Maximum Likelihood Approach to HF Radar Performance Characterization

Craig Carthel[1], Stefano Coraluppi[1], Peter Willett[2], Marco Maratea[3], and Alain Maguer[1]

[1]NATO Undersea Research Centre (NURC)
Viale S. Bartolomeo 400, 19126 La Spezia, Italy
E-mail: carthel@nurc.nato.int, coraluppi@nurc.nato.int, maguer@nurc.nato.int

[2]University of Connecticut – ECE Department
Storrs, CT 06269, USA
E-mail: willett@engr.uconn.edu

[3]Università degli Studi di Genova – Dipartimento di Informatica Sistemistica e Telematica
16145 Genova, Italia
E-mail: marco@dist.unige.it

***Abstract*—As part of NURC maritime surveillance research activities, there is a need to characterize the performance of HF radar vessel-detection processing, in support of subsequent automatic tracking as well as fusion with other available sensor data. In particular, HF radar detection and localization performance for a given contact SNR threshold must be determined. This paper examines two approaches to do so: one based on the maximum-likelihood criterion, the other using a localization threshold on target-originated contacts. The differences in and merits of the two approach are discussed with simulated and real maritime HF radar data. We introduce a multi-stage *track-extract-track* approach to HF radar tracking that is effective in a heavily-cluttered environment. Finally, we introduce scalar information metric for sensor characterization and discuss its relevance to sensor selection and detection threshold determination.**

***Index Terms*—Detection, localization, target tracking, maritime surveillance, HF radar**

I. Introduction

In the absence of precise characterization for the quality of detection-level data, it is of interest to characterize its quality for effective subsequent tracking and fusion. Our interest in doing so is in the context of HF radar maritime surveillance data. Thus, we will first address the problem more generally, and then apply our approach to HF radar data from recent at-sea experimental activities.

Sections 2-3 introduce the precise problem formulation and the localization-based and maximum likelihood approaches to determining target probability of detection and the corresponding localization error. Sections 4 examines these approaches in the context of one-dimensional simulated data, and section 5 examines the multi-dimensional case. Sections 6-7 present the results of our methodology with sea-trial HF radar data that includes range, bearing, and range-rate measurements, as well as an illustration of preliminary HF radar target tracking results. Sections 8-9 study a information-based scalar metric for quantifying sensor quality. Concluding remarks and future directions are in section 10.

A recent overview of NURC research in maritime surveillance is provided in [1]. In the context of target tracking, the maximum likelihood (ML) approach is generally applied to determine the target position and velocity states under non-maneuvering target assumptions; this approach and an extension to dynamic (maneuvering-target) scenarios are discussed in [2].

II. Problem Formulation and Thresholding Approach

Let a sequence of sets of measurements $Z = \{Z_i, 1 \le i \le N\}$ be given, with each set further described by $Z_i = \{Z_{ij}, j = 1,..., m_i\}$, and let the corresponding sequence of true target states $X = \{X_i, 1 \le i \le N\}$ be given as well. The relationship between target states and measurement sets is given by the following. With probability P_D, the set Z_i contains a target-induced measurement $Z_{ij'}$, $1 \le j' \le m_i$ of the true target state, given by:

$$Z_{ij'} = f(X_i) + W_i,\ W_i \sim N(0, \Sigma). \tag{1}$$

All other measurements are false. The number of false measurements is Poisson distributed with known parameter λ, and these are uniformly distributed in measurement space.

It is of interest to estimate the detection probability P_D and the measurement error covariance Σ in equation (1). The parameter λ is known as it can be estimated from the data.

Recently, in [3] we introduced the *system operating characteristics* (SOC) curve that generalizes the well known *receiver operating characteristics* (ROC) curve. The former replaces the *probability of detection* by the *probability of localization*: that is, the probability that there exists a sensor measurement within a specified Mahalanobis distance in measurement space from the error-free measurement. We argue in [3] that the SOC curve may be more operationally relevant that the ROC curve; interestingly, the probability of localization is a function of the probability of detection, localization accuracy, and false alarm statistics.

We take a related approach to the problem of estimating P_D and Σ for the dataset Z. Given a distance threshold, a simple estimate for P_D is given by the fraction of measurement sets in Z with at least one valid (i.e. close-enough) sensor measurement; the closest valid measurement is tagged as target originated. Assuming uncorrelated measurement errors, the RMS error between target-originated and error-free measurements[1] provides standard deviations of all measurement components, in turn leading to a (diagonal) estimate for the measurement covariance matrix Σ.

As we will see in section 4, a weakness in this first approach is that it relies on a distance threshold to define valid measurements, and thus the estimates for P_D and Σ require an effective approach for its selection.

III. The Maximum Likelihood Approach

It can be shown that the likelihood function for the dataset Z is the following [4]:

$$\Lambda = \prod_{i=1}^{N} \left\{ \frac{1-P_D}{u^{m_i}} \mu_\lambda(m_i) + \frac{P_D \mu_\lambda(m_i - 1)}{u^{m_i - 1} m_i} \sum_{j=1}^{m_i} p(Z_{ij} \mid X_i, \Sigma) \right\}. \quad (2)$$

In equation (2), $\mu_\lambda(\cdot)$ is the Poisson pmf with parameter λ, $p(\cdot \mid \cdot, \cdot)$ is the Gaussian pdf conditioned on a given mean and covariance, and u is the volume of the surveillance region.

The maximum likelihood estimates $\overline{P}_D$ and $\overline{\Sigma}$ are then:

$$(\overline{P}_D, \overline{\Sigma}) = \arg \max_{0 \le P_D \le 1, \Sigma > 0} \Lambda(P_D, \Sigma). \quad (3)$$

A limitation in this second approach to sensor characterization is a numerical one: the maximum likelihood function is not always amenable to simple optimization approaches, and may require a good initial guess for the parameters to be estimated. (In practice, for numerical stability it is best to utilize the log-likelihood function.)

IV. The One-Dimensional Case

Before examining the HF radar setting, it is instructive to focus on the scalar case with simple positional measurements. The scenario details are listed in table 1. The target is fixed and located at the measurement space origin.

Table 1. Parameters in one-dimensional example.

Number of scans N	150
Detection probability P_D	0.8
Measurement covariance Σ	1m^2
False alarm rate λ	5
Measurement space	[-30m, 30m]

Figure 1 illustrates a realization of Z. Recall that our task is to extract estimates of P_D and Σ from the data Z and the target ground truth trajectory X.

Applying a distance threshold as described in section 2 over a range of settings yields the curves illustrated in figure 2. As we see, a clear choice of distance threshold is not apparent from the curves. Further, in general there does not exist a choice that leads to accurate estimates for both P_D and Σ.

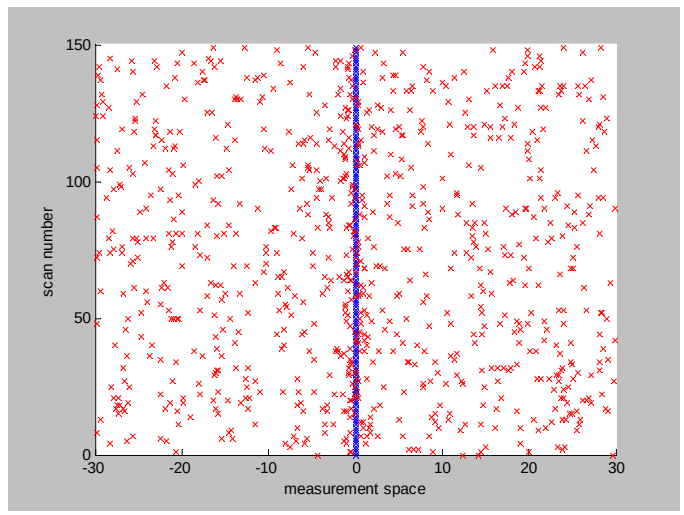

Figure 1. Scenario realization; red crosses denote measurements and blue crosses denote true target locations.

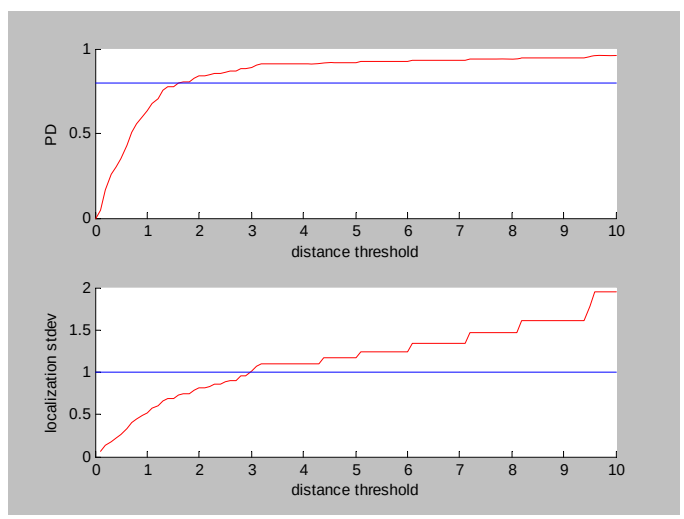

Figure 2. Threshold-based estimation of detection and localization performance (red), as a function of a distance threshold (true values in blue).

[1] The error-free measurement is obtained by transforming the true target state into measurement space.

Next, we consider the likelihood function formalism according to equation (2) and apply two optimization approaches to solving equation (3): a brute-force approach that evaluates the likelihood over a grid of values, and numerical optimization using the `FMINSEARCH` function in `MATLAB`. For the simple example that we examine here, the latter approach is fast, and succeeds in matching the brute-force solution. Figure 3 illustrates the results based on an initial guesses of $P_D = 0.633$ and $\Sigma = 0.518$ that correspond to a distance threshold of $\xi = 1$ in figure 2.

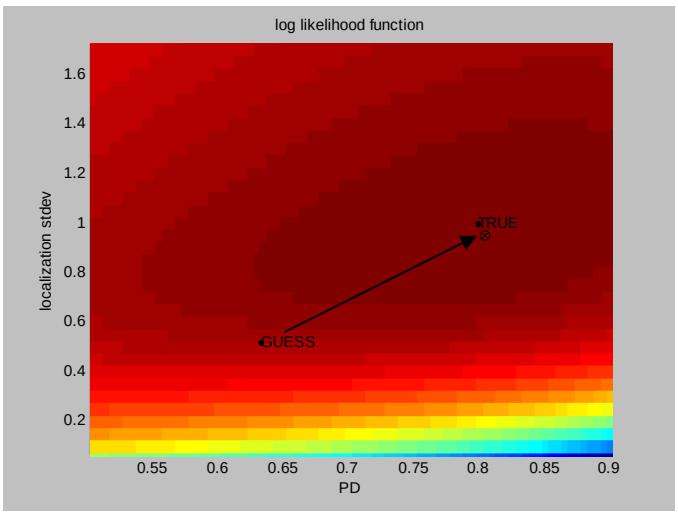

Figure 3. Close match between log-likelihood function maximization with numerical optimization ('x') and brute-force evaluation ('o').

It is interesting to note that the likelihood function evaluated at the *true* values of P_D and Σ does not precisely achieve the maximum. Indeed, the data Z may not always be sufficiently rich to support accurate determination of sensor data quality.

A further point worth noting is that, when employing a numerical optimization technique like `FMINSEARCH`, it is helpful to have a good enough initial guess of parameter values to ensure fast convergence. This points to the usefulness of the threshold-based curves in figure 2 in providing self-consistent initial guesses for P_D and Σ to be used in subsequent maximum likelihood processing.

In using `FMINSEARCH` for (unconstrained) numerical optimization, it is important to introduce a penalty for non-meaningful P_D values, i.e. less than zero or greater than unity. Additionally, it is important not to include the second term in eqn. (2) for those data scans that have no measurements, i.e. for every scan *i* with $m_i = 0$.

As illustrated in figures 4-5, we find consistent, rapid convergence to the maximum likelihood solution with one-dimensional data, for a number of expected false contact rates and initial guesses for P_D and Σ.

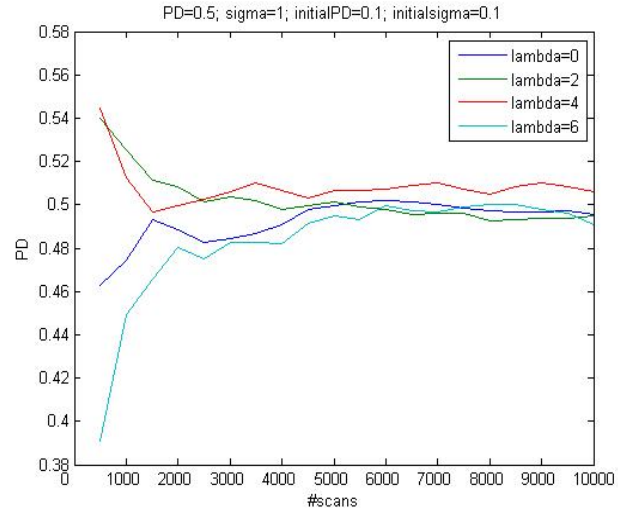

Figure 4. P_D convergence for different values of λ.

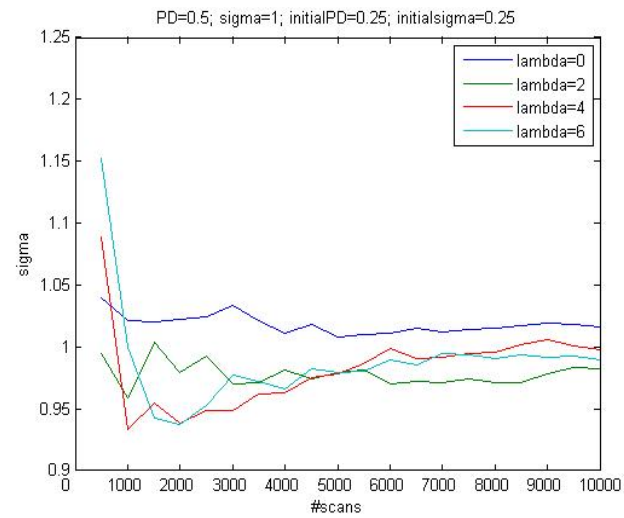

Figure 5. Σ convergence for different values of λ.

V. The Multi-Dimensional Case

Having completed simulation-based validation of the maximum likelihood approach to sensor performance characterization, it is of interest to confirm that the methodology is effective with multi-dimensional measurement data. In particular, it is important to examine convergence in examples with significant dynamic range in measurement error statistics. We perform this study with a generalization to the data simulation described in the previous case; in particular, the measurement space is set to [-10σ,10σ] in each measurement component dimension.

Figure 6 illustrates normalized convergence results with three-dimensional data, with initial guesses for P_D and Σ of half the true values and error statistics of $\sigma_1 = 1$, $\sigma_2 = 10$, and $\sigma_3 = 100$. Likewise, we find correct convergence to the true value of P_D as well.

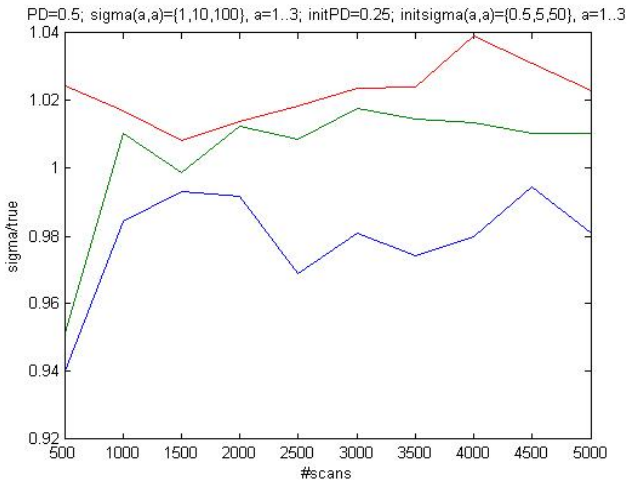

Figure 6. Convergence for multi-dimensional data with very different measurement component error statistics.

VI. Application to Real HF Radar Data

In May 2008, the N.R.V. Alliance participated in an experimental campaign in conjunction with the French company ACTIMAR, based in Brest, France. In particular, ACTIMAR acquired roughly 6hrs of HF radar data using its two shore-based radars.

Using *Automatic Identification System* (AIS) transponder data collected by the Alliance, it is possible to compare large-vessel traffic with the HF radar contact data that results from ship-detection processing [5]. Overlays of AIS tracks and HF radar contacts are illustrated in figures 7-8.

A quick inspection illustrates the difficulty in assessing the quality of the HF radar detection data. Our approach is to rely on N.R.V. Alliance navigational information as ground truth, and to apply the maximum likelihood methodology introduced in this paper. Note that the measured data is a nonlinear function of the target kinematic state, and includes range, bearing, and Doppler. A further note is that, for the purposes of this analysis, detections on other vessels are treated as false contacts.

We have limited the analysis to a time window for which the Alliance and a bounding region around her are fully contained in the footprint of a given HF radar. In particular, we have 111 detection files for each sensor, and the special window is given by 20km in range, 30deg in bearing, and 10m/s in Doppler. Further, the maximum likelihood (ML) processing assumes known false contact statistics. Thus, as a first step, we compute the average number of returns, i.e. λ in eqn. (2). Of course, there is variability in the number of returns per scan: see figure 9 for scans from 100 to 210. We perform our analysis for those contacts with SNR>25dB.

The HF radar performance characteristics as estimated with our maximum likelihood methodology are summarized in table 2. We note that the FAR is directly related to the spatial window defined above. The remaining statistics in principle are independent of the spatial window, though we have noted some statistical fluctuations as we vary its size, an issue that will require further analysis.

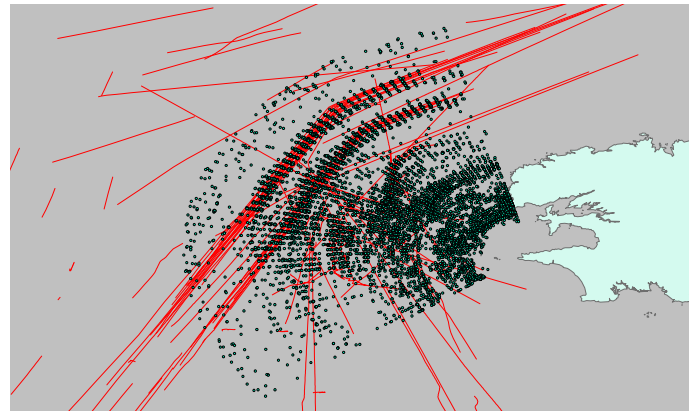

Figure 7. AIS tracks (red) and HF radar contacts from Brezellec.

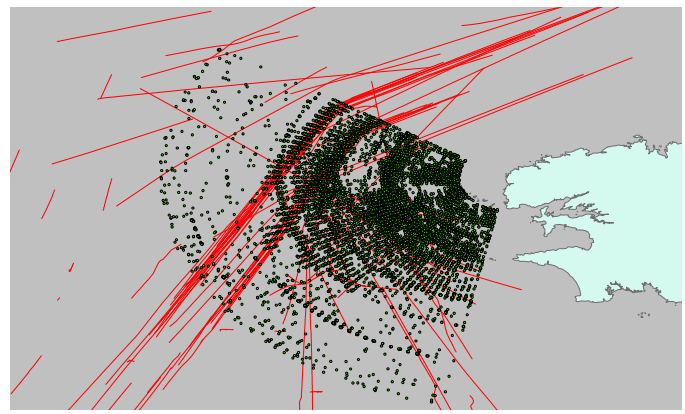

Figure 8. AIS tracks (red) and HF radar contacts from Garchine.

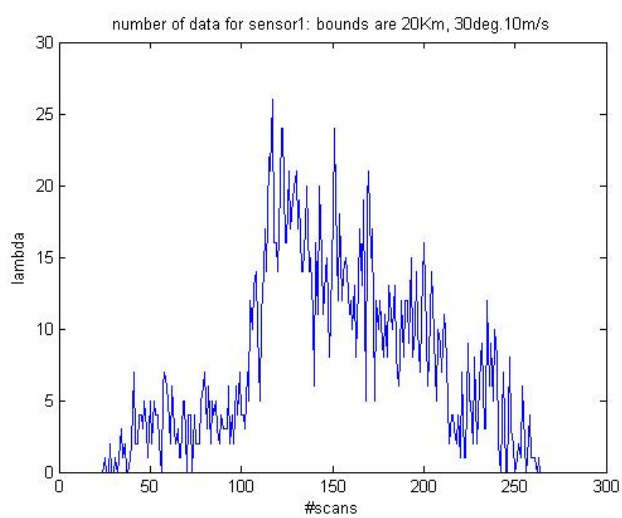

Figure 9. Number of contacts per scan from one of the sensors; relevant scans are 100 to 210, the time period during which bounding region around N.R.V. Alliance is fully within the sensor footprint.

Table 2. Estimated HF radar performance characteristics: measurement error standard deviations, and detection statistics.

sensor	*range*	*bearing*	*Doppler*	P_D	*FAR*
Brezellec	1.05km	2.5deg	0.17m/s	0.28	13.0
Garchine	818m	2.1deg	0.11m/s	0.51	9.0

We note that the two sensors do not yield the same performance characteristics. This is not surprising, as we have focused on a specific vessel of opportunity and detection

performance is known to depend on target range and aspect. It is important to note that we have consistent results in terms of the relative quality of detection and localization performance, as these are known to be coupled: higher contact SNR leads to improved localization performance [6].

VII. A First Look at Tracking Performance

From a maritime surveillance perspective, it is of interest to generate HF radar tracks for subsequent correlation with AIS and other available sensor feeds. To do so, we have extended the NURC *distributed multi-hypothesis tracker* (DMHT) that was first developed for active sonar tracking [7]. The salient extensions that enable maritime surveillance are the following:

- Inclusion of RADAR, SAR, and AIS sensor models;
- Track-specific tangential coordinate system allowing for global surveillance;
- Track management logic to handle cooperative sensor data, like AIS, that exhibits persistent identity information.

A discussion of multi-sensor maritime surveillance is beyond the scope of this paper, and will be documented in a forthcoming publication. Here, we provide a first illustration of HF radar tracking performance with the data acquired in Brest. The high-confidence radar tracks (those containing at least 40 radar contacts) are shown in figure 10.

Figure 10. AIS tracks (red) and HF radar tracks (purple).

It is clear that, from a first assessment, effective HF radar tracking is quite challenging. The radars yield few tracks in the more distant vessel sea lane: analysis of the AIS data revealed that only large tankers are tracked. There is greater success with the closer sea lane, though here again performance is limited. Many near range vessels identified by AIS are not found in the HF radar tracks. Interestingly, there is a significant HF radar track with a northeast heading, likely to be a large vessel, which is not present in the AIS data. This illustrates the potential of multi-sensor coverage to reveal anomalous vessel behavior.

We address next the potential advantages of a *track-extract-track* multi-stage tracking architecture. In this approach, we extract contact data from the first stage of tracking, and proceed with a second stage of tracking with the remaining contacts; the process can be iterated with a third stage of processing, and so on. The key advantage of the approach is that it allows for additional, weaker target tracks to be extracted from the data. This is accomplished by increasing the data correlation gates in the tracker, as well as lowering the track-confirmation criterion. In principle, a similar result could be achieved with centralized processing, but with a much more complex adaptive-tracking methodology.

We illustrate the *track-extract-track* processing approach in the context of HF radar data in figures 11-14, which provides a display of additional tracks beyond those identified in figure 10. Note that the superposition of all tracks (figure 15) yields a display that is heavily cluttered, yet comparable performance proved to be impossible to achieve with a single stage of (non-adaptive) target tracking.

In particular, the additional track extraction based on the *track-extract-track* multi-stage processing paradigm is as follows: previous tracks (fig. 6.1) included >40 contacts; the additional track shown here include >35 contacts (figure 11), >30 contacts (figure 12), >25 contacts (figure 13), and > 20 contacts (figure 14).

We recommend continued investigation into novel multi-stage data fusion architectures for application in a variety of surveillance settings. These approaches often hold considerable potential for dramatic performance improvements that are difficult to obtain with refinements in sensor data processing or nonlinear filtering algorithms, whether analytical or numerical.

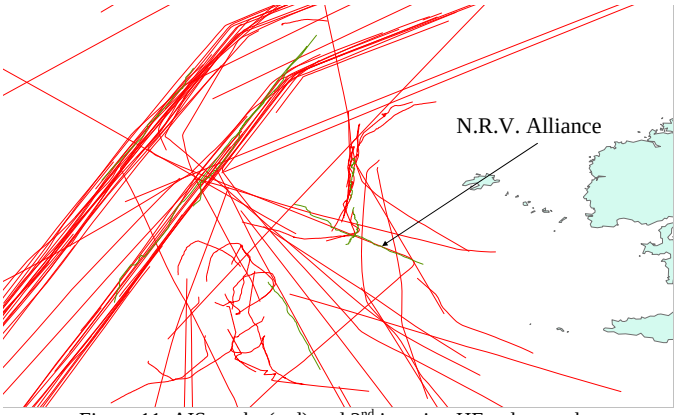

Figure 11. AIS tracks (red) and 2nd iteration HF radar tracks.

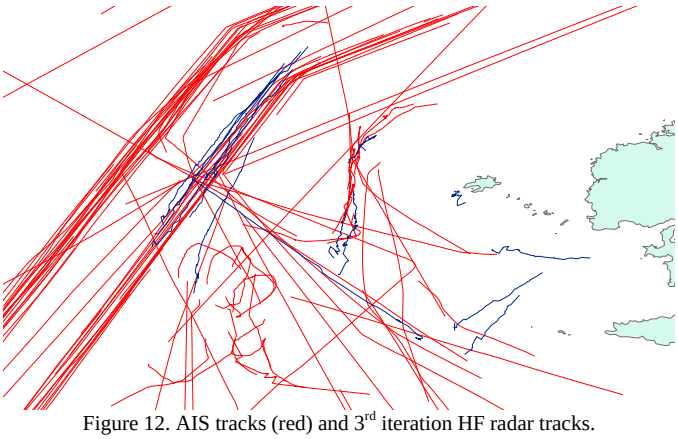

Figure 12. AIS tracks (red) and 3rd iteration HF radar tracks.

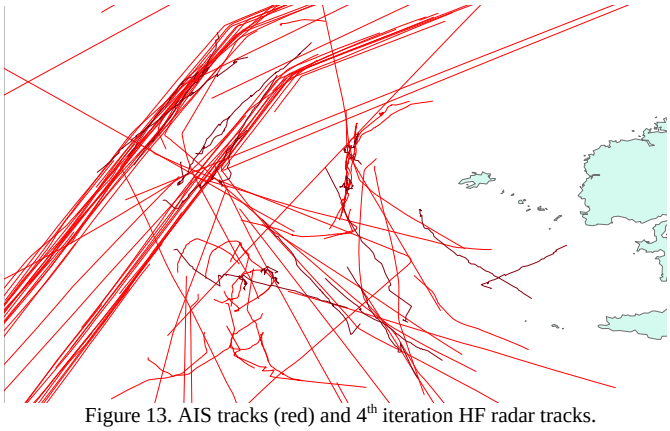

Figure 13. AIS tracks (red) and 4th iteration HF radar tracks.

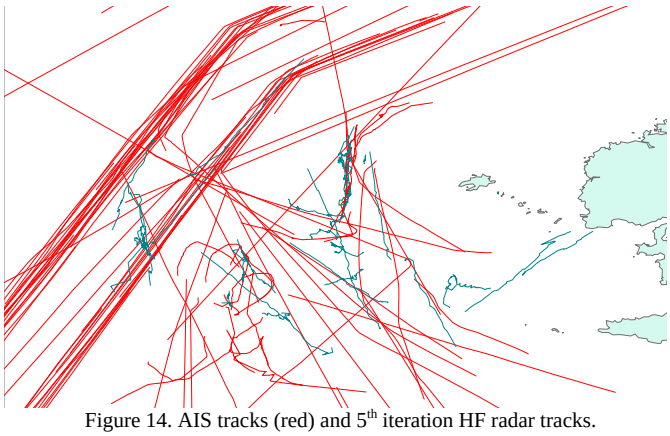

Figure 14. AIS tracks (red) and 5th iteration HF radar tracks.

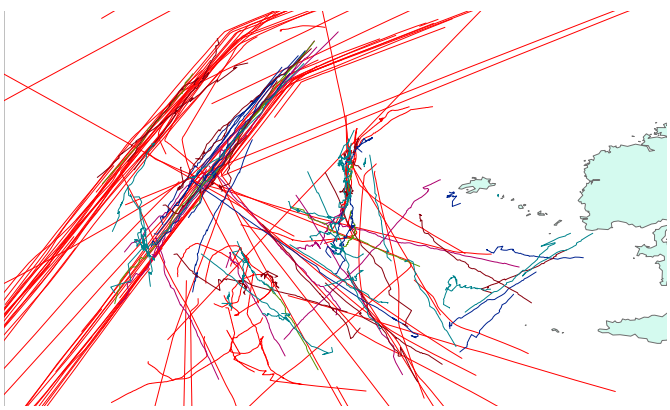

Figure 15. Consolidated HF radar adaptive tracking result based on an iterative *track-extract-track* scheme.

VIII. Information-Based Sensor Quality

As we have seen, subject to sufficient care in numerical optimization processing, the maximum likelihood criterion leads to statistically consistent estimates of sensor P_D and Σ. Combined with the false alarm rate λ, the parameter triple (P_D, Σ, λ) provides a complete characterization of sensor performance.

It is interesting to ask whether a single scalar measure of performance can be computed as a function of this parameter triple. Such a metric would have direct applicability, for instance, to sensor selection and to detection threshold selection. We wish to determine a simple answer to the question: is one triple better than another?

We proceed with a simple information-based approach. Assume for simplicity that the target is located at the measurement space origin. We consider a mixture pdf that is achieved by appropriately weighing the target-induced measurement pdf $N(0,\Sigma_X)$ with the false-alarm induced measurement pdf $U(0,\Sigma_Y)$, where $U(0,\Sigma_Y)$ is the zero-mean uniform distribution with covariance Σ_Y. We assume $\Sigma_Y > \Sigma_X$: indeed a target-originated measurement is more informative than a false alarm.

Let $X \sim N(0,\Sigma_X)$, $Y \sim U(0,\Sigma_Y)$, and $\gamma \sim B\left(P_D/\lambda+P_D\right)$, where $B(p)$ is the Binomial pmf that equals one with probability *p*. Define the mixture random variable Z as follows:

$$Z = \gamma X + (1-\gamma)Y \,. \tag{4}$$

Note that the weight given to X in this mixture matches the (normalized) frequency of occurrence of target-induced measurements.

It can be shown that, assuming independence of X, Y, and γ, the covariance of Z is given by the following:

$$\Sigma_Z = \Sigma_Y - \frac{P_D}{\lambda + P_D}(\Sigma_Y - \Sigma_X) \,. \tag{5}$$

At each sensor scan, the expected number of returns is $\lambda + P_D$. For each return, the determinant of the inverse of Σ_Z provides an information measure that characterizes the value of the measurement concerning the target state. Thus, the overall information measure is given by:

$$J = (\lambda + P_D)\det\left(\Sigma_Y - \frac{P_D}{\lambda + P_D}(\Sigma_Y - \Sigma_X)\right)^{-1} . \tag{6}$$

As we expect from our intuition, analysis of equation (6) leads to the following:

- For $\Sigma_Y \to 0$ (and thus $\Sigma_X \to 0$): $J \to \infty$;
- For $\Sigma_X \to \infty$ (and thus $\Sigma_Y \to \infty$): $J \to 0$;
- For $\lambda \to 0$: $J \to P_D \det \Sigma_X^{-1}$;
- For $P_D \to 0$: $J \to \lambda \det \Sigma_Y^{-1}$.

From a target tracking perspective, it is of interest to explore the implications of the simple metric discussed here compared to approaches to sensor selection based on tracking filter errors (see e.g. [8-9]).

IX. Optimal Choice of Detection Threshold

Proceeding with the metric *J* introduced in the previous section, it is of interest to examine the optimal selection of detection threshold when the target *SNR* is known. Assuming Rayleigh distributed returns, a fixed scaling parameter *A*, and

a detection threshold *DT*, we have the following expressions for P_D and λ:

$$P_D = \exp\left(-\frac{DT}{1+SNR}\right), \tag{7}$$

$$\lambda = A\exp(-DT). \tag{8}$$

Substituting expressions (7-8) into equation (6), we have a simple objective function that, for known scenario parameters, may be optimized to determine the optimal sensor detection threshold DT^{opt}.

Table 3 lists the parameter settings for a simple scalar illustrative example, and figures 11-12 plot $\det\Sigma_Z^{-1}$ and the information measure *J* for a range of detection thresholds. Interestingly, we find that $\det\Sigma_Z^{-1}$ increases monotonically with increasing detection threshold (i.e. measurements are on average more informative, as the mixture tends towards the target-induced pdf). On the other hand, an increasing detection threshold reduces the expected number of returns $\lambda + P_D$. Overall, figure 12 suggests that there is an optimal choice of detection threshold that maximizes information on the target. In the example, optimality is achieved at $DT^{opt} = 15.1$, leading to $P_D = 0.253$ and $\lambda = 2.77 \cdot 10^{-5}$.

Table 3. Parameter settings for detection threshold optimization example.

Target *SNR*	10
Number of detection cells	100
Σ_Y	1000m^2
Σ_X	1m^2

It should be noted that degeneracy is observed for scenarios with a large number of detection cells (i.e. large scaling parameter *A*). In this setting, optimality is achieved for a vanishingly small detection threshold, as illustrated in figure 13. The explanation for this phenomenon is that, with a finite Σ_Y, false returns do provide information, and a sufficient number of these is available with large *A* and low detection threshold to exceed the local maximum observed at a higher (non-degenerate) detection threshold.

It will be of interest to compare the information-based analysis documented here with model-based tracker ROC curves [10] and, ultimately, with simulation-based and real-data tracking results. Note that the results in [10] exhibit a similar effect with non-monotonic tracker ROC curves.

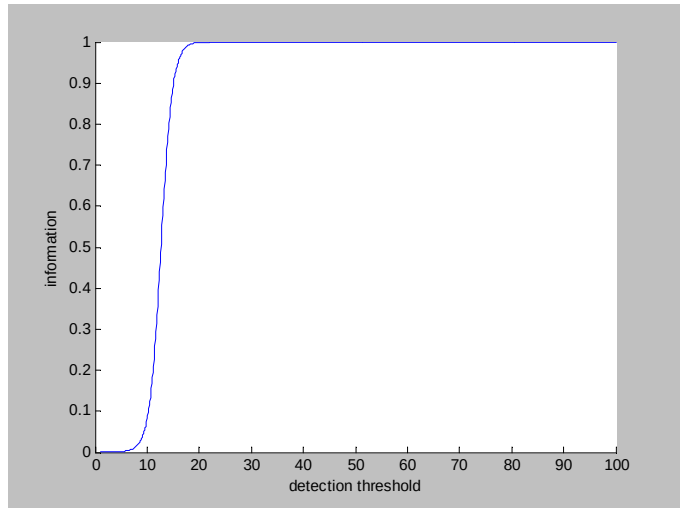

Figure 11. Expected information per return as a function of sensor detection threshold.

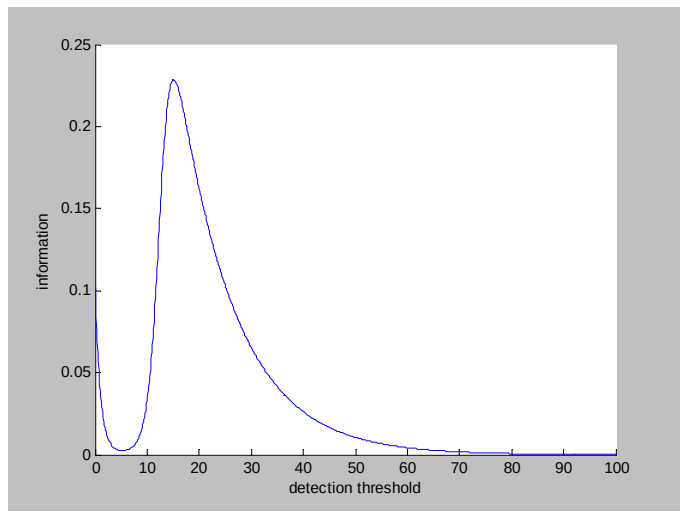

Figure 12. (Aggregate) information as a function of sensor detection threshold.

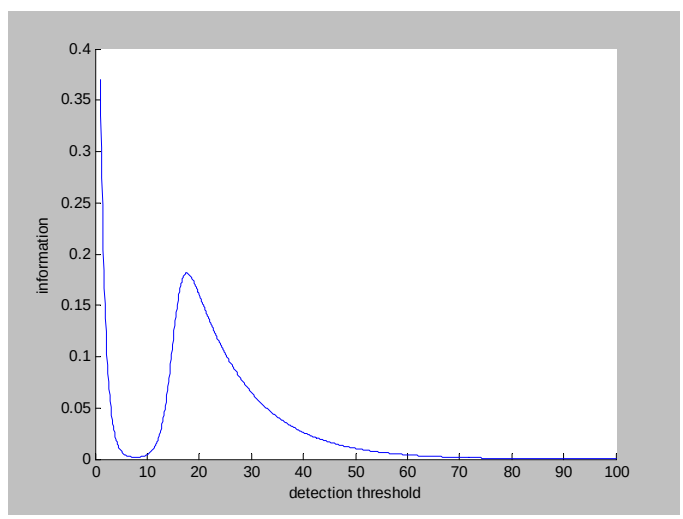

Figure 13. The same example, with large number of detection cells: degeneracy observed whereby optimality is achieved at high contact rates.

X. Conclusions and Future Work

This paper addresses the estimation of sensor detection and localization statistics. We propose a maximum likelihood (ML) methodology, and validate the approach with simulated data. A recent, related work of interest is [11]. In [11] the authors do not have target ground truth information, and estimate noise variances as a by-product of the tracking process.

We apply our ML methodology to HF radar data from a recent sea trial, for which coincident AIS data is available, and provide a preliminary assessment of the quality of the data to support automatic tracking. Our approach to tracking utilizes an iterative *track-extract-track* technique, which provides better performance than non-adaptive single-stage tracking and is simpler than a complex, adaptive track initiation schemes to identify weak and strong targets in a single processing stage.

Finally, the paper proposes a scalar sensor-quality metric that is a function of detection and localization performance. We apply this metric to determine the optimal sensor SNR threshold. It is of interest to investigate whether those thresholds that provide informative sensor data lead to good automatic tracking results. Encouragingly, we know that tracker output ROC curves exhibit a similar non-monotonic behavior as found here with the proposed information metric.

Acknowledgment

The authors wish to thank Nicholas Thomas (thomas@actimar.fr) and ACTIMAR (www.actimar.fr) for providing the HF radar data used in this study.

P. Willett was supported by the Office of Naval Research under contract N00014-07-1-0429.

References

[1] S. Coraluppi, C. Carthel, and A. Maguer, Maritime Surveillance Research at NURC, in *Proceedings of the NATO RTO SET Symposium on Sensors and Technology for Defence Against Terrorism*, April 2008, Mannheim, Germany.

[2] W. Blanding, P. Willett, and S. Coraluppi, Sequential ML for Multistatic Sonar Tracking, in *Proceedings of OCEANS 2007*, June 2007, Aberdeen, Scotland.

[3] S. Coraluppi, M. Guerriero, and P. Willett, Contact Fusion in Large Sensor Networks: Operational Performance Analysis, in *Proceedings of the NATO RTO SET Panel Symposium on Sensors and Technology for Defence Against Terrorism*, April 2008, Mannheim, Germany.

[4] T. Kirubarajan and Y. Bar-Shalom, Low observable target motion analysis using amplitude information, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32(4), 1996.

[5] T. Gorski, J.-M. Le Caillec, A. Kawalec, W. Czarnecki, M. Lennon, and N. Thomas, Target Detection using HF Radar Data, in *Proceedings of OCEANS 2007*, June 2007, Aberdeen, Scotland.

[6] W. Burdic, *Underwater Acoustic System Analysis*, Prentice Hall, 1984.

[7] S. Coraluppi and C. Carthel, Distributed Tracking in Multistatic Sonar, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41(3), July 2005.

[8] D. Kershaw and R. Evans, Waveform Selective Probabilistic Data Association, *IEEE Transaction on Aerospace and Electronic Systems*, vol. 33(4), October 1997.

[9] B. La Scala, M. Rezaeian, and B. Moran, Optimal Adaptive Waveform Selection for Target Tracking, in *Proceedings of the 8th International Conference on Information Fusion*, July 2005, Philadelphia PA, USA.

[10] S. Coraluppi, M. Guerriero, and P. Willett, Optimal Fusion Performance Modeling in Sensor Networks, in *Proceedings of the 11th International Conference on Information Fusion*, July 2008, Cologne, Germany.

[11] R. Osborne, Y. Bar-Shalom, and T. Kirubarajan, Radar Measurement Noise Variance Estimation with Several Targets of Opportunity, *IEEE Transaction on Aerospace and Electronic Systems*, vol. 44(3), July 2008.