.f. $p(\theta)$ can be carried out using:

1. The maximum a posteriori (MAP) method. This method employs the maximization of the posterior p.d.f., i.e.,

$$\hat{\theta}^{MAP}(\mathbf{Y}) = \operatorname{argmax}_{\theta} p(\theta | \mathbf{Y}) = \operatorname{argmax}_{\theta} p(\mathbf{Y} | \theta) p(\theta)$$
(1.6)

The MAP estimate is hence also a random variable.

2. The Minimum Mean Square Estimator (MMSE). It is the counterpart of

the LS estimator for the nonrandom parameter. It is defined by, [15],

$$\hat{\theta}^{MMSE}(\mathbf{Y}) = \operatorname{argmin}_{\theta} E\{(\hat{\theta} - \theta)^2 | \mathbf{Y})\}$$
(1.7)

The solution to (1.7) is the conditional mean of θ

$$\hat{\theta}^{MMSE}(\mathbf{Y}) = E\{\theta|\mathbf{Y}\} = \int_{-\infty}^{+\infty} \theta p(\theta|\mathbf{Y}) d\theta \qquad (1.8)$$

The mathematical tools for trajectory estimation rely on techniques of system theory which use (kinematic) state-space representation of linear/nonlinear deterministic/stochastic dynamic systems and optimal linear and nonlinear filtering. In the case of nonlinear dynamics and non-Gaussian noise models filter approximations or suboptimal solutions are applied. Besides system theory, the statistical analysis tools can be used to build a statistical trajectory model which describes all known and unknown quantities of interest (e.g. parameters of a polynomial that fits the data). In this case, either frequentist or Bayesian approach may be used to draw inferences about the unknown parameters.

In this thesis the research is limited to a single ship trajectory estimation from discrete random heteroscedastic measurements, where the ship's position is considered to be a random variable. Restriction to random models implies the need to employ Bayesian approaches to trajectory estimation such as Bayesian filtering, as well as Bayesian approach to inferencing about probabilistic multivariate regression models for piecewise polynomial fitting. Therefore, a background on these Bayesian approaches is provided next. It includes a review of Bayesian approach to filtering for the purpose of trajectory estimation from multisensor measurements, and the principles of Bayesian statistical modeling and analysis.

1.2.1 Bayesian Approach to Linear/Nonlinear Filtering

In the Bayesian approach to dynamic state estimation one attempts to construct the posterior probability density function (p.d.f.) of the state, based on all available information, including the sequence of received measurements [17]. If either the system or a measurement model is nonlinear, the posterior p.d.f. is non-Gaussian. In principal, an optimal (with respect to a criterion) estimate of the state may be obtained from the posterior p.d.f. For common tracking algorithms an estimate is required every time a measurement is received. In this case, a recursive filter is a standard solution where the received data are processed sequentially rather than in a batch. Such a filter consists of two steps: prediction (where the system model is used to predict the state p.d.f. forward from one measurement time to the next) and the information update (where the latest measurement is used to modify the predicted p.d.f.). The update step is accomplished using Bayes' theorem as a mechanism for updating the knowledge about the state.

Let the ship state vector be denoted by $x_k \in \mathbb{R}^n$, where *n* is the dimension of the state vector and \mathbb{R} is a set of real numbers, $k \in \mathbb{N}$ is time index and \mathbb{N} is the set of natural numbers. The index *k* is assigned to a continuous time instant t_k , and the sampling interval $t_k - t_{k-1}$ may be time dependent. The target evolves according to the following discrete-time stochastic model:

$$\mathbf{x}_{\mathbf{k}} = f_{k-1}(\mathbf{x}_{\mathbf{k}-1}, \mathbf{w}_{\mathbf{k}-1}) \tag{1.9}$$

where f_k is a known linear or nonlinear function of the state \mathbf{x}_{k-1} and \mathbf{w}_{k-1} is a process noise sequence. The process noise models the error in the motion model and exogenous disturbances. The objective of filtering is to recursively estimate \mathbf{x}_k from the measurements $\mathbf{y}_k \in \mathbb{R}^m$. The measurements are assumed to be related to the target state by the measurement equation

$$\mathbf{y}_{\mathbf{k}} = h_k(\mathbf{x}_{\mathbf{k}}, \mathbf{v}_{\mathbf{k}}) \tag{1.10}$$

where h_k is a known linear or nonlinear function and \mathbf{v}_k is a measurement noise sequence. The noise sequences \mathbf{w}_k and \mathbf{v}_k are assumed white with known probability density functions and mutually independent. The initial state is assumed to have a known p.d.f. $p(\mathbf{x}_0)$, which is independent of noise sequences.

The Bayesian filtering algorithm seeks filtered estimate of \mathbf{x}_k based on the sequence of all available measurements $\mathbf{Y}_k \triangleq \{\mathbf{y}_i, i = 1, \dots, k\}$ up to time k from the posterior distribution $p(\mathbf{x}_k | \mathbf{Y}_k)$, given the initial p.d.f. of the state vector, $p(\mathbf{x}_0) \triangleq p(\mathbf{x}_0 | \mathbf{Y}_k)$. The value of $p(\mathbf{x}_k | \mathbf{Y}_k)$ can be obtained in two steps, the prediction and the information update as following.

Suppose that the required p.d.f. $p(\mathbf{x_{k-1}}|\mathbf{Y}_{k-1})$ at time k-1 is available. Then, the prediction stage uses the system model (1.9) to obtain the prediction (or prior) density of the $\mathbf{x_k}$ using Chapman-Kolmogorov equation, [17]:

$$p(\mathbf{x}_{k}|\mathbf{Y}_{k}) = \int p(\mathbf{x}_{k}|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1}) d\mathbf{x}_{k-1}$$
(1.11)

where $p(\mathbf{x_k}|\mathbf{x_{k-1}}, \mathbf{Y}_{k-1}) = p(\mathbf{x_k}|\mathbf{x_{k-1}})$ as (1.9) describes a Markov process of order one.

At time k, when the measurement $\mathbf{y}_{\mathbf{k}}$ is available, the information update step is per-
formed as follows:

$$p(\mathbf{x}_{\mathbf{k}}|\mathbf{Y}_{k}) = p(\mathbf{x}_{\mathbf{k}}|\mathbf{y}_{\mathbf{k}}, \mathbf{Y}_{k-1})$$

$$= \frac{p(\mathbf{y}_{\mathbf{k}}|\mathbf{x}_{\mathbf{k}}, \mathbf{Y}_{k-1})p(\mathbf{x}_{\mathbf{k}}|\mathbf{Y}_{k-1})}{p(\mathbf{y}_{\mathbf{k}}|\mathbf{Y}_{k-1})}$$

$$= \frac{p(\mathbf{y}_{\mathbf{k}}|\mathbf{x}_{\mathbf{k}})p(\mathbf{x}_{\mathbf{k}}|\mathbf{Y}_{k-1})}{p(\mathbf{y}_{\mathbf{k}}|\mathbf{Y}_{k-1})}$$
(1.12)

where the normalizing constant

$$p(\mathbf{y}_{k}|\mathbf{Y}_{k-1}) = \int p(\mathbf{y}_{k}|\mathbf{x}_{k}) p(\mathbf{x}_{k}|\mathbf{Y}_{k-1}) \mathbf{d}\mathbf{x}_{k-1}$$
(1.13)

depends on the likelihood function value $p(\mathbf{y_k}|\mathbf{x_k})$ defined by the measurement model (1.10) and the known statistics of the measurement noise sequence $\mathbf{v_k}$. In the information update step, the measurement $\mathbf{y_k}$ is used to modify the prior (predicted) density to obtain the posterior density. The relations (1.11) and (1.12) represent the basis for the optimal Bayesian solution which solves the problem of exact and complete characterization of the posterior density in a recursive manner. From (1.12), the optimal state estimate can be obtained according to any of the criteria for a random model, MAP (1.6) or MMSE (1.7).

The solution given by (1.11) and (1.12) is only conceptual since its implementation would require the storage of the entire (possibly non-Gaussian) p.d.f., which would involve an infinite dimensional vector. Only in the linear Gaussian case the recursion (1.11) and (1.12) becomes the well known Kalman filter [18]. In other situations, approximations such as the Extended Kalman filter (EKF), a Gaussian sum filter [19], the use of numerical methods [20], [21], or suboptimal Bayesian algorithms based on sequential Monte Carlo approximations, are necessary.

Grid-based methods provide an optimal recursion of the filtered density, $p(\mathbf{x}_k|\mathbf{Y}_k)$, if

the state space is discrete and consists of a finite number of states [17]. Also, Benes and Daum filters provide solutions for a class of nonlinear systems and a linear measurement equation for which the posterior p.d.f. admits "sufficient statistics" of a constant finite dimension [22].

Particle filters are suboptimal filters which perform sequential Monte Carlo (SMC) estimation based on point (or "particle") representation of probability densities. The basic SMC idea was introduced in 1950' in [23] and later extended in [24]. However, only recently, due to increased computational power, and to the introduction of the re-sampling step [25], the particle filters became a popular filtering approach. An extensive description of the use of particle filters for tracking applications can be found in [17].

1.2.2 Bayesian Statistical Analysis for Trajectory Estimation

When the knowledge about the motion model does not exist or is insufficient, the standard tools for dynamic state estimation, such as linear or nonlinear filtering for producing a trajectory of motions cannot be used. Also, the advantage of commonly used sequential procedures for real time tracking may not be fully exploited if the data are only available in batches. In this case, techniques of (trajectory) model fitting to data, such as regression modeling and stochastic optimization procedures for data fitting (e.g. genetic algorithms) may be viable alternatives for trajectory estimation.

The Regression Problem

The regression problem involves determining the relationship between some response variable Y and a set of p predictor variables $X = (X_1, \ldots, X_p)$. The responses are assumed to be related to predictors through some deterministic function f and an additive random error ϵ , i.e.

$$Y = f(X) + \epsilon \tag{1.14}$$

where $\epsilon \sim N(0, \Sigma)$ is independent of X, and $f(\cdot)$ is a smooth function [26]. The function $f(\cdot)$ is usually referred to as the *regression function*. The function is unknown and cannot be determined exactly but needs to be estimated over predictor values $x \in \mathcal{X}$. In this setting, the "regression analysis or modeling" refers to methods for statistical inference about the regression function, $f(\cdot)$. The regression modeling attempts to find the approximation using the observed data set \mathbf{Y} at some known predictor locations so that $\mathbf{Y} = \{y_i, x_i\}_{i=1}^n$.

The simplest choice for the approximation of f is a linear function. A more general assumption can be made on the approximation of f such as a linear combination of basis functions and corresponding coefficients, i.e.,

$$f(x) \sim \sum_{i=1}^{k} \beta_i B_i(x), \qquad x \in \mathcal{X}$$
(1.15)

where $\beta = (\beta_1, \dots, \beta_k)'$ is the set of coefficients corresponding to basis functions $\mathbf{B} = (B_1, \dots, B_k)$. The basis functions in (1.15) can be natural cubic splines, B-splines, truncated polynomials, etc.

Possible assumptions on $f(\cdot)$ determine a suitable regression modeling methodology for (1.14): *parametric* or *nonparametric* regression modeling. A parametric regression model assumes that the form of $f(\cdot)$ is known except for the finite number of unknown parameters. More specifically, it is assumed that there exists a vector of parameters $\beta = (\beta_1, \ldots, \beta_p)' \in \mathcal{B}$, where $\mathcal{B} \subset \mathbb{R}^p$ and a function $f(\cdot)$ such that $f(\cdot) = f(\cdot; \beta)$. Hence, in parametric regression model, the inference about f amounts to inference about β . Parametric models can depend on parameters in linear or nonlinear fashion. For linear models, the classical method of estimating the regression coefficients in model (1.14) is the LS estimator. Hypothesis testing can be used for further inference about f and β . For nonlinear models, the most common parameter estimation method is the Gauss-Newton method [27]. In the case of both linear and nonlinear model, the resulting estimate is a curve that is selected from the family of curves allowed under the model and which conforms to the data in some fashion.

Nonparametric regression methodologies rely rather on data than on possible information about f. To construct a nonparametric regression model, an appropriate function space is chosen for which f is believed to belong to. Then, the data is used to determine an element of this function space that represents the unknown regression curve. The methods to perform nonparametric regression include deconvolution kernel functions methods, smoothing splines using wavelets, polynomials or other types of basis functions, and Bayesian inference [26].

Bayesian Approach to Regression Modeling

The Bayesian approach to regression modeling makes use of the Bayes' Theorem to update beliefs about the aspects of the model given the set of observed data. Bayesian analysis treats all parameters as random, assigns prior distributions to characterize knowledge about parameters prior to collection of data, and uses the joint posterior distribution of parameters to make inferencing about the unknown parameters given the data [28].

Suppose a likelihood $p(\mathbf{Y}|\theta)$ for the observed data \mathbf{Y} given the vector of unknown parameters θ , is known. The vector of unknown parameters θ may contain the set of coefficients corresponding to basis functions $\mathbf{B} = (B_1, \ldots, B_k)$. It may also include other parameters such as number and locations of change-points (e.g. for a spline or a piecewise truncated polynomial basis). Also, let a prior for θ be given in terms of probability density function $\pi(\theta|\eta)$, where η is an assumed known vector of hyperparameters, i.e. $\pi(\theta) = \pi(\theta|\eta)$. Inference about θ is based on its posterior distribution

$$p(\theta|\mathbf{Y}) = \frac{p(\mathbf{Y},\theta)}{p(\mathbf{Y})} = \frac{p(\mathbf{Y},\theta)}{\int p(\mathbf{Y},\theta) d\theta}$$
(1.16)

$$=\frac{p(\mathbf{Y}|\theta)\pi(\theta)}{\int p(\mathbf{Y}|\theta)\pi(\theta)d\theta}$$
(1.17)

where $\int p(\mathbf{Y}|\theta)\pi(\theta)d\theta$ denotes the marginal distribution of Y.

If η is unknown, the fully Bayesian approach adopts a hyperprior distribution $h(\eta)$ and computes the posterior distribution as

$$p(\theta|\mathbf{Y}) = \frac{\int p(\mathbf{Y}|\theta)g(\theta|\eta)h(\eta)d\eta}{\int \int p(\mathbf{Y}|u)g(u|\eta)h(\eta)dud\eta} = \int p(\theta|\mathbf{Y},\eta)h(\eta|\mathbf{Y})d\eta$$
(1.18)

The posterior distribution, $p(\theta|\mathbf{Y})$, often cannot be computed analytically, but can be only approximated or simulated.

Bayesian Computation

Determination of posterior distributions involves evaluation of complex and often highdimensional integrals. Computation of moments of the posterior distribution also leads to more integration. The use of conjugate priors may enable partial analytic evaluation of these integrals. Asymptotic methods, such as normal approximation or Laplace's method [29] for possibly asymmetric posterior approximation, can also be used. When approximate methods are intractable or are not sufficiently accurate, numerical integration methods such as Monte Carlo must be used. Particularly powerful are the iterative (or sequential) Monte Carlo methods (SMC), which produce a Markov chain, the output of which corresponds to a (correlated) sample from the joint posterior distribution [16], i.e. the Markov Chain Monte Carlo methods.

The MCMC methods are used when it is not possible to sample θ directly from $p(\theta|\mathbf{Y})$. Instead, the samples are drawn from an approximate distribution that evolves closer and closer to the posterior distribution $p(\theta|\mathbf{Y})$ by correcting the draws. The samples are drawn sequentially depending on the last value drawn, i.e. they are processed by employing a *transition distribution* $T_t(\theta^k|\theta^{k-1})$ of the Markov chain. The transition distribution $T_t(\theta^k|\theta^{k-1})$ must be constructed so that the Markov chain converges to a unique stationary distribution, i.e. the posterior distribution $p(\theta|\mathbf{Y})$.

Various methods are devised for constructing and sampling from transition distributions for arbitrary posterior distributions defined on spaces of known and fixed dimensions among which Metropolis [30] and Metropolis-Hastings [30], [31] and Gibbs sampler [32] are the most commonly used sampling schemes. For sampling of probability distributions defined on spaces of variable dimensions, the schemes such as reversible jump Markov Chain Monte Carlo [33] must be used.

1.3 Problem and Research Goals

This thesis delivers and discusses algorithms for trajectory estimation when dynamic or kinematic model of the target motion is unknown. The algorithms are restricted to a single target case. Only sensors that provide positional measurements with associated elliptical areas of uncertainty (e.g. ESM or ELINT) are considered. Therefore, the data are assumed to have heteroscedastic and correlated measurement errors (i.e. the measurement error ellipses of different sizes and orientations), obtained at irregular time intervals, and available in batches. This allows the problem of estimating a trajectory to be formulated in two ways:

- 1. As a linear stochastic optimal filtering problem which will rely on a combination of two stochastic processes to describe the ship's motion model, namely the Integrated Ornstein-Uhlenbeck process (IOU) and the random walk (RW) process. The model will be constructed on the basis of common knowledge on ship navigation. The assumptions on linear modeling and Gaussian distribution of measurement errors will then allow for the use of Kalman filtering and Rauch-Tung-Striebel optimal smoothing. In the open literature, the first complete description and functionality assessment of this algorithm, with a special focus on the importance of the measurement error modeling and appropriate incorporation into the tracking algorithm was provided by Hadzagic and Michalska (2010) [3]. The Rauch-Tung-Striebel optimal smoothing was added to improve the position estimates.
- 2. As a curve (i.e. trajectory) fitting problem to the measurements with heteroscedastic and correlated errors in which the Bayesian approach to the nonparametric regression modeling is used to make inference about the trajectory model and its characteristics. Specifically, the trajectory will be considered to be identified in the form of a piecewise spline function in the 2-D Euclidian plane with an unknown number of change points (or knots) which correspond to the locations of a ship's maneuvers.

In light of the above, the research goals were stated as:

• To develop trajectory estimation algorithms that successfully address practical issues which transgress the usual assumptions found in the target tracking literature such as high and regular data rates, relatively small and circular areas of measurement uncertainty, known motion models and the availability of data for processing. Precisely, the new algorithms should not assume the knowledge of the motion model, and should assume data with heteroscedastic and correlated errors, as well as the measurements arriving in batches.

- To perform sensitivity analysis of both algorithms with respect to perturbations in the model parameters.
- To assess the efficiency of both algorithms for various ground truth scenarios. For this purpose, a ground truth scenario that corresponds to real ship motion in 200 NM x 200 NM surveillance area is to be developed.
- To compare both algorithms with another batch stochastic optimization algorithm, namely a genetic algorithm (GA), used to form a trajectory from data with heteroscedastic and correlated errors [3].
- To analyse the results of the comparisons, provide conclusions about the novel algorithms and outline possible future research avenues.

To put the contribution of this thesis in a broader context, previous results in trajectory estimation from multisensor data, as well as piecewise function regression methods from data with and without errors, are presented and their advantages or shortcomings with respect to maritime surveillance issues are summarized below.

1.4 Literature Survey

Karl Friedrich Gauss (1777-1855) was the first to study the problem of trajectory estimation from uncertain observations. His results in estimation of trajectories of celestial bodies using the method of least squares [34] were published in 1809, later than those of Adrien-Marie Legendre (1752-1833) who independently published his work on the least squares method in 1806 [35]. Subsequent historical developments of estimation techniques include Fisher's interpretation of the least square method [36] and definition of the maximum likelihood method, Wiener [37] and Kolmogorov [38] development of the minimum mean square error method, and Kalman's formulation of a discrete time, recursive, minimum mean square filtering, the Kalman filter [39]. The history of estimation techniques is summarized in [19]. Extensive literature concerning recent target trajectory estimation algorithms, particularly the sequential ones, is presented in Blackman [13] and Bar-Shalom [40]. General assumptions usually include the knowledge of a linear or nonlinear model and positional radar measurements. The latter are usually assumed to be received at high and regular data rates and to provide good localization information about the target [13]. In tracking applications, linear or nonlinear recursive filtering, are common solutions to trajectory estimation. For a linear motion model, the Kalman filter in conjunction with a Gaussian noise model gives the optimal estimate. This case can be extended to a multiple model approach using the Interacting Multiple Model (IMM) filter to address trajectory estimation for a maneuvering target [15], [13]. However, this complexity and computational cost of this approach that employs banks of Kalman filters makes it inefficient for fast maneuvering ships found in counter-drug operation scenarios [41]. The assessment of the IMM algorithm in the context of tracking fast maneuvering aircrafts can be found in [42], [43] and [44]. For nonlinear and non-Gaussian models in state space forms, there exist analytic approximations such as Extended Kalman Filtering (EKF) [15], and variants such as linearization and Gaussian sum filter [40], [45]. Numerical approximations include grid based methods, Gaussian mixtures [46] and unscented filter [47]. Monte Carlo approximations include sequential importance sampling (SIS) [24], Rao-Blackwellisation [48] and sequential Monte Carlo (SMC) methods which estimate the complete probability distribution. The latter are known as particle filters, Sequential Importance Resampling (SIR) filters, bootstrap filters,

etc. These methods are more appropriate for tracking applications which consider non-Gaussian measurement errors, high data rates, and nonlinear models such as bearing-only tracking. Their descriptions and the use in tracking applications can be found in [49] [17].

The state space Bayesian filtering approach is proved to be superior for radar and infrared data in low signal-to-noise ratio environments over the contact-based Kalman filtering using deterministic motion models because the likelihood functions allow to incorporate the detailed physics of the sensor response to the target in a noisy environment as well as other available prior information, such as restrictions of moves (e.g. land-avoidance likelihood function) [21]. State space Bayesian filtering proved also to be more robust for data with high uncertainty [50]. This applies to particle filters as well since they are based on sequential Monte Carlo methods and therefore sequential by nature.

In [50], it is said that a state space Bayesian filtering algorithm which employs the Integrated Ornstein-Uhlenbeck (IOU) process as the motion model is successfully used in applications such as tracking submarines and surface ships. However, no detailed algorithm description nor implementation was provided. In [51], the Rauch-Tung-Striebel (RTS) fixed-interval optimal smoothing was said to be used with the IOU to improve the estimates of position and velocity. The first complete description, implementation, application, and the accuracy and functionality assessment of the IOU process used for ship trajectory estimation is reported in [52], [3]. It is shown that the IOU process can be used for ship motion modeling only in specific combination with the random walk (RW) process. The IOU model is assumed unknown and is constructed using basic knowledge on ship navigation and RW process parameters. Furthermore, in [3] the importance of adequately modeling and completely accounting for measurement error uncertainty in ship trajectory estimation was demonstrated by comparing the IOU-based trajectory estimation algorithm with a genetic algorithm based trajectory estimation. Other alternative approaches such as track-before-detect [53], [54], evidence theory [55], fuzzy set theory [56], possibility theory [57] can be also used for situations where standard approaches fail (e.g. low data rate, low signal-to-noise ratio, ambiguous information about the target, negative information).

Results in trajectory estimation pertaining specifically to maritime surveillance applications include those reported in [58], [59] where the data obtained from HFSWRs are used to form a trajectory. However, a high-rate single type multi-sensor environment is only considered. Since the HFSWR produces measurements with circular errors, capturing the uncertainty was not an issue. A dissimilar multisensor environment is considered in [60] where the Distributed Multi-Hypothesis Tracker (DMHT) is employed to process data obtained from an AIS sensor network and a coastal radar. The focus is on track fusion and the global maritime surveillance system design rather than on (single) trajectory estimation. Another dissimilar multisensor environment pertaining to maritime surveillance is addressed in [2]. The described algorithm for ship trajectory estimation is based jointly on a "hybrid" genetic algorithm, i.e. a combination of a genetic algorithm, a simulated annealing and a chemotaxis algorithm, and the track templates mechanism that employs non-sensor information. Although it addresses the availability of measurements for processing by using a batch stochastic optimization procedure for data fitting, the algorithm does not completely account for measurement uncertainty because it considers only the approximation of the measurement error obtained by projection.

For measurements with heteroscedastic and correlated errors and available in a batch, a robust line fitting solution using Hough Transform is presented in [11] and is used to detect lines in pictures. However, the algorithm is limited to image processing applications since it involves the intensity of image pixel in the calculations.

As batch procedures, the regression methods have also found their applications in tra-

jectory estimation. The problem of trajectory estimation of a maneuvering target can be translated to a change point analysis for regression in a multi-phase random linear model with known/unknown number of change points, and continuous/discontinuous at change points with an arbitrary error. The continuity or discontinuity of the regression function at the change point influences the statistical inference about the parametric or nonparametric model. Estimation procedures involve both frequentist and Bayesian inferencing.

The results in estimation of multivariate piecewise function regression can be classified in terms of known number (e.g. one or multiple) or unknown number of change points (or knots), and random or non-random approach to fitting.

Single change point. Results that consider both known and unknown single change point include those in [61], [62], [63] [64]. Two-phase nonlinear regression with smooth transition using two-stages LS fit is reported in [65], while in [66] a two phase step function is considered. An ML estimator is used for both random and nonrandom regression in [67] and [68], while for random parametric regression, in [68], a two-phase linear model was addressed. Another procedure for a single change point determination for a nonlinear model was presented in [69]. Asymptotics of M-estimators in two-phase linear regression model was examined in [70]. Rukhin and Vajda(1997) [71] consider the change point estimation problem as a nonlinear regression problem and prove that M-estimators can localize the change point and establish the consistency of a class of approximate M-estimators. In all the results above only a single change point is considered. However it is less likely that a ship in the surveillance area performs only one maneuver, at either known or unknown location.

Multiple change points. The problem of multi-phase linear and a non-linear model with known number of change points was addressed in [72]. In [73], a multivariate nonparametric regression spline was presented, also for a known number of change points. The problem of determination of multiple changes in a piecewise linear model for unknown number of change points is addressed first in [74] and does not include the measurement error. Later in [75] the consistency of estimators of the change points is proven, and tests for multiple changes and constructed confidence intervals for the break points are provided. Additionally, in the same paper efficient algorithms for computing the estimates are proposed, including several methods (one of which is purely sequential) for determining the number of breaks. A method for determining the number of change points in a type of multi-phase piecewise linear regression model, the structural break model, by sequential testing is presented in [76]. Estimation of the number of change points in a multi-phase generalized linear model, with known model parameters was considered in [77]. A very significant work related to variable number of locations of break-points (knots) is presented by Green [33]. Denison et al. [78] generalized this approach to higher order free-knot splines. However, by not specifying a prior distribution on spline parameters, and using LS estimation at each stage instead, the fully Bayesian approach is omitted. Zhou & Shen [79] apply spatially adaptive regression spline as frequentist, iterative method. Di Matteo et al. [1] use Bayesian adaptive regression free-knot spline fitting to scalar measurements with the nonparametric model error assumed as constant. None of the above mentioned procedures for either single or multiple change point considers errors in data, and in particular the heteroscedastic and correlated errors. Also, none of the above-mentioned regression methods addresses the case of decreasing value of predictor variable which in ship trajectory estimation occurs in the case of a turning maneuver. Heteroscedastic errors (or errors-in-variables problem) has been previously addressed by [80] using deconvolution method for nonparametric regression and in [81] for nonlinear regression using Fourier transform. Approaches in nonparametric regression modeling with correlated errors using kernel methods, spline smoothing and wavelet regression can be found in [82].

Regression in tracking. Results concerning regression in trajectory estimation applications are reported by El-Hawary and Jing (1995) [83]. The authors proposed a robust regression-based EKF for tracking underwater targets using high uncertainty measurements, while assuming a known motion model. Another work by Fruhwirth et al. [84] provides results on reconstructing trajectories (the so called secondary vertices) from real position data points which represent points of particle collision using M-estimation. However, the measurement uncertainty issue was not addressed. Splines have been previously used in [85] for ML estimation of paths in tracking using bearings-only data with errors that follow Von Mises distribution. In [86] deterministic annealing procedure is incorporated into a robust M-estimator and applied to determine the target motion parameters based on non-Gaussian and non-stationary measurements derived from passively observed transient signals radiated by a target. In [87], a batch estimation using regression techniques proved as successful as conventional recursive approaches in applications such as ballistic trajectory launch point estimation, adaptive flight control, and radio frequency target triangulation. Toledo et al. (2007) [88] used locally weighted regression to estimate vehicle trajectories from a set of positional data obtained at discrete time intervals with possible missing observations. However, only regular and relatively small time intervals are considered. Finally, Agrawal, Singh (2008) [89] apply nonlinear (second order polynomial) regression for trajectory estimation of a moving object in space but the heteroscedascity of measurement errors was not addressed.

1.5 Approach and Contributions

The issues pertaining to maritime surveillance such as low and irregular data rates, the availability of surveillance data, the large measurement error in form of elliptical uncertainty, the absence of knowledge about the motion model, and the known maneuver behaviour of ships in the surveillance area, suggest a possibility of using batch processing and piecewise regression techniques for data fitting for estimation of trajectories, rather than commonly used recursive methods.

The first approach to trajectory estimation is based on fitting the measurements with heteroscedastic and correlated errors to a piecewise cubic spline with the unknown number of knots (as points of maneuvers), where the Bayesian inference will be used in nonparametric regression modeling. The approach is based on the work of [1], however assuming data with known heteroscedastic and correlated errors as well as accommodating for decreasing values of the predictor variable (i.e. accounting for backwards turning maneuvers). Further, the regression will be performed on two nonparametric models, i.e. one which which uses the error projections and the other one which completely accounts for measurement uncertainty. Complete accounting for the measurement uncertainty anticipates better performance than other batch stochastic optimization algorithms for trajectory fitting which use error approximations such as the one reported in [2].

Another proposed approach is the stochastic linear filtering approach in which the motion model is constructed from the common knowledge about the ship motion. Specifically, the Integrated Ornstein-Uhlenbeck (IOU) process and the random walk (RW) process will be combined to describe the motion. The assumptions on linear modeling and Gaussian errors will allow for the use of Kalman filtering and Rauch-Tung-Striebel (RTS) optimal smoothing to produce the positional estimates.

This thesis presents the following results in trajectory estimation from data with heteroscedastic and correlated measurement errors:

• A novel batch single ship trajectory estimation algorithm is introduced. It assumes no

knowledge about the ship motion model, while assuming standard ship maneuvers. As a batch processing algorithm, it addresses the availability of data for processing, an issue specific to maritime surveillance applications. The algorithm developed here is based on the algorithm found in Di Matteo et al. [1] which employs Bayesian inference to spline fitting using a nonparametric regression model. However, the work of Di Matteo et al. [1] fails to accommodate explicitly for heteroscedastic and correlated errors, as well as for decreasing values of predictor variable (in case of turning maneuvers), which are inherent to the problem of trajectory estimation in maritime surveillance. Additionally, Bayesian inference is performed for nonparametric regression models with two different choices of response and predictor variables. A complete account for the measurement uncertainty by this algorithm guarantees improved performance over other batch stochastic optimization algorithms for ship trajectory estimation based on error approximations such as the one reported in [2].

- Assessment of the robustness of the proposed algorithm by sensitivity analysis is carried out. This includes verifying if the results of the Bayesian analysis remain unchanged in the presence of perturbation in prior parameters.
- The first time development, implementation and assessment in the open literature of the stochastic linear filtering algorithm which employs the Integrated Ornstein-Uhlenbeck processes in conjunction with Kalman filtering and Rauch-Tung-Striebel optimal smoothing, is provided. It is presented here that such motion modeling is possible only when the IOU process is used in the specific combination with the random walk process. The assessment of the IOU-based trajectory estimation algorithm with special focus on the importance of the modeling of measurement uncertainty was reported by Hadzagic and Michalska (2010) [3], [4].

- Sensitivity analysis of the linear stochastic filtering algorithm with respect to perturbations in model parameters is performed.
- Assessment of the efficiency of both algorithms for various ground truth scenarios in terms of the chosen performance measures is carried out. A separate analysis is performed for each nonparametric regression model.
- Comparative analysis of the Bayesian nonparametric regression spline based algorithm (both models) with the batch linear stochastic filtering algorithm for trajectory estimation is carried out.
- Respective comparative analysis of the Bayesian nonparametric regression spline algorithm (both models) and the IOU-model based algorithm with another batch stochastic optimization algorithm, namely a genetic algorithm, used to form a trajectory from heteroscedastic data [3] is provided.
- Analysis and discussion of the results of the comparisons, conclusions about the novel algorithms are provided, and the possible future research avenues are outlined.
- A software platform for maritime surveillance which includes the ground truth scenario integrated with three algorithms for trajectory estimation: the IOU-KF-RTSS algorithm, Bayesian spline regression algorithm, and the GA-based trajectory estimation algorithm, which are all conceptually different. It is developed in C++ and includes 6701 lines of code. The Bayesian spline regression algorithm represents a modified open source C implementation, which is available here [5]. The modification involves injection of the measurement error information and the rotation of coordinate system to accommodate maneuvers in x y plane when the position in x-direction is decreasing.

The algorithms presented in this thesis apply to a broader class of problems than the ones defined by the problem of batch stochastic ship trajectory estimation, since the curve fitting algorithms, as well as the stochastic linear filtering algorithm employing the IOU as a model process, are found in applications in the fields of finance (e.g. estimation of integrated volatility in stochastic volatility models) and biostatistics (e.g. derivative tracking of viral and other marker data). Other applications for batch trajectory estimation algorithms include specifying behavioral patterns, relationships between locations observed at adjacent times, or forming records of the history of past locations.

Chapter 2

Problem Statement

Two different Bayesian approaches to trajectory estimation from batch of data with heteroscedastic and correlated measurement errors are considered: (1) a stochastic optimal filtering approach and (2) a nonparametric regression based curve fitting approach using Bayesian inference. Hence, two formulations of the trajectory estimation problem are stated here. Prior to that, it is useful to characterize the measurement data and the ship motion features with respect to goals of ship trajectory estimation.

2.1 Characterization of the Measurement Data, Measurement Errors and the Surveillance Region

The multisensor surveillance data are assumed to originate from a single ship which is moving in an environment with no clutter or countermeasures, therefore no data association procedure is considered. In maritime surveillance, a single measurement indicating ship's position is referred to as the *ellipse contact report*. The ellipse contact report (measurement) vector will be defined below. As illustrated in Fig. 2.1, the ellipse contact report is graphically perceived as an elliptical area of uncertainty (AOU) for the ship's location, i.e. the minimum area having a specified probability of containing the measurement, see Appendix A. The positional observation is the centre of the ellipse. The positional errors exist in both directions of the x - y Euclidean surveillance region. However, due to the assumptions on dissimilar multisensor environment, the errors do not follow directions of axes of the surveillance region causing the errors in x and y directions to be correlated. Furthermore, the errors in x and y are assumed nonconstant across measurements in time, rendering the measurement error heteroscedastic [61], [11]. It is assumed that the ellipse specifies two-sigma or equivalently a 96% containment region, i.e. the probability that the target is found inside the area of the specified ellipse is assumed to be 0.96, see Appendix A.

It is assumed that a batch of surveillance data $\mathbf{Z} = {\{\mathbf{z}_k\}_{k=1}^n}$, obtained over the known discrete time interval $[t_1, t_n]$, is available, where \mathbf{z}_k is the k-th ellipse contact report vector, obtained at a known irregular time instant $t_k, t_k \in [t_1, t_n]$, and is defined as

$$\mathbf{z}_{\mathbf{k}} = \begin{bmatrix} \mathbf{y}_{\mathbf{k}} & I_{sen}^{k} \end{bmatrix} = \begin{bmatrix} x_{k} & y_{k} & a_{k} & b_{k} & \theta_{k} \end{bmatrix}.$$
 (2.1)

The total number of measurements indexed in time is n. The two-dimensional positional subvector $\mathbf{y}_{\mathbf{k}}$ comprises of the components x_k and y_k and represents the centre of the ellipse, i.e. the 2-D measured geographical position, (x_k, y_k) , with x_k and y_k being longitude and latitude measured in degrees, respectively. The subvector, I_{sen}^k , contains the information which characterizes the sensor error. The components a_k and b_k represent the magnitudes of the semi-major and semi-minor axis of the ellipse measured in nautical miles [nm], respectively, while θ_k is the angle of the orientation of the ellipse, measured in degrees clockwise from the true north, as shown in Fig. 2.1. It is assumed that $a_i \neq a_j$, $b_i \neq b_j$,



Fig. 2.1 The positional measurement with elliptical area of uncertainty.

and $\theta_i \neq \theta_j$ for $i \neq j$. Furthermore, given the ellipse AOU confinement probability P, with the assumption on the normal distribution of the measurement error, the vector I_{sen}^k can be transformed into the covariance R_k of a bivariate Gaussian distribution $\mathcal{N}(0, R_k)$ which characterizes the error of measurement k, see Appendix A. The mean is identified as the ellipse centroid, while the covariance matrix can be obtained from the parameters of the ellipse geometry. The covariance matrix, R_k , of the bivariate Gaussian distribution corresponding to the position \mathbf{y}_k with the ellipse orientation θ_k , and the magnitudes of the semi-major and semi-minor axes, a_k and b_k , respectively, is computed as

$$R_{k} = \begin{bmatrix} \cos \theta_{k} & -\sin \theta_{k} \\ \sin \theta_{k} & \cos \theta_{k} \end{bmatrix} \begin{bmatrix} a_{k}^{2}/2 & 0 \\ 0 & b_{k}^{2}/2 \end{bmatrix} \begin{bmatrix} \cos \theta_{k} & \sin \theta_{k} \\ -\sin \theta_{k} & \cos \theta_{k} \end{bmatrix}$$
(2.2)

The background on the measurement error ellipse, including the equations which relate the characteristics of the ellipse geometry of the positional measurement $\mathbf{y}_{\mathbf{k}}$ and the corresponding covariance matrix R_k of the Gaussian bivariate error distribution are found in Appendix A.

Since the error ellipses are small in size as compared with the size of the surveillance region, a local flat-Earth approximation in the plane tangent to the measured position, (x_k, y_k) (i.e. the ellipse centroid), is used for constructing the bivariate Gaussian error distribution, $\mathcal{N}(0, R_k)$.

The surveillance region is assumed to be a two-dimensional Euclidean space where the Cartesian coordinates are used to express distances in north-south and east-west directions, respectively. The positional measurements are obtained in geographical coordinates, latitude and longitude, whose definitions assume Earth as a sphere. Nevertheless, the Earth can be considered flat for short distances, and calculating the straight line distance between two points can be done using flat-Earth approximation formulas. The details on flat-Earth approximation are provided in Appendix A.

2.2 Characterization of Model Features and Estimation Goals

The ship motion model is assumed unknown. Only general assumptions on the surface ship navigation are used to make assumptions on the possible shapes of trajectories.

It is assumed that a single surface ship motion occurs in the 2-D Euclidean plane of geographical coordinates, longitude and latitude. A ship is assumed to follow standard trajectories which include maneuvers such as zig-zag and turning circles, [90]. The zig-zag maneuver corresponds to navigating through water between waypoints (i.e., points of maneuver or change-points) with constant velocity. The number of change-points as well as their locations along the trajectory are assumed unknown. The ship's motion does not assume acceleration. Since the motion occurs in horizontal plane of longitude and latitude, a flat-Earth approximation is assumed for calculating the positions along the trajectory.

Relative to the knowledge on ship navigation and the rest of the assumptions on the ship motion, the motion model in the stochastic linear filtering algorithm is constructed. On the basis of the same knowledge, the trajectory is assumed as a spline for another Bayesian approach.

2.3 The Linear Stochastic Bayesian Filtering Problem

Let the ship state vector be denoted by $\mathbf{x}_{\mathbf{k}} \in \mathbb{R}^n$, where n = 4 is the dimension of the state vector and \mathbb{R} is a set of real numbers, $k \in \mathbb{N}$ is time index, and \mathbb{N} is the set of natural numbers. The index k is assigned to a continuous time instant t_k , and the sampling interval $t_k - t_{k-1}$ may be time dependent. It is assumed that the state evolves according to the discrete-time stochastic model (1.9):

$$\mathbf{x}_{\mathbf{k}} = f_{k-1}(\mathbf{x}_{\mathbf{k-1}}, \mathbf{w}_{\mathbf{k-1}})$$

where f_{k-1} is a "known" linear function of the state \mathbf{x}_{k-1} , and \mathbf{w}_{k-1} is a process noise sequence. The model is constructed from the common knowledge on surface ships' behaviour with respect to the assumptions in Section 2.2. The model development will be presented in Section 3.

The measurements are assumed to be related linearly to the target state, i.e.,

$$\mathbf{y}_{\mathbf{k}} = M\mathbf{x}_{\mathbf{k}} + \epsilon_{\mathbf{k}};$$
 $k = 1, 2, \dots n;$ $t_{k+1} > t_k > t_0$ (2.3)

where $\mathbf{y}_{\mathbf{k}}$ is the 2-dimensional position vector, M is a known 2 × 4 matrix,

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$
 (2.4)

 $\mathbf{x}_{\mathbf{k}}$ is the 4-dimensional state vector of the process defined as

$$\left[\begin{array}{ccc} x_k & y_k & u_k & v_k \end{array}\right]^T \tag{2.5}$$

where x_k and y_k , and u_k and v_k are 2D-position and 2D-velocity components, respectively, and $\epsilon_{\mathbf{k}}$ is a 2-dimensional vector of independent white Gaussian sequence, i.e. $\epsilon_{\mathbf{k}} \sim N(0, R_k), R_k > 0$ where R_k is known measurement covariance matrix obtained from sensor information characteristics and given by (2.2). The covariance matrices are assumed nonconstant, i.e., for the positions obtained at the time instants j and k, $R_j \neq R_k$. The positional measurement at time k is the position of the ship at time k plus a bivariate Gaussian error with zero mean and the covariance R_k . The initial target state is assumed to have a known probability density function, $p(x_0)$, which is assumed independent of noise sequences.

Objective. The objective in this approach is to estimate the state $\mathbf{x}_{\mathbf{k}} \sim \mathcal{N}(\mathbf{x}_{\mathbf{k}}; \mu_{\mathbf{k}}, \Sigma_{\mathbf{k}})$ based on the sequence of all available positional measurements $\mathbf{Y}_{k} \triangleq \{\mathbf{y}_{\mathbf{i}}, i = 1, \dots, k\}$ from the posterior distribution $p(\mathbf{x}_{\mathbf{k}} | \mathbf{Y}_{k})$, given the initial probability distribution of the state vector, $p(\mathbf{x}_{0}) \sim \mathcal{N}(0, \Sigma_{0})$.

2.4 The Nonparametric Regression Spline Problem

Relative to the characterization of the measurement errors and the assumptions on the ship motion (i.e. navigation with constant speed along straight lines between waypoints and performing turning maneuvers at waypoints), the trajectory is assumed to be of the form of piecewise cubic splines. Hence, the ship trajectory estimation can be stated as a nonparametric curve (spline) regression problem from a batch of heteroscedastic and correlated surveillance data.

Assume mutually independent and normally distributed pairs of positional data, $\mathcal{D} = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, obtained at known irregular time intervals t_k , where $k = 1, \ldots, n$, for which the conditional densities satisfy

$$Y_k|X_1,\ldots,X_n \sim \mathcal{N}(y_k|f(x_k),\sigma) \tag{2.6}$$

where f is a real-valued function on [a, b], σ is a known conditional marginal covariance of

jointly distributed variables X_k and Y_k , and x_k and y_k represent the observed values of the random variables X_k and Y_k following the standard notation.

The data are considered to be generated by two possible nonparametric models:

• Model 1: A nonparametric regression model in which each response variable, the position in each dimension, x_k and y_k , is related to the predictor variables $T = \{t_1, \ldots, t_n\}$ by its own regression model (through some deterministic functions, f_x and f_y , and some additive random errors ϵ_x and ϵ_y , respectively), i.e.,

$$\mathbf{y}_{\mathbf{k}} = f(t_k) + \epsilon_{\mathbf{k}} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} f_x(t_k) \\ f_y(t_k) \end{bmatrix} + \begin{bmatrix} \epsilon_{xk} \\ \epsilon_{yk} \end{bmatrix}$$
(2.7)

where $\epsilon_k \sim \mathcal{N}(0, R_k)$ for k = 1, ..., n, n being the total number of measurement pairs, and R_k is known nonconstant and nondiagonal matrix.

• Model 2: A nonparametric regression model which relates each response variable y_k , i.e. the position in y-direction to the predictor variable x_k , i.e. the position in xdirection, through a deterministic function f and the additive random error ϵ_k . The predictor variable X_k is assumed to be observed through random variable Z_k , the so-called latent regressor, with error ϵ_{yk} , therefore the model can be written as

$$y_k = f(z_k) + \epsilon_{yk} \tag{2.8}$$

$$x_k = z_k + \epsilon_{xk} \tag{2.9}$$

The variable Z_k is considered independent from the measurement error ϵ_{xk} , and ϵ_{xk} and ϵ_{yk} are diagonal elements of the known measurement covariance matrix R_k .

For both nonparametric regression models, it is assumed that f(x) defined on [a, b] can be approximated between a and b by a cubic spline with an unknown number of l knots at unknown locations $\xi = (\xi_1, \ldots, \xi_l)$, where $a < x_{(1)} < \xi_1 \leq \cdots \leq \xi_k < x_{(n)} < b$ where $x_{(1)}$ and $x_{(n)}$ are the minimum and the maximum sampling points, respectively. A cubic B-spline is a piecewise cubic polynomial over four intervals which is defined by a recursive convolution formula in terms of fourth-order divided differences. A grid of l + 4 knots generates l cubic B-splines. For an extensive background on B-splines, see [91].

Let $b_j(x)$, j = 1, ..., k + 2, denote the *j*-th function in a cubic B-spline basis with natural boundary constraints, i.e. linear outside the [a, b] interval.

Let $B_{l,\xi}$ be the matrix such that the (i, j)-th component is $b_{l,\xi}^j(x_i)$ which depends on the knot configuration, (l,ξ) . For notational simplicity, denote $b_{l,\xi}^j(x_i) = b_j(x_i)$

Under these assumptions, the function f can be approximated as

$$f(x) = \sum_{j=1}^{l+2} b_j(x)\beta_j(x)$$
(2.10)

where $\beta = (\beta_1, \dots, \beta_{k+1})$, i.e. $B_{l,\xi}\beta = f(x) \equiv (f(x_1), \dots, f(x_n))$ at observed data points.

Objective. The objective of this approach is to obtain the curve estimate $\hat{f}(x)$ using Bayesian inference, where $\hat{f}(x)$ represents the pointwise posterior mean,

$$\hat{f}(x) = E\{f(x)|y\} = E\{E\{f(x|y,\xi,k\}\} \approx B_{l,\xi}\hat{\beta}$$
(2.11)

The number of knots, l, and the set of knot locations, $\xi = (\xi_1, \ldots, \xi_l)$ are to be determined during the estimation procedure. For nonparametric Model 1, the inference involves estimation of f_x and f_y .

Chapter 3

Trajectory Estimation Using Linear Stochastic Filtering with Integrated Ornstein Uhlenbeck Process

This Chapter presents, as recently reported in Hadzagic and Michalska (2010) [3], for the first time in the open literature, the development and the description, of the stochastic linear filtering algorithm which employs the Integrated Ornstein-Uhlenbeck processes in conjunction with Kalman filtering [50]. The trajectory is obtained as a solution to a linear stochastic optimal filtering problem. The algorithm employs a combination of two stochastic processes, namely the Integrated Ornstein-Uhlenbeck process (IOU) and the random walk (RW), process to describe the ship's motion. The assumptions on linear modeling and bivariate Gaussian distribution of measurement errors allow for the use of Kalman filtering and Rauch-Tung-Striebel optimal smoothing.

3.1 Development of the Ship Motion Model

In physical modeling the Wiener process is usually used to represent the position of the Brownian particle. Although the sample paths of the Wiener process are continuous functions, they are considered as nowhere differentiable, i.e the velocity of the Brownian particle cannot be defined. Ornstein and Uhlenbeck [92] compensated for this drawback of the Wiener process by modeling directly the velocity of the Brownian particle. The resulting process is known as the Ornstein-Uhlenbeck process.

Assume that the motion process $\{X_t, t \ge 0\}$ in the *n* dimensional state space *S* satisfies the linear (Itô) stochastic differential equation (LSDE) [93],

$$dX_t = [F_t X_t + a_t]dt + G_t dw_t \tag{3.1}$$

where F_t and G_t are $n \times n$ deterministic matrices, a_t is a deterministic *n*-dimensional vector, and w_t is an *n*-dimensional vector of independent Wiener processes with unit variance. This is a natural class of models for target motion since if X_0 is Gaussian, the solution to (3.1) is a Gaussian-Markov process [93]. Models from this class of processes can be employed in continuous-discrete Kalman filter recursion. The functions F_t , a_t and G_t have natural interpretations. The functions F_t and a_t determine the drift of the motion process, and G_t determines the diffusion. If the drift and the diffusion are constant, i.e. if $F_t = F$, $a_t = a$, and $G_t = G$, for $t \ge 0$, then (3.1) becomes

$$dX_t = [FX_t + a] dt + G dw_t \tag{3.2}$$

Under the assumption that the initial condition X_0 is normally distributed, i.e. $X_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$, and with X_0 independent of $\{w_t; t \ge 0\}$, (3.2) can be solved explicitly using

stochastic integration to obtain the motion process X_t as follows:

$$X_t = e^{Ft} \left[X_0 + \int_0^t e^{-Ft} (a \, d\tau + G \, dw_\tau) \right]$$
(3.3)

The mean, the variance, and the covariance of the process X_t can be obtained as:

$$E\{X_t\} = e^{Ft} E\{X_0\} + \int_0^t e^{F(t-\tau)} a \, d\tau$$

= $e^{Ft} \mu_0 + \int_0^t e^{F(t-\tau)} a \, d\tau$ (3.4)

$$Var\{X_{t}\} = Var\left\{e^{Ft}\left[X_{0} + \int_{0}^{t} e^{-F\tau}G \,dw_{\tau}\right]\right\}$$
$$= e^{Ft}\Sigma_{0}(e^{Ft})^{T} + \int_{0}^{t} e^{F(t-\tau)}GG^{T}(e^{F(t-\tau)})^{T} \,d\tau$$
(3.5)

$$Cov[\{X_{t_0}, X_{t_1}\}] = Cov[e^{Ft_0}X_0, e^{Ft_1}X_0] + Cov[\int_0^{t_0} e^{F(t_0-\tau)}G \, dw_\tau, \int_0^{t_1} e^{F(t_1-\tau)}G \, dw_\tau] = e^{Ft_0}\Sigma_0(e^{Ft_1})^T + \int_0^{\min(t_1, t_0)} e^{F(t_0-\tau)}GG^T(e^{F(t_1-\tau)})^T \, d\tau$$
(3.6)

It follows from (3.3) that the motion process X_t is a Gaussian process with the mean and the covariances given by (3.4)-(3.6).

For X_t to model the velocity, (3.2) needs to be modified so that the drift is linear in tand induce motion toward the origin. Thus, it is assumed that $F = -\beta I$, where I is the identity matrix, a = 0, $G = \sigma I$, β and σ are scalar parameters. Then (3.2) becomes

$$dX_t = -\beta X_t \,\mathrm{d}t + \sigma \,dw_t \tag{3.7}$$

The motion process X_t that satisfies (3.7) is called the Ornstein-Uhlenbeck (OU) process. If the motion takes place in a two-dimensional Euclidean space, the OU process can be thought of as Brownian motion with a drift that tends to move the Brownian particle toward the origin. Also, it can be seen that the drift is proportional to the displacement of the particle from the origin. The solution to (3.7) is

$$X_t = X_0 e^{-\beta t} + C \int_0^t e^{-\beta(t-\tau)} dw_\tau$$
(3.8)

From (3.8), the mean, the variance and the covariance of the OU process can be easily obtained. The mean is computed as

$$E\{X_t\} = 0 \tag{3.9}$$

and the variance is given by

$$E\{X_t^2\} = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})$$
(3.10)

Thus, as $t \to \infty$ the OU process has an asymptotic stationary distribution which is normal with zero mean and the variance $\frac{\sigma^2}{2\beta}$, i.e. the variance of the OU process is increasing over time and reaches a limiting constant maximum value $\frac{\sigma^2}{2\beta}$. This value is sometimes called the stationary velocity variance. The covariance of the process decreases exponentially, i.e.

$$E\{X_t X_s\} = \frac{\sigma^2}{2\beta} e^{-\beta|t-s|} \tag{3.11}$$

A generalization of the OU process is the Integrated Ornstein-Uhlenbeck (IOU) process $\{Y_t\}$ which can be used to model the position of a Brownian particle when started at the origin, $Y_0 = 0$

$$Y_t = \int_0^t X_s \, ds \tag{3.12}$$

The integral over a Gaussian process is again a Gaussian process, therefore Y_t is completely characterized by its mean and variance

$$E\{Y_t\} = E\{\int_0^t X_s \, ds\} = \int_0^s E\{X_s\} \, ds = 0 \tag{3.13}$$

$$E\{Y_t^2\} = \frac{\sigma}{\beta^3}(e^{-\beta t} + \beta t - 1), \qquad (3.14)$$

and its covariance, see [94] for details:

$$E\{Y_tY_s\} = \frac{\sigma^2}{\beta^2}t + \frac{\sigma^2}{2\beta^3}[e^{-\beta t} - 1 + e^{-\beta s} - e^{-\beta(s-t)}]$$
(3.15)

Here, the (integrated) Ornstein-Uhlenbeck process is used to represent the ship's motion. Specifically, if the state space is four-dimensional and the state is defined as

$$\left[\begin{array}{ccc} x_t & y_t & u_t & v_t \end{array}\right]^T \tag{3.16}$$

where x_t and y_t , and u_t and v_t are 2D-position and 2D-velocity components, respectively,

and

$$F = \begin{bmatrix} I & 0 \\ 0 & -\beta I \end{bmatrix} \qquad \text{and } G = \begin{bmatrix} 0 & 0 \\ 0 & -\sigma I \end{bmatrix}$$

then (3.7) can be expressed in terms of a system of LSDEs:

$$dx_t = u_t dt \tag{3.17}$$

$$dy_t = v_t dt \tag{3.18}$$

$$du_t = -\beta u_t + \sigma dW_t \tag{3.19}$$

$$dv_t = -\beta v_t + \sigma dW_t \tag{3.20}$$

Equations (3.17)-(3.20) describe the (integrated) Ornstein-Uhlenbeck process as the ship motion model. The initial condition for this process is assumed as $X_0 \sim N(\mu_0, \Sigma_0)$ so that $X \sim N(\mu_t, \Sigma_t), \forall t \ge 0$. According to (3.9) and (3.10), the mean velocity in the direction of the x axis is given by

$$E\{u_t\} = u_0 e^{-\beta t} (3.21)$$

while the covariance of the x-direction velocity process u_t is given by

$$Cov\{u_t u_s\} = \frac{\sigma^2}{2\beta} e^{-\beta I|t-s|} \text{ for } t \le s.$$
(3.22)

For t = s, the variance of the process u_t is

$$E\{u_t^2\} = \frac{\sigma^2}{2\beta}, \qquad \forall t \ge 0 \qquad (3.23)$$

i.e. the stationary limiting velocity variance. It is not difficult to see that the same expressions hold for v(t). The parameter β controls the rate at which the ship velocity changes occur. Hence, the velocity change rate can be controlled by the choice of the parameter β .

The IOU process is driven by Gaussian white noise (as the derivative of the Wiener process) and hence is "physically non-realizable" [95]. On the other hand, the Random Walk (RW) family of models is physically realizable and can represent the ship motion. As described in the next Section and with the goal of producing a more realistic model, the IOU process will be represented as a random walk by a specific choice of parameters.

3.1.1 Random Walk

To relate the IOU motion process to the RW process it is first helpful to understand how a ship motion can be identified with a RW.

To this end, let the initial ship position be the origin of the (X = East, Y = North)coordinate system, see Figure 3.1. The initial ship course (i.e. the ship heading at t = 0) is then chosen randomly from the uniform distribution over the interval $[0^o, 360^o]$, while the ship speed is assumed to be a known constant V. All ship courses θ_i are samples drawn independently from the probability density function

$$f(t;\alpha) = \begin{cases} 0 & t < 0\\ \alpha e^{-\alpha t} & t \ge 0 \end{cases}$$

where α is the rate parameter. Hence, the ship course is understood to be an exponential random variable with mean value $1/\alpha$ (i.e. the average time between course changes is $1/\alpha$ and the average rate at which the velocity changes occur is α). Since such a model is



Fig. 3.1 Random Walk process with the initial course that is chosen randomly from $U[0^0, 360^0]$ and $\theta \sim exp(1/\alpha)$, i.e. average time between the course changes is $1/\alpha$.

non-Gaussian, Kalman filtering cannot be applied directly which justifies its approximation in terms of an IOU model which is Gaussian.

For the purpose of one such approximation, define the processes R_t and V_t to represent the range and speed:

$$R_t^2 = x_t^2 + y_t^2 (3.24)$$

$$V_t^2 = u_t^2 + v_t^2 (3.25)$$

(3.26)

Also, let

$$h(\beta t) = e^{-\beta t} - 1 + \beta t \tag{3.27}$$

From (3.13)-(3.15) it follows that, $\forall t$,

$$E\{x_t\} = 0$$

$$Var\{x_t\} = E\{x_t^2\} = \frac{\sigma^2}{\beta^3}h(\beta t)$$

$$E\{y_t\} = 0,$$

$$Var\{y_t\} = E\{y_t^2\} = \frac{\sigma^2}{\beta^3}h(\beta t)$$

$$Cov\{x_t, y_t\} = 0$$

Furthermore,

$$E\{R_t^2\} = E\{x_t^2 + y_t^2\}$$
(3.28)

$$= E\{x_t^2\} + E\{y_t^2\}$$
(3.29)

$$=\frac{2\sigma^2}{\beta^3}h(\beta t) \tag{3.30}$$

$$E\{V_t^2\} = 2E\{u_t^2\} = \frac{\sigma^2}{\beta}$$
(3.31)

3.1.2 Link between the IOU Model and the RW Model

As the number of parameters in both the RW and the IOU models is equal to two, one can attempt to approximate the RW model by the IOU model by way of equating the first two moments of both processes. To this end, it is helpful to evaluate the averages $E\{R_t^2\}$ for both models, which yields

$$E\{R_t^2\}_{IOU} = \frac{2\sigma^2}{\beta^3}(e^{-\beta t} - 1 + \beta t)$$
(3.32)

$$E\{R_t^2\}_{RW} = \frac{2V^2}{\alpha^2}(e^{-\beta t} + \beta t - 1)$$
(3.33)
The above expressions coincide if the parameters are related as follows:

$$\beta = \alpha \tag{3.34}$$

$$\frac{\sigma^2}{\beta} = V^2 \text{ or equivalently } \sigma^2 = \alpha V^2 \tag{3.35}$$

Equations (3.34) and (3.35) provide the main link between the IOU model and the real world ship motion. Invoking (3.36)-(3.40), it is now easy to see that (3.35) entails equality of the first two moments of the RW and the IOU processes, i.e.,

$$E\{x_t\}_{RW} = E\{x_t\}_{IOU} = 0 \tag{3.36}$$

$$E\{y_t\}_{RW} = E\{y_t\}_{IOU} = 0 \tag{3.37}$$

$$Var\{x_t\}_{RW} = Var\{x_t\}_{IOU} = E\{x_t^2\}$$
$$= \frac{V^2}{\alpha^2}(e^{-\alpha t} - 1 + \alpha t)$$
(3.38)

$$Var\{y_t\}_{RW} = Var\{y_t\}_{IOU} = E\{y_t^2\}$$
$$= \frac{V^2}{\alpha^2}(e^{-\alpha t} - 1 + \alpha t)$$
(3.39)

$$Cov\{x_t, y_t\} = 0$$
 (3.40)

From the equality of moments (3.36)-(3.40), and because the Gaussian processes are fully characterized by their first and second moments, it follows that the resulting IOU process gives the best possible second order approximation to the RW model among all Gaussian models.

In terms of (3.17)-(3.20) and with given initial velocity conditions, the mean state and the variance of the IOU process used to model ship motion, are obtained using (3.4) and (3.5):

$$E\{X_t\} = e^{Ft}\mu_0 \tag{3.41}$$

 $\mathbf{52}$

$$Var\{X_t\} = e^{Ft} \Sigma_0 (e^{Ft})^T + \sigma^2 \begin{bmatrix} b_{11}(t)I & b_{12}(t)I \\ b_{21}(t)I & b_{22}(t)I \end{bmatrix}$$
(3.42)

where

$$e^{Ft} = \Phi(t) = \begin{bmatrix} I & \frac{1}{\beta}(1 - e^{-\beta t})I\\ 0 & e^{-\beta t}I \end{bmatrix}$$
(3.43)

$$b_{11}(t) = \frac{1}{\beta^2} \left[t - \frac{2}{\beta} (1 - e^{-\beta t}) + \frac{1}{2\beta} (1 - e^{-2\beta t}) \right]$$
(3.44)

$$b_{12}(t) = b_{21}(t) = \frac{1}{\beta^2} [(1 - e^{-\beta t}) - \frac{1}{2}(1 - e^{-2\beta t})]$$
(3.45)

$$b_{22}(t) = \frac{1}{2\beta} (1 - e^{-2\beta t}) \tag{3.46}$$

3.2 Continuous-Discrete Kalman Filter Recursion for the IOU Motion Process

In this Section, a continuous-discrete Kalman filtering algorithm is developed to produce filtered estimates of ship position as it varies in time.

In continuous-discrete Kalman filtering, as applied to ship trajectory estimation it is assumed that the actual motion takes place in continuous time, while observations are only available at discrete instants of time t_k . The measurement equation is usually given by:

$$y_k = Mx_k + \epsilon_k; \ k = 1, 2, \dots; \qquad t_{k+1} > t_k > t_0 \tag{3.47}$$

where y_k is the 2-dimensional vector of observations in geographical coordinates, M is a given 2×4 matrix,

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$
 (3.48)

 x_k is the 4-dimensional state vector of the process as defined in (2.5), and ϵ_k is a 2dimensional vector of white Gaussian sequence, $\epsilon_k \sim N(0, R_k)$, $R_k > 0$ with known R_k . With reference to the practical problem of ship position estimation in a dissimilar multisensor environment, see Section 2, the covariance matrix, R_k , must generally be considered as non-diagonal and non-constant, due to different sizes and orientations of the elliptical AOUs, rendering the error heteroscedastic and correlated. The covariance matrix R_k is determined from the parameters which define the geometry of the area of uncertainty for the measurement \mathbf{y}_k , i.e. from the values a_k, b_k and θ_k according to (2.2) as

$$R_{k} = \begin{bmatrix} \cos \theta_{k} & -\sin \theta_{k} \\ \sin \theta_{k} & \cos \theta_{k} \end{bmatrix} \begin{bmatrix} a_{k}^{2}/2 & 0 \\ 0 & b_{k}^{2}/2 \end{bmatrix} \begin{bmatrix} \cos \theta_{k} & \sin \theta_{k} \\ -\sin \theta_{k} & \cos \theta_{k} \end{bmatrix}$$

With the above representation of the covariance matrix in terms of the ellipse parameters, the extended measurement mean vector d_k for the contact y_k (i.e. the centroid of the ellipsoid) and the measurement information matrix α_k , can be defined for convenience of the Kalman filter recursion as

$$d_k = \begin{bmatrix} y_k & 0_2 \end{bmatrix}^T, \qquad \qquad \alpha_k = \begin{bmatrix} R_k^{-1} & 0_{2\times 2} \\ 0_{2\times 2} & 0_{2\times 2} \end{bmatrix}$$
(3.49)

Since the components of the contact report data are measured in different units, the conversion of all input surveillance data to the algorithm's format must be done so that the units agree.

3.2.1 The Algorithm

The continuous-discrete Kalman recursion as applied to the IOU process with state $X \sim \mathcal{N}(\mu, \Sigma)$ with the Gaussian p.d.f. $p(x; \mu, \Sigma)$ is summarized by the following sequence of steps [93]:

Algorithm 1 IOU-KF-RTSS algorithm	_
1: Initialize with $p(0, x) = \mathcal{N}(x; \mu_0, \Sigma_0)$, with	_

$$\mu_{0} = \begin{bmatrix} 0\\0\\0\\0 \end{bmatrix}, \qquad \Sigma_{0} = \begin{bmatrix} 0_{2\times 2} & 0_{2\times 2}\\0_{2\times 2} & \frac{\sigma^{2}}{2\beta}I_{2\times 2} \end{bmatrix}$$
(3.50)

2: Do motion update, i.e. using (3.4)-(3.5), the predicted state mean, compute $\mu_{k+1|k}$, and the state covariance, $\Sigma_{k+1|k}$,

$$\mu_{k+1|k} = \Phi(t_{k+1}, t_k) \mu_{k|k}$$

$$\Sigma_{k+1|k} = \Phi(t_{k+1}, t_k) \Sigma_{k|k} \Phi(t_{k+1}, t_k)^T + C(t_{k+1}, t_k)$$

where $C(t_{k+1}, t_k)$ is the second term in (3.5) and $\Phi(t_{k+1}, t_k)$ is defined by (3.43). 3: Do information update, i.e. compute

$$\Sigma_k = \Sigma_{k+1|k} [I + \alpha_k \Sigma_{k+1|k}]^{-1}$$

$$\mu_{k+1|k+1} = \mu_{k|k} + \Sigma_{k+1|k+1} \alpha_{k+1} (d_{k+1} - \mu_{k+1|k})$$

The overview of the algorithm is given in Fig. 3.2.

Since the ship motion is assumed to occur in a horizontal plane and the measurement data assume a spherical Earth, a flat-Earth approximation is used for calculating (updating) the positions. This involves a projection down from the tangent plane of the estimated data to sphere in order to obtain both the positional vector and the associated covariance matrix on the sphere. From the point of tangency and the corresponding covariance matrix, the point of tangency on the sphere (lon, lat) in radians and the displacements of longitude and latitude are obtained. The flowchart in Fig. 3.3 shows the principal steps in calculating the trajectory on the sphere. The single-point expansion is added functionality to the algorithm which provides position, heading and speed estimated for a ship at a specified time. The algorithm flowchart for the single-point expansion is given in Fig. 3.4.

3.3 Rauch-Tung-Striebel Fixed Optimal Smoothing

To improve on the quality of the estimated state the output of the continuous discrete Kalman filter is further passed through a smoothing filter. A particularly convenient choice seems to be discrete-time fixed interval Rauch-Tung-Striebel (RTS) optimal smoothing algorithm as it is known to be insensitive to computational round-off errors present when processing large batches of data [18].

With the same assumptions on the system as in Sections 3.1 and 3.2, and assuming that the output of the Kalman filter $(\mu_{k|k}, P_{k|k})$ is available, the RTS algorithm seeks a recursion for $\mu_{k|N}$, k < N, given the observations Y_N , where N is the time of the last measurement received. The latter is derived in [18] and cited here in the form:

$$\mu_{k|N} = \mu_{k|k} + S_k [\mu_{k+1|N} - \mu_{k+1|k}], \qquad (3.51)$$

where $S_k = P_{k|k} \Phi_{k+1,k} P_{k+1|k}^{-1}$ and $\mu_{N|N}$ is the filtered state estimate for k = N - 1, while $\Phi_{k+1,k}$ is the state transition matrix of the IOU process defined by (3.43). The above recursion is performed backwards in time with the initial condition $\mu_{N|N}$. The recursion for the smoothing error covariance matrix is given by

$$P_{k|N} = P_{k|k} + S_k [P_{k+1|N} - P_{k+1|k}] S_k^T$$
(3.52)

It is also computed backwards in time, with $P_{N|N}$ as initial condition.



Fig. 3.2 The overview of the algorithm.



Fig. 3.3 Flowchart of function that calculates the trajectory.



Fig. 3.4 Flowchart of function that calculates a single point expansion.

Chapter 4

Bayesian Statistical Inference Approach to Trajectory Estimation

This Chapter presents a novel batch stochastic optimization algorithm for trajectory estimation which does not assume knowledge about the motion model, but only the standard maneuvering ship behavior. The algorithm performs curve fitting to the measurements with heteroscedastic and correlated measurement errors using Bayesian statistical inference for nonparametric regression. Specifically, the trajectory is considered to be in the form of a cubic spline with an unknown number of knots in two-dimensional Euclidean plane of longitude and latitude. The function estimate is determined from the data which are assumed Gaussian distributed. A fully Bayesian approach is adopted by defining the prior distributions on all unknown parameters: the spline coefficients, the number and the locations of knots. The calculation of the posterior distributions is performed using Markov Chain Monte Carlo (MCMC) and reversible jump Markov sampling due to the varying dimensions of subspaces where the searches are performed. Previous work on Bayesian inference to spline fitting using nonparametric regression model found in Di Matteo et al. [1]. However, the work of DiMatteo et al. fails to accommodate for heteroscedastic and correlated errors which are inherent to the problem of trajectory estimation in dissimilar multisensor environment such as maritime surveillance. Also, it fails to accommodate for decreasing values of predictor variables which occurs in backwards turning maneuvers. In addition here, Bayesian inference is performed for two nonparametric regression models with different choices of response and predictor variables: time t versus position in x or y direction, and position in x versus position in y. For the t, x and t, y combinations of regression variables, the error projections in respective directions are used for calculating the known errors' ϵ_x and ϵ_y variances.

4.1 Method Development

Assume any of the nonparametric regression models from Section 2.4. Depending on the choice on the response and predictor variables, the goal is to obtain the curve estimate $\hat{f}(x)$ (or $\hat{f}_x(t)$ and $\hat{f}_y(t)$) using Bayesian inference as

$$\hat{f}(x) = E\{f(x)|y\} = E\{E\{f(x|y,\xi,l\}\} \approx B_{l,\xi}\hat{\beta}$$
(4.1)

where $\hat{f}(x)$ is the pointwise posterior mean, l is the unknown number of knots, ξ is the set of unknown knot locations, and $\hat{\beta}$ is the set of spline parameters to be estimated. For each knot configuration pair, (l,ξ) , there is a corresponding set of spline coefficients, $\beta = (\beta_1, \ldots, \beta_{k+2})$. Therefore, the inference about the unknown parameters has to be done one at a time, i.e. the inference on β is made after inferencing on ξ and k. In the setting of Bayesian analysis, this requires calculating the marginal posterior distribution of each parameter of interest β, ξ and l from the joint posterior distribution $p(\beta, \xi, l|y)$ which models all the unknown parameters.

4.1.1 Bayesian Model

The joint posterior probability distribution $p(\beta, \xi, l|y)$ for the set of unknown parameters ξ, l, β can be found as, [28],

$$p(\beta,\xi,l|y) \propto p(y|\beta,\xi,l)p(\beta,\xi,l)$$
(4.2)

As shown in [28], the joint posterior density (4.2) can be factored to yield the marginal posterior distribution of the unknown spline parameters β , $p(\beta|y)$, defined as

$$p(\beta|y) = \int p(\beta|\xi, l, y) p(\xi, l|y) d\xi dl$$
(4.3)

The expression (4.3) shows that the posterior distribution $p(\beta|y)$ is the mixture of the conditional posterior distributions given (ξ, l) and y, where $p(\xi, l|y)$ is a weighting function for the different possible values of (ξ, l) . The weights denote the posterior density of (ξ, l) , hence are combinations of evidence from data and the prior model.

In general, the integral in (4.3) is computed by both marginal and conditional simulation, i.e. by first drawing (ξ, l) from its marginal posterior $p(\xi, l|y)$ and then β from its conditional posterior distribution $p(\beta|\xi, l, y)$, given the drawn value (ξ, l) . Following [1], the prior distribution on β , $\pi(\beta|\xi, l)$, is chosen as normal so that $p(\beta|\xi, l, y)$ can be solved analytically given (ξ, l) , i.e. the prior distribution on β , $\pi(\beta|\xi, l)$, is chosen as

$$\pi(\beta|\xi,l) = N_{k+1}(0,\sigma^2 n(B'_{\xi,l}B_{\xi,l})^{-1})$$

In [1], the prior distribution $\pi(\beta|\xi, l)$ is referred to as the conjugate Normal prior on β , while in [96] as the unit-information prior because the amount of information in the prior, represented in the covariance matrix, is equal to the amount of information in one observation, as represented by Fisher information matrix.

For a full Bayesian formulation the prior distributions on the number of knots l and the knot locations ξ are chosen as uniform as in [1] i.e.

$$\xi | l \sim U[\xi_1, \xi_{l_0}] \tag{4.4}$$

$$l \sim U[1, l_0] \tag{4.5}$$

where l_0 is the initial known number of knots, and $\{\xi_1, \ldots, \xi_{l_0}\}$ are the initial knot locations. The uniform distribution on ξ is induced by the uniform prior over the standard k-simplex by rescaling ξ to [a, b].

4.1.2 Bayesian Simulation: Reversible-Jump Metropolis-Hastings Markov Chain MC

Drawing (ξ, l) from its marginal posterior $p(\xi, l|y) \propto p(y|\xi, l)p(\xi, l)$ is done by Markov chain Monte Carlo simulations on the knot set (ξ, l) . However, MCMC model search is required over the collection of spaces of variable dimension, hence a *trans-dimensional* Markov chain simulation for variable dimension model selection must be used. One such scheme is the reversible jump Markov Chain MC [33] in which the Markov chain moves among candidates of models.

Reversible-jump Metropolis-Hastings Markov Chain MC generates a Markov chain (with the marginal posterior on (ξ, l) as the stationary distribution), that can "jump" between models with parameter spaces of different dimensions, while retaining aperiodicity, irreducibility, and detailed balance conditions necessary for MCMC convergence.

The moves are accepted with Metropolis-Hastings acceptance probability ρ , [31], [97],

defined as

$$\rho = \min\{1, \text{likelihood ratio} \times \text{prior ratio} \times \text{proposal ratio}\}$$
$$= \min\left\{1, \frac{p(y|\xi^c, l^c)}{p(y|\xi, l)} \cdot \frac{\pi_{\xi, l}(\xi^c, l^c)}{\pi_{\xi, l}(\xi, l)} \cdot \frac{q(\xi, l|\xi^c, l^c)}{q(\xi^c, l^c|\xi, l)}\right\}$$
(4.6)

where (ξ, l) and (ξ^c, l^c) are the current state and the candidate state of the chain, respectively, $\pi_{\xi,l}(\xi, l) = \pi_{\xi}(\xi|l)\pi_l(l)$, and q is the proposal density.

The proposal ratio involves calculations of probabilities of the possible addition, deletion and relocation, respectively, according to [1] as:

$$b_l = c \cdot \min\{1, p(l+1)/p(l)\}$$
(4.7)

$$d_l = c \cdot \min\{1, p(l-1)/p(l)\}$$
(4.8)

$$\eta_l = 1 - b_l - d_l \tag{4.9}$$

where p(l) denotes the prior probability of having l number of knots.

The choice for the prior on β as a normal distribution, given the knot set (ξ, l) , $\pi(\beta|\xi, l)$, allows for $p(y|\xi, l)$ defined as

$$p(y|\xi,l) = \int p(y|\beta,\xi,l)\pi(\beta|\xi,l)d\beta_{\xi}.$$
(4.10)

to be solved analytically. This further facilitates computation of the likelihood ratio $p(y|\xi^c, k^c)/p(y|\xi, l)$ used in reversible-jump algorithm to determine whether or not to move, e.g. in case of addition:

$$\frac{p(y|\xi^c, l^c)}{p(y|\xi, l)} = \frac{1}{\sqrt{n+1}} \left(\frac{y^T \{I_n - n(n+1)^{-1} B_{l,\xi} (B_{\xi,l}^T B_{l,\xi})^{-1} B_{l,\xi}^T \} y}{y^T \{I_n - n(n+1)^{-1} B_{l,\xi^c} (B_{l,\xi^c}^T B_{l,\xi^c})^{-1} B_{l,\xi^c}^T \} y} y^T \right)^{n/2}$$
(4.11)

Finally, the conditional posterior expectation of the unknown function f, $E\{f|\xi, l, y\} = \hat{f}$, can be evaluated for given values of x_i , i = 1, ..., n and given $\hat{\beta}$. The conditional posterior expectation for $f(x_i)$, for any x_i , i = 1, ..., n, can be obtained by averaging (4.1) over (ξ, k) samples as

$$\hat{f} = E\{f|\xi, l, y\} = \frac{n}{n+1} B_{l,\xi} (B_{l,\xi}^T B_{l,\xi}) B_{l,\xi}^T y \approx B_{l,\xi} \hat{\beta}$$
(4.12)

For the nonparametric regression model (2.7), in which the predictor variable is the time t, the respective obtained functions are \hat{f}_x and \hat{f}_y . The function \hat{f} which describes the trajectory in the x - y plane is reconstructed by combining \hat{f}_x and \hat{f}_y values obtained at the identical time instants t.

4.1.3 The Bayesian Spline Regression Algorithm

Algorithm 2 Bayesian Regression Algorithm
1: Input data, initialize, set the number of MCMC iterations
2: Declare initial knot set ξ
3: for $j = 1 : N_b, N_b =$ number of burn-in iterations: do
4: {Knot step which produces $\xi^{(j)}$ }
5: if addition-birth step then
6: do addition
7: else if deletion-death step then
8: do deletion
9: else if relocation step then
10: do relocation
11: else
12: Metropolis-Hastings step (4.6)
13: end if
14: Integrating the posterior marginal density $p(y \xi, l)$, where $y = (y_1, y_2, \dots, y_n)$.
15: Generate $\beta_{\xi}^{(j)}$
16: Obtain fitted values as $f^{(j)}(t) = \sum b_{\xi,l}(t)\beta_{\xi,h}^{(j)}$
17: Obtain f_{max}
18: end for

Steps 3-15 of the algorithm 2 produces $\xi^{(j)}$. In these steps, the Metropolis-Hastings ratio (4.6) is evaluated as the proposal to add, delete or relocate the knots. The procedures of addition, deletion and relocation are performed as follows.

Algorithm 3 addition-birth step. If in model M_l and proposing jump to model M_{l+1} .

- 1: Choose one knot, ξ_{j^*} uniformly from the set of existing knots $\{\xi_1, \ldots, \xi_l\}$
- 2: Do sampling ξ_{cand} from a proposal distribution h_B .
- 3: Generate the candidate new knot, ξ_{cand} , centered at ξ_{j^*} , with the known spread parameter τ_B and having density $h_B(\xi_{cand}|\xi,\tau_B)$.
- 4: Calculate the probability of a jump from the model M_k to the model M_{k+1} as

$$q(M_{l+1}|M_l) = b_l \frac{1}{l} \sum_i h_B(\xi_{cand}|\xi_i, \tau_B)$$

where b_l represents the probability of a new knot being added (prior birth probability).

$$b_l = c \cdot min\{1, p(l+1)/p(l)\}$$

and p(l) is the prior probability of having k number of knots, i.e the prior probability on model l.

Algorithm 4 deletion-death step. If in model M_l and proposing a jump to model M_{l-1} .

- 1: Choose a knot to be removed uniformly from the set of existing knots $\{\xi_1, \ldots, \xi_l\}$.
- 2: Calculate the probability of a jump from the model M_l to the model M_{l-1} as

$$q(M_{l-1}|M_l) = d_l \frac{1}{l}$$

where the death-probability, d_l is defined as

$$d_l = c \cdot min\{1, p(l-1)/p(l)\}$$

Algorithm 5 relocation step.

- 1: Choose one knot, ξ_{j^*} uniformly from the set of existing knots $\{\xi_1, \ldots, \xi_l\}$. Now the current sequence of knots is $\xi = (\xi_1, \ldots, \xi_{j^*-1}, \xi_{j^*}, \xi_{j^*+1}, \ldots, \xi_l)$.
- 2: Generate the candidate new knot location, ξ_{cand} , for the knot ξ_{j^*} , centered at ξ_{j^*} , with the known spread parameter τ_R and having density $h_B(\xi^c|\xi,\tau_R)$. The candidate new sequence of knots is $\xi = (\xi_1, \ldots, \xi_{j^*-1}, \xi_{cand}, \xi_{j^*+1}, \ldots, \xi_l)$. The candidate new knot location does not have to be the j^* -th element.
- 3: Choose a knot to be removed uniformly from the set of existing knots $\{\xi_1, \ldots, \xi_l\}$.
- 4: Calculate the probability of a jump from the model M_l to the model M_{l-1} as

$$q(M_{l-1}|M_l) = d_l \frac{1}{l}$$

where the death-probability, d_l is defined as

$$d_l = c \cdot min\{1, p(l-1)/p(l)\}$$

5: Compute the probability of jumping from the current model to the candidate model as

$$q(M_{cand}|M_{curr}) = \eta_l \frac{1}{l} h_R(\xi_{cand}|\xi_{j^*})$$
(4.13)

where the probability of relocation step, η_l , is defined as

$$\eta_l = 1 - l_l - d_l$$

4.1.4 Convergence of the Algorithm

The convergence of the algorithm depends on the balance equations for the reversible jump Markov chain on the knot set (ξ, l) . The balance is not influenced by heteroscedastic nature of the measurement error because the error covariance σ is known, therefore the proof of the validity of the balance equations is essentially the same as the one found in [1]. The complete proof as it applies to our case is stated here for completeness.

The goal is to show that the following holds:

$$\pi(M_l)P(M_{l-1}|M_l) = \pi(M_{l-1})P(M_l|M_{l-1})$$
(4.14)

where M_l denotes the model with l knots, $M_l = \{l, \xi_1, \ldots, \xi_l\}, l = 1, 2, \ldots, \xi_i \in (0, 1)$, and $\pi(M_l)$ represents the target distribution from which the samples are drawn. $\pi(M_l)$ is also the posterior distribution of M_l , and it is defined as

$$\pi(M_l) = \frac{p(y|\xi_1, \dots, \xi_l)p(\xi_1, \dots, \xi_l, l)p(l)}{p(y)}$$
(4.15)

 $P(M_{l-1}|M_l)$ and $P(M_l|M_{l-1})$ are the probabilities of transition from the model M_l to model M_{l-1} and from M_{l-1} to M_l , respectively, the so called Markov transition kernel.

$$M_{l} = \{l, \xi_{1}, \dots, \xi_{j^{*}-1}, \xi_{j^{*}}, \xi_{j^{*}+1}, \dots, \xi_{l}\}$$

$$(4.16)$$

$$M_{l-1} = \{l - 1, \xi_1, \dots, \xi_{j^*-1}, \xi_{j^*+1}, \dots, \xi_l\}$$
(4.17)

Then the transition probabilities $P(M_l|M_{l-1})$ and $P(M_{l-1}|M_l)$ can be now written as fol-

lows:

$$P(M_l|M_{l-1}) = P(l|l-1) \cdot P(\text{add } \xi_{j^*}|l-1) \cdot \{\text{acceptance probability}\}$$
$$= b_{l-1} \frac{1}{l-1} \sum_i h_B(\xi_{j^*}|\xi_i) \min\{1, B\}$$
$$P(M_{l-1}|M_l) = P(l-1|l) \cdot P(\text{delete } \xi_{j^*}|l) \cdot \{\text{acceptance probability}\}$$
$$= d_l \frac{1}{l} \min\{1, D\}$$

where

$$B = \frac{\pi(M_{l-1})}{\pi(M_l)} \frac{b_{l-1}\frac{1}{l-1}\sum_i h_B(\xi_{j^*}|\xi_i)}{d_l \frac{1}{l}}$$
$$D = \frac{\pi(M_l)}{\pi(M_{l-1})} \frac{d_l \frac{1}{l}}{b_{l-1}\frac{1}{l-1}\sum_i h_B(\xi_{j^*}|\xi_i)} = \frac{1}{B}$$

If B < 1 then $\alpha_d = B$ and $\alpha_b = 1$. In this case the left-hand side of the Equation (4.14) can be rewritten as

$$\pi(M_l)P(M_{l-1}|M_l) = \pi(M_l)d_l\frac{1}{l}B$$

= $\pi(M_l)d_l\frac{1}{l} = \frac{\pi(M_{l-1})}{\pi(M_l)}\frac{b_{l-1}\frac{1}{l-1}\sum_i h_B(\xi_{j^*}|\xi_i)}{d_l\frac{1}{l}}$
= $\pi(M_{l-1})b_{l-1}\frac{1}{l-1}\sum_i h_B(\xi_{j^*}|\xi_i)$
= $\pi(M_{l-1})P(M_l|M_{l-1})$

Also for the completeness, the standard Metropolis-Hastings step is presented below.

4.1.5 Metropolis-Hastings

The Metropolis-Hastings algorithm associated with the *target* density f and the conditional density q produces a Markov chain $(X^{(t)})$ through the following transition: Given $x^{(t)}$,

• Generate $Y_t \sim q(y|x^{(t)})$.

• Take

$$X^{(t+1)} = \begin{cases} Y_t \text{ with probability } \rho(x^{(t)}, Y_t) \\ x^{(t)} \text{ with probability } 1 - \rho(x^{(t)}, Y_t) \end{cases}$$
(4.18)

where

$$\rho(x,y) = \min\left\{\frac{f(y)}{f(x)}\frac{q(x|y)}{q(y|x)}, 1\right\}$$

The distribution q is called the *proposal* distribution and the probability $\rho(x, y)$ the *Metropolis-Hastings acceptance probability*.

4.2 Maneuvers in x - y Plane When x is Decreasing

Trajectory estimation during backward turning maneuvers of ships in the x - y plane involves descending values in the x-coordinate such as illustrated in Fig. 4.1, and hence cannot be directly accommodated for in standard regression procedures. Since the data arrives in batches a direct and simplest remedy for this situation is to transform the data received. One such transformation is a rotation of the entire data set with respect to the origin in such a way as to insure strictly increasing ordering in the values of the xcomponents of the measurements received. Needless to say, such a deterministic rotation of data does not affect any of the statistical characteristics of the data and is essentially performed as a rotation of a "rigid body composed of all measurement points together with their associated areas of uncertainty".



Fig. 4.1 Turning maneuver.

Since all the measurements with associated time tags are available in a batch, a temporal alignment allows to discern when the position in x begins to decrease. Using any of the measurements with a decreased value in x it is possible to determine an angle of rotation of the coordinate system needed to secure a strictly increasing ordering of the x-components of the measurement points. The angle of rotation can be obtained by vector algebra, see Fig. 4.2, by first computing the vector \vec{v}

$$\vec{v} = (\vec{a} + \vec{b})/2 - \vec{c}$$



Fig. 4.2 The rotation of data.

Assuming that \vec{v} is normalized to $||\vec{v}|| = 1$ and that \vec{w} , with $||\vec{w}|| = 1$ represents the unit basis vector of the w axis, perpendicular to \vec{v} , the desired angle of rotation is then determined from the vector product $\langle \vec{x}, \vec{w} \rangle = \cos \alpha$, where $||\vec{x}||$ is the unit basis vector of the axis x.

Chapter 5

A Genetic Algorithm for Trajectory Estimation

To facilitate the interpretation of the simulation results, the concept of a batch optimization procedure, which employs a genetic algorithm (GA) and the least square (LS) criterion, is presented in this Chapter. The GA based method is conceptually similar to the one described in [2]. However, the method described here uses a basic GA in the optimization procedure instead of a hybrid one. Also, it does not account for the non-sensor information, i.e. it does not use track templates.

5.1 Genetic Algorithms

Genetic algorithms are stochastic algorithms for global optimization inspired by the mechanisms of natural selection and genetics [98]. When searching the space of optimal candidate solutions with a GA, only a small fraction of candidates (i.e. chromosomes) need to be examined. The likely optimal candidates are obtained using the basic GA operators such as selection, crossover and mutation [98]. The chromosomes in a GA population typically take the form of binary bit strings. Each chromosome can be thought of as a point in the search space of candidate solutions. The GA processes populations of chromosomes, successfully replacing one such population with another. A fitness function assigns a score (fitness) to each chromosome in the current population. The fitness of a chromosome depends on how well the chromosome solves the problem at hand.

5.2 Trajectory Estimation Using a GA

With the goal of developing a trajectory estimation algorithm using a GA, the trajectory representation is considered the same as the one presented in [2]. A trajectory is represented by a collection of parametrized line segments, as ships usually navigate between waypoints along straight lines. Different states along a trajectory are represented by chromosomes made up from contacts ordered in time. Each contact represents a gene in a track chromosome. A gene may have two possible values, 0 or 1. The value 1 indicates the beginning of a track and the beginning of a line segment in the track. Each track is assumed to contain at least one segment. Each segment is assumed to contain at least four contacts. Figure 5.1 shows a possible chromosome representation for a twenty-contact trajectory consisting of three segments.



Fig. 5.1 Chromosome representation of a three-segment trajectory.

The initial population is randomly generated. To improve the search space, two additional track chromosomes are generated and added to the initial population. The first one contains only one segment while the other is obtained by calculating and comparing the sequential χ^2 best fit of contacts ordered in time. If the new χ^2 fit (i.e. one that includes the new contact in the sequence) differs in excess of 20% from the previous χ^2 fit, then a new segment is created.

The fitness function for each trajectory chromosome is calculated while assuming that the probability distribution for the contact report data is Gaussian, which allows for a least-square fit of the contacts [99]. Assuming that the ground-truth trajectory model is a collection of parameterized line segments, with parameters α_q and β_q , the least-square fit of the contact report data for each trajectory is obtained as

$$FIT_{ls} = \sum_{s=1}^{n_s} \left(\frac{\chi^2}{\nu}\right)_s \tag{5.1}$$

where s is the segment index, n_s is the total number of segments, and

$$\frac{\chi^2}{\nu} = \frac{1}{2n-4} \left\{ \sum_{q=x,y} \sum_{i}^{n} \left(\frac{q_i - (\alpha_q + \beta_q t_i)}{\sigma_{q_i}} \right)^2 \right\}$$
(5.2)

where *n* is the number of contacts in a segment, *q* characterizes the 2-D measured, geographical position of the contact (*x*-axis for longitude, *y*-axis for latitude), t_i is the time tag for the *i*th contact, and α_q and β_q are the parameters that define the line segment in the slope-intercept form. It is the α_q and β_q that are sought by way of global optimization. The values σ_{q_i} represent the maximum measurement error projections of the ellipse contact report *i* in the x - y plane, see Fig. (2.1), i.e.,

$$\sigma_{x_i} = \max\{a_i \sin \theta_i, \, b_i \cos \theta_i\} \tag{5.3}$$

$$\sigma_{y_i} = \max\{a_i \cos \theta_i, \, b_i \sin \theta_i\} \tag{5.4}$$

where a_i and b_i are the lengths of the semi-major and semi-minor axes of the ellipse, respectively, and θ_i is the ellipse orientation angle for the *i*-th report.

As in [2], the GA algorithm is exited when the state with the lowest cost remains the same for ten generations in a row. Then, the parameters α_q and β_q can be obtained from the "fittest" chromosome, i.e. from the one with the minimal cost, as

$$\alpha_q = \frac{\sum_i^n \frac{t_i^2}{\sigma_i^2} \sum_i^n \frac{q_i}{\sigma_i^2} - \sum_i^n \frac{t_i}{\sigma_i^2} \sum_i^n \frac{q_i t_i}{\sigma_i^2}}{\sum_i^n \frac{t_i^2}{\sigma_i^2} \sum_i^n \frac{1}{\sigma_i^2} - \left(\sum_i^n \frac{t_i}{\sigma_i^2}\right)^2}$$
(5.5)

$$\beta_q = \frac{\sum_{i}^{n} \frac{q_i}{\sigma_i^2} - \alpha_q \sum_{i}^{n} \frac{1}{\sigma_i^2}}{\sum_{i}^{n} \frac{t_i}{\sigma_i^2}}$$
(5.6)

In (5.5) and (5.6), the term σ_{q_i} is replaced by σ_i for simplicity. The maximum measurement error projections, σ_{q_i} , are the approximations of the real measurement error values used to simplify the calculations and decouple the errors in (5.5) and (5.6).

Chapter 6

Summary of Simulation Results for Trajectory Estimation Algorithms

This Chapter provides the results of the performance evaluation of the presented trajectory algorithms in the statistical sense. It includes the sensitivity analysis of the algorithms' parameters with respect to perturbations in parameters, and the discussion of the obtained results.

6.1 Scenario

For the purpose of evaluation of the performance of the ship trajectory estimation algorithms the SimTrack application [100] is developed to simulate the ground truth and the multi-sensor environment. The ground truth trajectories are generated so as to represent actual ships' routes. The scenario produces random ship trajectories of random lengths (in terms of time) in a 200NM×200NM surveillance area, with the length of a trajectory being uniformly distributed on the interval of [6, 12] hours. For the Monte Carlo (MC) simulations, the default value for the maximum number of segments, i.e. "legs", is assumed to be three. This value can be changed. The time-on-leg (TOL) is considered uniformly distributed on the interval of [2, 6] hours. The default constant speed between waypoints is assumed to be V = 12 knots. This value can be changed as well.

The application includes a "generic" sensor that produces a positional ellipse contact report at a random sampling rate; the rate is chosen randomly from the uniform distribution on [5, 60] minutes interval. The measurement is the center of the error ellipse, with the lengths of semi-major and the semi-minor axes being uniformly distributed on [1, 2] nm and [3, 5] nm intervals, respectively. The measurement ellipse has a bearing θ , which is measured in degrees clockwise from the true North of the semi-major axis, and is randomly chosen from uniform distribution on the interval [0°, 360°]. This type of measurement and the measurement error are chosen to adequately depict the multisensor environment.



Fig. 6.1 Scenario specifications.

6.2 Choice of Performance Measure

Estimated trajectories using all presented algorithms are data dependent, hence their performance is random and is evaluated in a statistical sense using Monte Carlo simulations. The selection and proper interpretation of the metrics used for measuring the performance and determining the characteristics of the each estimation algorithm is of great importance for the evaluation.

In this thesis the usual measure of estimation error, the root mean squares error (RMSE), is replaced by the average Euclidean error (AEE) as suggested in [101]. Having clear physical interpretation and paying less attention to large errors than the RMSE, it proves to be better suited for the positional estimation than the RMSE [101]. The AEE is defined as

$$AEE(\hat{x}) = \frac{1}{M} \sum_{i=1}^{M} \|\tilde{x}_i\|$$

where estimate (i.e. the ground truth position), estimate, and estimation error are denoted by x, \hat{x} , and $\tilde{x} = x - \hat{x}$, respectively. The term Euclidean error stems from the concept of Euclidean distance or Euclidean norm. Subscript *i* stands for quantities pertaining to the *i*-th run of a Monte Carlo simulation consisting of a total of M independent runs, and thus \tilde{x}_i and \tilde{x}_j are independent for $i \neq j$.

In the next Section, the results pertaining to the IOU-KF-RTSS algorithm are presented.

6.3 IOU-KF-RTSS Algorithm Performance Evaluation

The default IOU model parameters were chosen so as to capture ships' average dynamics, i.e. the average speed, V = 12 knots and the average time on leg TOL = 4 hours. The parameters are consequently obtained from (3.34) and (3.35) as $\beta = 1/\text{TOL} = 0.25$ and $\sigma = \sqrt{\beta}V = 6$.

To evaluate trajectory estimation performance of the IOU-KF-RTSS algorithm several scenarios were considered and the corresponding simulation results were obtained. First, for a single ground truth 3-segment trajectory, the average Euclidean errors (AEEs) and the associated standard deviations (STDs) were obtained for the track models produced by the tracker. MC simulations included 100 measurement noise realizations (i.e. different ellipse locations, the same ground truth and the same sizes and orientations of ellipses). The AEEs and the associated STDs are shown in Figures 6.2. It can be seen that the tracker performs well with an average error value of 0.7 NM, a value that is smaller than the minimum magnitude of semi-minor axis of the area of uncertainty, i.e. 1 NM. The first peak in the error plot is due to the track initiation process.



Fig. 6.2 IOU-KF-RTSS: The AEEs and the STDs of produced tracks for a 3-segment trajectory.

The analysis was extended to 1000 different ground truth trajectories which were gen-

erated by the SimTrack. For each ground truth track it was assumed that sensors report different number of measurements at different random rates. To be able to evaluate simultaneously modeling of 1000 ground truth trajectories of different lengths with different numbers of contacts per trajectory, the AEEs were averaged along each trajectory with respect to the number of contacts per trajectory. The AEEs along each trajectory were calculated again for 100 measurement noise realizations, at the times of contacts. The histograms in Figure 6.3 present the number of trajectories with corresponding errors and standard deviations for 1000 produced trajectories. The error spread is consistent across tracks, i.e. 99% of errors have the values between 0.5 - 1.0 NM.



Fig. 6.3 IOU-KF-RTSS: Averaged AEEs and STDs for estimated trajectories of 1000 ground truth tracks of different lengths and different number of contacts per track.

Furthermore, trajectory estimation performance was investigated in the vicinity of a turning maneuver. For this purpose, 1000 two-segment ground truth trajectories were generated and the AEEs and the STDs of the estimated trajectories in the vicinity of the "leg" change point were calculated. The calculations included at least one measurement before and after the change point. The evaluation time interval was then chosen to be 15 minutes so as to capture the complete turning maneuver of a ship moving at the average speed of 12 knots. The errors and the average standard deviations per track are presented in Figure 6.4. The tracker models the transition from one leg to another rather accurately



Fig. 6.4 IOU: Averaged AEEs and STDs of 1000 ground truth tracks in the vicinity of turning maneuver.

as most of the errors have approximately the value between 0.5-2.0 NM with the associated average standard deviation of 0.02 - 0.04 NM.

6.4 Sensitivity Analysis in the IOU Parameters

The sensitivity analysis is performed by two different simulations which both included 50 ground truth trajectories and 100 measurement noise realizations. First, for a ship moving through water with constant speed V = 12 knots, the σ parameter was varied between [6, 20] while the parameter β was varied between [0.25, 1]. The obtained average Euclidean error values are shown in Fig. 6.5. The varying intervals correspond to tailoring the algorithm to a ship moving with speeds between [12, 20] knots as depicted in Fig. 6.6. The AAE



Fig. 6.5 IOU: AEEs for varying IOU parameters σ and β .

values of approximately 0.6 NM indicate no deterioration of performance, see Fig.6.3.

The second approach to sensitivity analysis of the proposed algorithm was to investigate the its performance for the default values of the IOU process parameters, (i.e. $\beta = 0.25$ and



Fig. 6.6 IOU-KF-RTSS: AEEs for varying IOU parameters σ and β in terms of *TOL* and *V*.

 $\sigma = 6$ and corresponding to V = 12 knots), while tracking a ship which travels with speed V = 20 knots, the average speed of a cruising ship. The AEE plots in Fig. 6.7 and Fig. 6.8, respectively, indicate slightly deteriorated performance with respect to the situation when tracking ships travelling at V = 12kn. Nevertheless, the AEEs have values of $\propto 1$ NM which is very small. The choice of β controls the rate of the speed change. Therefore, a skilled choice of β is necessary if a better trajectory estimation accuracy is required.



Fig. 6.7 IOU-KF-RTSS: AEEs for tracking ships traveling with speed, V = 20knots, for $\beta = 0.25$ and $\sigma = 6$.



Fig. 6.8 IOU-KF-RTSS: STDs for tracking ships traveling with speed, V = 20knots, for $\beta = 0.25$ and $\sigma = 6$.

6.5 IOU-KF-RTSS Algorithm Comparison with a Genetic Algorithm (GA) for Trajectory Estimation

A comparison in trajectory estimation performance between IOU-KF-RTSS and GA-based algorithms was performed for several scenarios for which the corresponding simulation results are obtained. The first set of results pertains to a single ground truth 3-segment trajectory. The average Euclidean errors (AEEs) and the associated standard deviations (STDs) are obtained for the estimated trajectories using both algorithms. The MC simula-



Fig. 6.9 IOU-KF-RTSS vs. GA: The difference of the AEEs and the STDs of produced trajectories for a 3-segment ground truth trajectory.

tions again included 100 measurement noise realizations. The AEE and the associated STD are shown in Figures 6.9, and suggests a slight superiority of the IOU based method over the GA based one, which is manifested by positive difference between the AEEs of both algorithms. The exception is at the beginning of the track may be attributed to the track
initiation process in the IOU based method. The peaks in the error plot coincide with the change points along the trajectory (i.e. the beginning of different legs of the trajectory).

Next, the analysis is extended to 1000 different ground truth trajectories generated by SimTrack. Again, for each ground truth track it is assumed that sensors report different number of measurements at different random rates. As in Section 6.3, to be able to simultaneously evaluate the estimation of 1000 trajectories of different lengths with different numbers of contacts per track, the AEEs were averaged along each trajectory with respect to the number of contacts received for this trajectory. The AEEs along each trajectory were calculated again for 100 measurement noise realizations at the times of contacts. The his-



Fig. 6.10 IOU-KF-RTSS vs. GA: Averaged AEEs for 1000 ground truth tracks of different lengths and different number of contacts per track.

tograms in Fig. 6.10 and Fig. 6.11 present number of trajectories with corresponding error and standard deviation values for each algorithm. The error spread for the IOU based algorithm is more consistent across tracks. Also, the errors and the standard deviations



Fig. 6.11 IOU-KF-RTSS vs. GA: Averaged STDs for 1000 ground truth tracks of different lengths and different number of contacts per trajectory.

have smaller values for the IOU based algorithm, i.e. $AEE \approx 0.5$ NM.

The same conditions as in Section 6.3 apply in the comparison of two algorithms' performance in the vicinity of a turning maneuver. The number of trajectories and corresponding average errors and the average standard deviations per track are presented in Fig. 6.12 and Fig. 6.13, respectively. The IOU-based method estimates the transition from one leg to another rather accurately as most of the errors have approximately the value between 0.5 - 2.0 NM with the associated average standard deviation of 0.02 - 0.04 NM. These results are better than the corresponding ones for the GA based algorithm.



Fig. 6.12 IOU-KF-RTSS vs. GA: Averaged AEEs of a 1000 ground truth tracks in the vicinity of turning maneuver.



Fig. 6.13 IOU-KF-RTSS vs. GA: Averaged STDs for a 1000 ground truth tracks in the vicinity of turning maneuver.

6.6 Performance Evaluation of the Bayesian Regression Spline Algorithm

Table 6.1 shows the default prior parameters values for the Bayesian regression spline algorithm:

parameter	value
number of burn-in iterations	100 or 0
number of MCMC iterations	300
initial number of knots	3
τ_R parameter for knot proposal distribution	50
reversible jump constant c	0.4
number of grid points	500

 Table 6.1
 Prior parameters for the Bayesian regression spline algorithm.

Similar to the IOU-KF-RTSS algorithm performance evaluation, each simulation for the Bayesian regression spline algorithm included 100 measurement noise realizations. However, when evaluating multiple trajectories of different number of segments and different number contact per trajectory, the total number of trajectories for all simulations had to be reduced from 1000 to 50 due to the computational requirements of the Bayesian spline regression algorithm.

6.6.1 Regression on x - y

A single 3-segment ground-truth trajectory from which 14 contacts were obtained by the sensors was considered in the first set of simulation experiments. The results obtained are compared with those obtained in the same simulation for the IOU-KF-RTSS algorithm and the GA-based trajectory estimation. The AEEs and the STDs for all three algorithms are summarized in Figures 6.14 and 6.15. In the Bayesian regression spline algorithm, the regression was performed in x-y plane, hence the complete information about measurement

uncertainty was used. The total errors across tracks show better trajectory estimation performance of both proposed algorithms, the BRS(x,y) and the IOU-KF-RTSS algorithm, than of the GA-based one. Although the difference in total error appears rather small, a better performance of the trajectory estimation algorithm can be cruicial in surveillance activities such as search and rescue.



Fig. 6.14 BRS, (x,y): AEE for a single 3-segment trajectory with 14 contacts versus IOU-KF-RTSS and GA.

To investigate the overall performance of the Bayesian regression spline algorithm, a simulation for 50 trajectories of different lengths, 3 - 7 segments and different number of contacts per trajectory was performed next.

The histograms in Fig. 6.16 and 6.17 show the number of trajectories with corresponding AEEs and STDs obtained by this method. Most of the AEE values are close to 1 NM, which is similar to the performance of the IOU-KF-RTSS algorithm. This was expected since the regression model in the algorithm completely accounted for the measurement uncertainty.



Fig. 6.15 BRS, (x,y): STD for a single 3-segment trajectory with 14 contacts versus IOU-KF-RTSS and GA.



Fig. 6.16 BRS, (x,y): Averaged AEEs for a 50 ground truth tracks.



Fig. 6.17 BRS, (x,y): Averaged STDs for a 50 ground truth tracks.

6.6.2 Regression on t - x and t - y

The regression in Bayesian regression spline algorithm was carried out with respect to t as well. In this case, the regression model only partially accounted for measurement uncertainty by using the measurement error projections. The values with the same time tag of two resulting splines, \hat{f}_x and \hat{f}_y , were then combined to obtain the resulting spline function \hat{f} in the x - y plane. The results of the numerical simulations in Fig. 6.18 show a performance similar to those of the GA-based trajectory estimation algorithm (AEEs $\approx 3 - 5$ NM) since both algorithms use measurement error projection in calculating the estimated trajectory. The obtained AEE values are also larger than those obtained when the regression was performed on x - y.



Fig. 6.18 BRS from (t, x) and (t, y): Averaged AEEs and STDs for 50 ground truth tracks obtained by regression with respect to t.

6.7 Sensitivity Analysis in Parameters of the Bayesian

Regression Spline Algorithm

The sensitivity to the number of contacts was examined for a 7-segment trajectory for which 14, 26, 32, or 40 contacts were obtained. The results are shown in Fig. 6.19 and Fig. 6.20 and they indicate that the Bayesian regression spline algorithm is sensitive to the number of measurements available for fitting. The values of both AEE and STD show that the performance becomes poorer as the number of contact decreases. The influence of the total number of measurements n on the accuracy of produced trajectory of the Bayesian regression spline algorithm was further investigated in comparison with the IOU-KF-RTSS and the GA-based algorithm. A 7-segment ground truth trajectory was generated from which first 14, and then 32 contacts, were obtained. The obtained



Fig. 6.19 BRS, (x,y): AEEs for 7-segment ground-truth trajectory with different number of contacts per trajectory.



Fig. 6.20 BRS, (x,y): STDs for 7-segment ground-truth trajectory with different number of contacts per trajectory.

values of the AEE shown in Fig. 6.21 suggest robustness of the IOU-KF-RTSS algorithm with respect to the total number of measurements, while both the GA-based and Bayesian regression spline algorithms exhibit larger errors. However, the total error across produced trajectory for both proposed algorithms, the BRS(x,y) and the IOU-KF-RTSS, are smaller than of the one obtained when using the GA-based. Also, it is confirmed that the accuracy of produced trajectories by the BRS(x,y) increases with the increase of the total number of measurements in a batch.



Fig. 6.21 AEEs for the BRS(x,y) vs. IOU-KF-RTSS and GA-based for a 7-segment ground truth trajectory for n = 14 and n = 32.

Finally, a sensitivity analysis was performed with respect to change in the initial number of knots, l, as a prior parameter of the Bayesian regression spline algorithm. This was investigated for ten 10-segments ground truth trajectory with 26 contact and 100 measurement noise realizations. The initial number of knots l_0 was set to be chosen randomly from a uniform distribution $l_0 \sim U[2, 10]$. The results shown in Dig. 6.22 that the Bayesian



regression spline algorithm's performance remains unchanged.

Fig. 6.22 AEEs for the BRS(x,y) algorithm for different initial number of knots, iknot = 3 and iknot = 7, for ten 10-segment ground truth trajectories.

6.8 On the Computational Note

The implementation of the IOU-KF-RTSS (both batch and recursive implementation) is much faster than the current implementation of the Bayesian regression spline algorithm. This is also due to the MCMC computations in the Bayesian regression spline algorithm. Using 1.3GHz Intel Core Duo processor with 3 GB of RAM, it takes 0.21 s for the IOU-KF-RTSS to estimate a single trajectory of a random length, while the same task takes the Bayesian regression spline 61.3 s.

Chapter 7

Concluding Remarks

7.1 General Summary

Due to large differences in their temporal resolution, multi-level security and management processing, the contact report data used in compiling the picture of activities in the maritime domain are usually available for batch processing. The data obtained from dissimilar sources have large measurement errors characterized by normal probability distributions with non-constant and non-diagonal covariance matrices (i.e. the heteroscedastic and correlated measurement errors). Therefore, the method used for processing this type of contact report data will significantly influence the accuracy of the estimated trajectories. More accurate representation of ship trajectories will support decision processes in the maritime domain, and consequently increase the maritime domain awareness.

This thesis presented two batch single ship trajectory estimation algorithms using Baysian approaches to estimation which assume no knowledge about the ship motion model, while assuming standard ship maneuvers: (1) a stochastic linear filtering algorithm which employs the Integrated Ornstein-Uhlenbeck processes in conjunction with Kalman filter-

ing and Rauch-Tung-Striebel fixed optimal smoothing, and (2) a curve fitting algorithm which employs Bayesian statistical inference to perform nonparametric regression. The description of development, the implementation and the assessment of the first presented algorithm, the linear stochastic filtering algorithm which employs the IOU process as the motion model, is first presented. In the implementation of the second algorithm, the trajectory is considered to be in the form of a cubic spline with an unknown number of knots in two-dimensional Euclidean plane of geographical coordinates: the longitude and the latitude. The function estimate is determined from the data which are assumed Gaussian distributed. A fully Bayesian approach is adopted by defining the prior distributions on all unknown parameters: the spline coefficients, the number and the locations of knots. The calculation of the posterior distributions is performed using Markov Chain Monte Carlo (MCMC) and reversible jump Markov sampling due to the varying dimensions of subspaces where the searches are performed. The algorithm developed here is inspired by the algorithm found in Di Matteo et al. [1]. However, the work of DiMatteo et al. fails to accommodate explicitly for heteroscedastic and correlated errors as well as for decreasing values of predictor variable which occurs in backwards turning maneuvers. Additionally, Bayesian inference is performed for nonparametric regression models with two different choices of response and predictor variables.

For the purpose of the evaluation of trajectory estimation algorithms an application SimTrack [100] was developed to simulate the ground truth in a dissimilar multi-sensor environment. The performance evaluation of the trajectory estimation procedures was carried out for several scenarios. It also included the sensitivity analysis with respect to perturbation in algorithm's parameters. The comparison of the novel algorithms with a GA-based algorithm for trajectory estimation was also performed.

7.2 Conclusions

The following conclusions are drawn from this study:

- In the absence of any knowledge about the motion model, the proposed batch stochastic optimization algorithm based on spline fitting proves to be an efficient method for trajectory estimation in application to ship track estimation.
- The proposed algorithm is sensitive to the total number of contacts in a batch and obviously performs better when this number is large. Bootstrap techniques could be used to improve this situation.
- When employed in the x y plane, the Bayesian regression spline algorithm performs better than the GA based algorithm because it completely accounts for the specifics of the uncertainty of measurements. It has similar performance to the GA-based algorithm when using the regression model on (t, x) and (t, y), since they both use measurement errors' projections.
- The IOU based algorithm exhibits robustness with respect to number of measurements and all types of ground truth trajectories. For improved accuracy, a skilled choice of parameters is necessary.
- The IOU-KF-RTSS algorithm has smaller execution time than the Bayesian regression algorithm in producing trajectories.
- Both proposed algorithms outperform the GA-based data fitting algorithm.
- Extensive simulation studies showed that both proposed algorithms represent competitive methodologies for trajectory estimation in maritime surveillance, especially

in the absence of knowledge of the motion models and having contacts with large heteroscedastic and correlated errors.

• In light of the above, it can be concluded that modeling and incorporating the measurement uncertainty plays an important role in the trajectory estimation algorithm by affecting the accuracy of the estimated trajectory. Any approximation of the measurement error degrades the accuracy of the produced trajectory especially in target environments where the availability of radar measurements is reduced.

Although primarily developed for maritime surveillance applications, both presented algorithms in this thesis apply to a broader class of parameter estimation algorithms such as those used in finance (e.g. estimation of integrated volatility in stochastic volatility models and biostatistics (e.g. derivative tracking such as tracking AIDS and other viral marker data). Trajectory estimation by means of regression splines in \mathbb{R}^1 may be of interest as a way to specify behavioral patterns or relationships between locations observed at adjacent times, as a record of the history of past locations, or to improve modeling in other applications than ship tracking.

7.3 Future Research Avenues

A direct extension of this research is the problem of combining ellipse contact reports which contain the same time tag. Furthermore, for slow maneuvering vessels, the Interacting Multiple IOU model would accommodate for more accurate trajectory estimation of ships in that class. The Bayesian regression spline method can be extended by using the empirical Bayesian inference in nonparametric regression modeling in which the prior parameters can be learnt from the data. The problem of tracking in maritime surveillance naturally extends to problem of multiple ship trajectory estimation, where the cross-entropy information could be used for various data association procedures and the ellipse combination. A reliable frequency determination/classification and ship identification from the ESM or ELINT measurements together with tracking is still an open problem especially in difficult maritime scenarios, such as illegal fishing or contraband activities.

Appendix A

The Relationship of the Error Ellipse Measurement to a Gaussian Bivariate Distribution

The information about the measurement *i* obtained at known time t_i is in the form of a vector $z_i = \begin{bmatrix} x_i & y_i & a_i & b_i & \theta_i \end{bmatrix}$, where $[x_iy_i]'$ represents 2-D measured geographical position (*x*-axis for longitude, *y*-axis for latitude), and a_i, b_i and θ_i are the parameters that characterize the geometry of the elliptical area of uncertainty associated with the position. The area of uncertainty (AOU) is defined as the minimum area having a specified probability of containing the measurement, see Fig. A.1.

Hence, each ellipse contact report is defined by the center position, the angle of orientation, magnitudes of the semi-axes, and the probability level. Assuming that the positional measurements are normally distributed, Gaussian statistics can be used to characterize the measurement error as well as assign the specified probability level to its area of uncertainty. The computation of the covariance matrix from the parameters that characterize the ellipse



Fig. A.1 The uncertainty ellipse with 96% of probability confinement region

geometry is performed as follows. The details can be found in [102] and [103].

Let $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}'$ be a Gaussian random vector. The joint probability distribution of \mathbf{y} is defined as

$$p(\mathbf{y}) = k \exp\{-\frac{1}{2}(\mathbf{y} - \mu)' B(\mathbf{y} - \mu)\}$$
 (A.1)

where $B \in \mathbb{R}^{2 \times 2}$, $\mu \in \mathbb{R}^2$, and k is determined by the normalization procedure of the total probability to one.

The vector $\mu = \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix}'$ represents the mean of the distribution, i.e.,

$$\mu = E\{\mathbf{y}\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{y} p(\mathbf{y}) dy_1 dy_2$$
(A.2)

The vector μ also represents the centre of symmetry of the distribution such that

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\mathbf{y} - \mu) p(\mathbf{y}) dy_1 dy_2 = 0$$
 (A.3)

The covariance matrix of the joint distribution $p(\mathbf{y})$ is defined as, [102]

$$R(\mathbf{y}) = E\{(\mathbf{y} - \mu)(\mathbf{y} - \mu)'\} = E\{(y_i - \mu_i)(y_j - \mu_j)\}$$
(A.4)

where the element $r_{ii}(y) = E\{(y_i - \mu_i)\}^2$ of the matrix $R(\mathbf{y})$ represents the variance of y_i , and the off-diagonal elements $r_{ij} = E\{(y_i - \mu_i)(y_j - \mu_j)\}$, for $i \neq j$, is the covariance of the elements y_i and y_j . Since $E\{(y_i - \mu_i)(y_j - \mu_j)\} = E\{(y_j - \mu_j)(y_i - \mu_i)\}$, the covariance matrix can be rewritten as

$$R = B^{-1} = E\left\{ \begin{bmatrix} (y_1 - \mu_1)^2 & (y_1 - \mu_1)(y_2 - \mu_2) \\ (y_2 - \mu_2)(y_1 - \mu_1) & (y_2 - \mu_2)^2 \end{bmatrix} \right\}$$
(A.5)
$$= \begin{bmatrix} E\{(y_1 - \mu_1)^2\} & E\{(y_1 - \mu_1)(y_2 - \mu_2)\} \\ E\{(y_2 - \mu_2)(y_1 - \mu_1)\} & E\{(y_2 - \mu_2)^2\} \end{bmatrix}$$
$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

where the use has been made of the symmetry property of the matrix R (i.e. $\sigma_{12} = \sigma_{21}$). The matrix B then can be obtained by inversion of R matrix as

$$B = R^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$
(A.6)

Let the reduced variables u_i be

$$u_i = \frac{y_i - \mu_i}{\sigma_i}, \qquad \qquad i = 1, 2 \tag{A.7}$$

such that $var\{u_1\} = var\{u_2\} = 1$ so that the correlation coefficient, ρ , can be written as

$$\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = cov\{u_1, u_2\} \tag{A.8}$$

The probability density for two joint variables u_1 and u_2 can now be rewritten as

$$p(u_1, u_2) = k \exp(-\frac{1}{2}u'Bu)$$
 (A.9)

where B is defined as

$$B = \frac{1}{1 - \rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$
(A.10)

The lines of constant probability density of (A.9) are determined from the following requirement

$$\frac{1}{1-\rho^2}(u_1^2+u_2^2-2u_1u_2\rho)=c \tag{A.11}$$

As shown in [104], for c = 1, (A.11) becomes,

$$\frac{(y_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{y_1 - \mu_1}{\sigma_1} \frac{y_2 - \mu_2}{\sigma_2} + \frac{(y_2 - \mu_2)^2}{\sigma_2^2} = 1 - \rho^2$$
(A.12)

A The Relationship of the Error Ellipse Measurement to a Gaussian Bivariate Distribution 107

which is equivalent to

$$\frac{1}{1-\rho^2}\frac{(y_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{y_1-\mu_1}{\sigma_1}\frac{y_2-\mu_2}{\sigma_2} + \frac{(y_2-\mu_2)^2}{\sigma_2^2} = (\frac{z_1}{\sigma_{z1}})^2 + (\frac{z_2}{\sigma_{z2}})^2 = \chi^2$$
(A.13)

where z_1 and z_2 are two independent random variables χ^2 distributed with two degrees of freedom. The equation (A.13) describes the ellipse centered at (μ_1, μ_2) . The semi-axes of the ellipse, a and b, have an angle θ with respect to y_1, y_2 axes. The orientation of the ellipse, i.e. the angle θ and the magnitudes of the semi-axes, a and b can be derived from (A.12) using the properties of conic sections:

$$\tan 2\theta = \frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \tag{A.14}$$

$$a^{2} = \frac{\sigma_{1}^{2}\sigma_{2}^{2}(1-\rho^{2})}{\sigma_{2}^{2}\cos^{2}\theta - 2\rho\sigma_{1}\sigma_{2}\sin\theta\cos\theta + \sigma_{1}^{2}\sin^{2}\theta}$$
(A.15)

$$b^{2} = \frac{\sigma_{1}^{2}\sigma_{2}^{2}(1-\rho^{2})}{\sigma_{2}^{2}\sin^{2}\theta - 2\rho\sigma_{1}\sigma_{2}\sin\theta\cos\theta + \sigma_{1}^{2}\cos^{2}\theta}$$
(A.16)

The ellipse with the magnitudes of semi-major axis, a, semi-minor axis, b, and the orientation, θ , characterize the ellipse of covariance of the bivariate Gaussian distribution.

By solving for σ_1 , σ_2 , and ρ , the covariance matrix R is obtained as

$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} a^2/2 & 0 \\ 0 & b^2/2 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$
(A.17)

The Equation (A.13) in matrix form can be rewritten as

$$(\mathbf{y} - \mu)' R^{-1} (\mathbf{y} - \mu) = \chi^2$$
 (A.18)

where **y** is a Gaussian random vector, μ is its mean and R is its covariance matrix. The chi-squared value, χ^2 , with two degrees of freedom corresponds to the probability that the observation falls within the ellipse described by (A.18).

Also, there is a relationship between the parameters of the ellipse geometry and the eigenvalues and eigenvectors of the covariance matrix R, i.e.

$$|R - \lambda I| = 0 \tag{A.19}$$

where I is the identity matrix. From (A.5) and (A.19), and taking the determinant yields

$$\lambda^2 - \lambda(\sigma_1^2 + \sigma_2^2) + \sigma_1^2 \sigma_2^2 - \sigma_{12}^2 = 0$$
(A.20)

From (A.20), the eigen values are obtained as

$$\lambda_{1,2} = \frac{\sigma_1 + \sigma_2 \pm \sqrt{(\sigma_1^2 + \sigma_2^2)^2 - 4(\sigma_1^2 \sigma_2^2 - \sigma_{12}^2)}}{2}$$
(A.21)

The magnitudes of the semi-axes are obtained as

$$a = \sqrt{\chi^2 \lambda_1} \tag{A.22}$$

$$b = \sqrt{\chi^2 \lambda_2} \tag{A.23}$$

where the two degrees of freedom χ^2 value corresponds to the specified probability P. The

corresponding χ^2 value with respect to a specified P can be found in the Table A, [102]. In this thesis the probability level of the ellipse contact report is assumed to be P = 0.96. This value corresponds to the probability area under the bivariate Gaussian distribution defined as

$$P(|\mathbf{y} - a| \le 2\sigma)$$

hence the name 2-sigma ellipse and the factor 1/2 in the calculations of the matrix R.

χ^2	P
2.30	0.683
4.61	0.90
6.17	0.954
9.21	0.99
11.9	0.9923
18.4	0.9999

Table A.1 Table of χ^2 distribution percentage points.

A.1 The Flat-Earth Approximation

Since the error ellipses are small in size as compared with the size of the surveillance region, a local flat earth approximation in the plane tangent to the measured position, (x_k, y_k) (i.e. the ellipse centroid), can be used for constructing the bivariate Gaussian distribution, $N(0, R_k)$ corresponding to each position (x_k, y_k) .

The centre of the measurement error ellipse represents the position expressed in longitude and latitude, (lon, lat). Longitude and latitude are spherical coordinates which assume that the Earth is round. If assume a spherical Earth with radius R, the two position locations in geographical coordinates (longitude and latitude) are (lon_1, lat_1) and (lon_2, lat_2) , and the straight line as the shortest distance d between two position locations in a plane, then a Pythagorean flat-Earth approximation can be used. This approximation assumes that lines of longitudes are parallel and that those of latitude are to a small extent different from great circles, and that great circles are negligibly different from straight lines [105]. However, close to the poles, the parallels of latitude are both shorter and curved. In, this case the polar coordinate flat-Earth approximation which calculates the distance d using polar coordinates and the planar law of cosines can be used, i.e.,

$$a = \pi/2 - lat_1$$

$$b = \pi/2 - lat_2$$

$$c = \sqrt{(a^2 + b^2 - 2ab\cos(lon_2 - lon_1))}$$

$$d = Rc$$

The latitudes lat_1 , lat_2 , and the intermediate result c must be expressed in radians, while the distance d is in the same units as the Earth radius R. For higher latitudes and greater distances, these formulas produce smaller errors than those obtained using just the Pythagorean Theorem.

Alternatively, with the same assumptions on the spherical Earth, the Haversine flat-Earth approximation [106] can be used, for more details, see [103].

References

- I. DiMatteo, C. R. Genovese, and R. E. Kass, "Bayesian curve fitting with free-knot splines," *Biometrika*, vol. 88, pp. 1055–1071, 2001.
- [2] E. Lefebvre and C. Helleur, "Automated association of track information from sensor sources with non-sensor information in the context of maritime surveillance," Proc. of Intl. Conference on Information Fusion, (Stockholm, Sweden), pp. 1251–1256, 2004.
- [3] M. Hadzagic and H. Michalska, "Batch stochastic optimization algorithms for trajectory estimation," *ELSEVIER Journal on Information Fusion, in review*, 2010.
- [4] M. Hadzagic and H. Michalska, "A linear stochastic process and a genetic algorithm for ship trajectory modeling," vol. 7697 of Proc. of SPIE Defence, Security and Sensing Conference on Sensor Fusion, and Target Recognition XIX, pp. 76970V1– 76970V11, 2010.
- [5] http://www.jstatsoft.org/v26/i01.
- [6] M. Davenport, *Maritime Surveillance-The Future*. Armed Forces Communications and Electronics Association (AFCEA), 2005.
- [7] É. Bossé, J. Roy, and A. Jouan, "Information fusion concepts for airborne maritime surveillance and C2 operations, TM 2004-281," tech. rep., Valcartier QUE (CAN), 2006.
- [8] D. Hall, Mathematical Techniques in Multisensor Data Fusion. Artech House, 1992.
- [9] A. N. Steinberg and C. L. Bowman, *Revisions to JDL Data Fusion Model*, ch. 2. Handbook of Multisensor Data Fusion, (D. L. Hall and J. Llinas eds.), CRC Press, Boca Raton, FL, 2001.
- [10] J. Llinas and G. Rogova, "Revisiting the JDL data fusion model II," Proc. of the Seventh International Conference on Information Fusion, (Stockholm, Sweden), 2004.

- [11] N. Kiryati and A. Bruckstein, "Heteroscedastic Hough transform (HtHT): an efficient method for robust line fitting in the errors in the variables problems," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 69–83, 2000.
- [12] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*. Academic Press, Boston, MA, 1987.
- [13] S. Blackman and R. Popoli, Design and Analysis of Modern Tracking Systems. Artech House, 1999.
- [14] L. A. Klein, Sensor and data fusion: a tool for information assessment and decision making. SPIE Publishing, 2004.
- [15] Y. Bar-Shalom and X.-R. Li, Estimation and Tracking: Principles, Techniques, and Software. Artech House, 1999.
- [16] B. Carlin and T. Louis, *Bayesian Methods for Data Analysis*. Chapman & Hall, 2009.
- [17] B. Ristic, S. Arulampalam, and N. Gordon, Beyond the Kalman Filter. Artech House, 2004.
- [18] A. Gelb, Applied Optimal Estimation. MIT Press, Cambridge, MA, 1974.
- [19] H. Sorenson, "Least square estimation: from Gauss to Kalman," IEEE Spectrum, pp. 63–68, 1970.
- [20] K. Kastella, Finite difference methods for nonlinear filtering and automatic target recognition, vol. III of Multitarget Multisensor Tracking, (Y. Bar-Shalom and W.D. Blair eds.), ch. 5. Norwood MA: Artech House, 2000.
- [21] L. Stone, A Bayesian approach to multiple target tracking, ch. 10. Handbook of Multisensor Data Fusion, (D. L. Hall and J. Llinas, eds.), CRC Press, Boca Raton, FL, 2001.
- [22] F. Daum, "Exact finite dimensional nonlinear filters," IEEE Trans. Automatic Control, vol. 31, no. 7, pp. 616–622, 1986.
- [23] J. Hammersley and K. Morton, "Poor's man Monte Carlo," Journal of Royal Statistical Society B, vol. 16, pp. 23–38, 1954.
- [24] J. E. Handschin and D. Q. Mayne, "Monte Carlo techniques to estimate the conditional expectation in multi-stage non-linear filtering," *International Journal of Control*, vol. 9, pp. 547–559, 1969.
- [25] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEE Proc.-F*, vol. 140, no. 2, pp. 107–113, 1993.

- [26] R. Eubank, Spline Smoothing and Nonparametric Regression. Marcel Dekker Inc, New York, NY, 1988.
- [27] M. Bartholomew-Biggs, Nonlinear Optimization with Financial Applications. Springer, USA, 2005.
- [28] A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*. Chapman & Hall, 2004.
- [29] L. Tierney and J. Kadane, "Accurate approximations for posterior moments and marginal densities," *Journal of American Statistical Association*, vol. 29, pp. 82–86, 1986.
- [30] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, pp. 1087–1092, 1953.
- [31] W. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.
- [32] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and Bayesian restoration of images.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1970.
- [33] P. Green, "Reversible jump Markov chain Monte Carlo computation in Bayesian model determination," *Biometrika*, vol. 82, pp. 711–732, 1995.
- [34] J. K. F. Gauss, Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium (Theorie der Bewegung der Himmelskrper, die die Sonne in Kegelschnitten umkreisen). Dover, New York, reprinted 1963 ed., 1809.
- [35] A. Legendre, Nouvelles méthodes pour la détermination des orbites des cometes (1805). Sur la Méthode des moindres quarrés appears as an appendix. 1806.
- [36] R. A. Fisher, "The goodness of fit of regression formulae, and the distribution of regression coefficients," J. Royal Statist. Soc., vol. 85, no. 189, pp. 597–612, 1996.
- [37] N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series. New York: Wiley, 1949.
- [38] A. Kolmogorov, Foundations of Theory of Probability. New York: Chelsea, 1956.
- [39] R. E. Kalman, "A new approach to linear filtering and prediction problems," Transactions of the ASME-Journal of Basic Engineering, vol. 82, no. Series D, pp. 35–45, 1960.

- [40] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software. John Wiley & Sons, Inc., 2001.
- [41] É. Bossé, J. Roy, and A. Jouan, "Investigating the performance of some tracking filter schema for the Advanced Shipboard Command and Control Technology (ASCACT) project DREV-TR-1999-208," tech. rep., Valcartier QUE (CAN), October 2000.
- [42] P. Valin, A. Jouan, and É. Bossé, "Demonstration of data/information fusion concepts for airborne maritime surveillance operations, TR 2004-283," tech. rep., 2006.
- [43] A. Jouan, B. Jarry, and H. Michalska, "Tracking closely maneuvering targets in clutter with an IMM-JVC algorithm," Proc. of the third Intl. Conference on Information Fusion, pp. MOD2/10 – MOD2/16 vol.1, 2000.
- [44] M. Hadzagic, H. Michalska, and A. Jouan, "IMM-JVC and IMM-JPDA for closely maneuvering targets," Proc. of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, pp. 1278–1282, 2001.
- [45] B. O. D. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ., 1979.
- [46] M. West, "Mixture models, Monte Carlo, Bayesian updating and dynamic models," Computer Science and Statistics, vol. 24, pp. 325–333, 1993.
- [47] S. Julier and J. Uhlmann, "A new extension of the Kalman filter to nonlinear systems.," Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls, 1997.
- [48] K. Akashi and K. Kumamoto, "Construction of discrete-time nonlinear filter by Monte Carlo methods with variance-reducing techniques," Systems and Control, vol. 19, 1975.
- [49] A. Doucet, N. De Freitas, and N. Gordon, Sequential Monte Carlo Methods in Practice. Springer-Verlag, 2001.
- [50] L. Stone, C. Corwin, and A. Barlow, *Bayesian Multiple Target Tracking*. Artech House, 1999.
- [51] I. Thiburon, Systems, "Over-the-horizon/detection classification and targeting (OTH/DC&T) system level specification ship tracking algorithm," tech. rep., Thiburon Systems Inc., San Jose, CA, 1991.

- [52] M. Hadzagic, H. Michalska, and D. Grenier, "Integrated Ornstein-Uhlenbeck process in application to ship tracking," Proc. of IASTED Signal and Image Processing Conference, (Kailua-Kona, HI, USA), pp. 623–87–623–93, 2008.
- [53] M. Hadzagic, E. Lefebvre, and H. Michalska, *Track-before-detect methods for track-ing low-observable targets*, p. 256. Sensors and Transducers, multisensor data and information processing ed., 2005.
- [54] M. Hadzagic, "Track-before-detect methods for tracking low-observable targets," tech. rep., Montreal, QC, Canada, 2005.
- [55] G. Shafer, A Mathematical Theory of Evidence. Princeton University Press, 1976.
- [56] J. Klir and M. Wierman, Uncertainty based information, vol. 15 of Studies of Fuzziness and Soft Computing. Physica-Verlag, 1999.
- [57] D. Dubois and H. Prade, Combination of Fuzzy Information in the Framework of Possibility Theory, vol. 15 of Data Fusion in Robotics and Machine Learning. Academic Press, 1992.
- [58] L. Sevgi, A. Ponsford, and H. Chan, "An integrated maritime surveillance system based on high-frequency surface-wave radars. Part 1," *IEEE Antennas and Propagation Magazine*, vol. 43, no. 4, pp. 28–43, 2001.
- [59] A. Ponsford, L. Sevgi, and H. Chan, "An integrated maritime surveillance system based on high-frequency surface-wave radars. Part 2," *IEEE Antennas and Propagation Magazine*, vol. 43, no. 5, pp. 52–63, 2001.
- [60] C. Carthel, S. Coraluppi, and P. Grignan, "Multisensor tracking and fusion for maritime surveillance," Proc. of the 10th Intl. Conference on Information Fusion, (Florence, Italy), pp. 1–6, 2007.
- [61] G. Seber, *Linear Regression Analysis*. John Wiley & Sons, Inc., 1977.
- [62] D. Bacon and G. Watts, "Estimating the transition between two intersecting straight lines," *Biometrika*, vol. 58, no. 3, pp. 525–534, 1971.
- [63] P. Feder, "Asymptotics of M-estimators in two-phase linear regression models," EL-SEVIER J. on Stochastic Processes and their Applications, vol. 3, pp. 84–97, 1979.
- [64] T. L. Lai, T. Robbins, and C. Z. Wei, "Strong consistency of least-squares estimates in multiple regression," J. on Multivariate Analysis, vol. 9, pp. 343–361, 1979.
- [65] R. Walkowiak and R. Kala, "Two-phase nonlinear regression with smooth transition," Journal of Communications in Statistics-Simulation and Computation, 2000.

- [66] Y. Yao and S. Au, "Least-squares estimation of a step function," Sankhya (Series A), vol. 51, no. 3, pp. 370–381, 1989.
- [67] P. K. Bhattacharya, "Some aspects of change-point analysis," MS Lecture Notes Monograph Series, vol. 23, pp. 28–56, 1994.
- [68] H. L. Koul and L. Qian, "Asymptotics of maximum likelihood estimator in two-phase linear regression model," J. Statist. Planning Inference, vol. 108, pp. 99–119, 2002.
- [69] G. Ciuperka, "Maximum likelihood estimator in a two-phase nonlinear random regression model," *Statistics and Decision*, vol. 22, p. 335, 2004.
- [70] H. L. Koul, L. Qian, and L. Surgailis, "Asymptotics of M-estimators in two-phase linear regression models," *ELSEVIER J. on Stochastic Processes and their Applications*, vol. 103, pp. 123–154, 2003.
- [71] A. L. Rukhin and I. Vajda, "Change-point estimation as a nonlinear regression problem," *Statistics*, vol. 30, pp. 181–200, 1997.
- [72] G. Ciuperca and N. Dapzol, "Maximum likelihood estimator in a multi-phase random regression model," *Journal of Theoretical and Applied Statistics*, vol. 42, no. 4, pp. 363–381, 2008.
- [73] J. H. Friedman, "Multivariate adaptive regression splines," Ann. Statist., vol. 19, no. 1, pp. 1–141, 1991.
- [74] J. Bai, "Estimating multiple breaks one at a time," *Econometric Theory*, vol. 13, pp. 315–352, 1997.
- [75] J. Bai and P. Perron, "Estimating and testing linear models with multiple structural changes," *Econometrica*, vol. 66, no. 1, pp. 47–78, 1998.
- [76] B. Strikholm and T. Terasvirta, "A sequential procedure for determining the number of regimes in a threshold autoregressive model," *Econometrics Journal*, vol. 9, pp. 472–491, 2006.
- [77] H. Kuchenhoff, "An exact algorithm for estimating breakpoints in segmented generalized linear model," *Computational Statistics*, vol. 12, no. 2, 1997.
- [78] D. Denison, B. Mallick, and A. Smith, "Automatic Bayesian curve fitting," J.J. Statist. Soc. B, vol. 60, pp. 330–350, 1998.
- [79] S. Zhou and X. Shen, "Spatially adaptive regression splines and accurate knot selection schemes," J. Am. Statist. Assoc., vol. 896, pp. 247–259, 2001.

- [80] A. Delaigle and A. Meister, "Nonparametric regression estimation in the heteroscedastic errors-in-variables problem," *Journal of the American Statistical Association*, vol. 102, no. 480, pp. 1416–1426, 2007.
- [81] S. Schennach, "Estimation of nonlinear models with measurement error," *Econometrica*, vol. 72, no. 1, pp. 33–75, 2004.
- [82] J. Opsomer, Y. Wang, and Y. Yang, "Nonparametric regression with correlated errors," *Statistical Science*, vol. 16, no. 2, pp. 134–153, 2001.
- [83] F. El-Hawary and J. Jing, "Robust regression-based EKF for tracking underwater targets," *IEEE Journal of Oceanic Engineering*, vol. 20, no. 1, pp. 31–41, 1995.
- [84] R. Fruhwirth, P. Kubinec, W. Mitaroff, and M. Regler, "Vertex reconstruction and track bundling at the LEP collider using robust algorithms," *Computer Physics Communications*, vol. 96, no. 189, 1996.
- [85] R. Anderson-Sprecher and R. Lenth, "Spline estimation of paths using bearings-only tracking data," *Journal of the American Statistical Association*, vol. 91, no. 433, pp. 276–283, 1996.
- [86] D. Carevic, "Robust estimation techniques for target-motion analysis using passively sensed transient signals," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 2, pp. 263–270, 2003.
- [87] E. B. Nelson, Nonlinear Regression Methods for Estimation. PhD thesis, Department of the Air Force, Air Force Institute of Technology, Air University, Wright-Patterson Air Force Base, Ohio, USA, 2005.
- [88] T. Toledo, N. Koustopoulos, and K. Ahmed, "Estimation of vehicle trajectories with locally weighted regression," *Journal of the Transportation Research Board*, no. 1999, pp. 161–169, 2007.
- [89] N. Agrawal and M. P. Singh, "Trajectory estimation of moving object in space," Proceedings of the Fifth IASTED Intl. Conference on Antennas, Radar and Wave Propagation, (Baltimore, Marlyland, USA), 2008.
- [90] J. Strom-Tejsen, "A digital computer technique for prediction of standard maneuvers of surface ships," Master's thesis, US Navy Department, Washington, DC, USA, 1965.
- [91] C. De Boor, A practical guide to splines. Springer-Verlag, 1978.
- [92] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Phys-ical Review*, vol. 36, pp. 823–841, 1930.

- [93] A. Jazwinski, Stochastic Processes and Filtering Theory. Academic Press, 1970.
- [94] W. Horsthemke and R. Lefever, Noise-Induced Transitions. Springer Verlag, Berlin, Germany, 1984.
- [95] P. Zarchan and H. Mussoff, Fundamentals of Kalman Filtering: A practical approach. AIAA, 2005.
- [96] R. Kass and L. Wasserman, "A reference Bayesian test for nested hypotheses and its relationships to the Shwartz criterion," J. Am. Statist. Assoc., vol. 90, pp. 928–934, 1995.
- [97] C. Robert and G. Casella, *Monte Carlo Statistical Methods*. Artech House, second ed., 2004.
- [98] M. Mitchel, An Introduction to Genetic Algorithm. MIT Press, Cambridge, MA, 1999.
- [99] B. P. Roe, *Probability and Statistics in Experimental Physics*. New York: Springer, 2001.
- [100] M. Hadzagic, SimTrack v.1.0 User Manual. Centre for Intelligent Machines, McGill University, Montréal, QC, Canada, 2008.
- [101] X. Rong Li and Z. Zhao, "Evaluation of estimation algorithms part i: Incomprehensive measures of performance," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 42, no. 4, pp. 1340–1358, 2006.
- [102] R. A. Johnson and J. Wichern, Applied Multivariate Statistical Analysis. Prentice Hall, 2002.
- [103] J. J. Orechovesky, "Single Source Error Ellipse Combination," Master's thesis, Naval Postgraduate School, Monteray, CA, USA, 1996.
- [104] S. Brandt, Data Analysis: Statistical and Computational Methods for Scientists and Engineers. Springer, NY, 1999.
- [105] "Conversion of geodetic coordinates to the local tangent plane." http://www.psas.pdx.edu.
- [106] R. W. Sinnott, "Virtues of the Haversine," Sky and Telescope, vol. 68, no. 2, p. 159, 1984.