Enhancing Performance and Reliability in Maritime Target Fusion Through a Context-Driven Approach

1st Pablo Rangel
Digital Systems Group / Program Systems Eng.
Brazilian Navy Research Institute / UFRJ
Rio de Janeiro, Brazil
pablo.rangel@marinha.mil.br

2nd José Gomes de Carvalho Jr.
Digital Systems Group
Brazilian Navy Research Institute
Rio de Janeiro, Brazil
junior.carvalho@marinha.mil.br

3rd Luiz Fernando Yuan Gouvêa
Digital Systems Group
Brazilian Navy Research Institute
Rio de Janeiro, Brazil
luizyuan@gmail.com

4th José Ricardo Potier de Oliveira
Digital Systems Group
Brazilian Navy Research Institute
Rio de Janeiro, Brazil
jricpotier@gmail.com

5th Karen da Silva Cardoso
Digital Systems Group
Brazilian Navy Research Institute
Rio de Janeiro, Brazil
karen.cardoso@marinha.mil.br

Abstract—Multi-sensor-based surveillance systems with overlapping detection areas must address the challenge of associating objects detected by different sensors and accurately determining whether they correspond to the same real-world entity. Notably, different sensors may have distinct state vectors (e.g. AIS, radar, and passive sonar), and in some cases, information from systems such as AIS may be unavailable.

Additionally, the area under surveillance may be vast and contain a large number of vessels, leading to a high computational cost for track association. This study aims to simulate the aforementioned scenario, which is commonly observed along the Brazilian coast today, and to apply a combination of association and fusion techniques to enhance problem-solving. Some models are classical, while others introduce contextual adaptations or modifications to their original conditions of use. Statistical results indicate that contextualization significantly reduces the execution time of multi-sensor association algorithms, while adaptations in classical algorithms improve target fusion across different sea conditions.

Index Terms—track-to-track association, track-to-track fusion, multi-sensor, heterogeneous track association.

I. INTRODUCTION

Coastal surveillance systems must integrate data from multiple sensors and external systems. Numerous studies in this field propose models with high accuracy and precision, often utilizing radar plots as the primary input for estimators. However, in real-world scenarios, the available equipment is often sourced from multiple vendors, featuring different brands and varying levels of information. For instance, some radars provide kinematic tracking data but lack detailed information about associated errors. These radars generally do not output raw plots, as they have already undergone internal estimation processes and only provide track data, such as position, speed, and course. A similar situation occurs with Automatic Identification System (AIS) equipment and tracks obtained from

external systems, including those from other governmental agencies.

Furthermore, some sensors provide only bearing information, such as passive sonars and Electronic Measurement Systems (EMS). This diversity of sensors creates scenarios where it is necessary to analyze state vectors of different types. Specifically, we can identify cases where [1]:

1) **Homogeneous track association:** refers to the process of associating tracks provided by sources that estimate the same state vector (position and velocity), such as radar and AIS tracks. The same applies to tracks originating from data links or external systems;

2) **Heterogeneous track association:** refers to the process of associating tracks from sensors that provide only bearing information, such as passive sonars and electromagnetic emission detectors.

This lack of knowledge about the estimation errors in the acquisition process and the heterogeneity of data provided highlights the need for a track-to-track association (T2TA) before track-to-track fusion (T2TF). The main question, therefore, is: How can such heterogeneous data be combined without knowledge of the inherent measurement errors over time?

T2TA relies on a pairwise combination process, which involves calculating the covariance and cross-covariance of measurements from two sources, as discussed in [2], [3]. In real-world surveillance scenarios, such as along the Brazilian coast with over 100,000 tracks, this combinatorial process leads to extremely high computational costs, even when leveraging optimized solutions.

As a result, a critical question is to design a computationally feasible model to associate and fuse homogeneous and heterogeneous tracks in real-world scenarios with thousands of vessels and dozens of sources. In order to achieve the goal of providing a complete tactical overview of this complex scenario, a practical and low-computational approach must be

sought. To address these questions, a robust model for associating and fusing tracks from homogeneous and heterogeneous sources is proposed that can handle a large number of tracks while achieving high precision and accuracy, as demonstrated in tests.

We hypothesize that restricting the spatial context can reduce the computational cost associated with pairwise tracking comparisons, since clustering candidate tracks within limited geographic areas, unnecessary computations will be avoided. Additionally, to address the lack of measurement background knowledge, we propose incorporating a temporal context. By analyzing kinematic data from different sensors within a short time window, we believe it is possible to calculate and infer potential similarities using statistical tests.

II. RELATED WORKS

Several studies explore ways to improve the performance of fusion techniques; however, they often do not consider computational costs, the context in which the methods will be applied, or the full sequence of operations required in real-world scenarios. These scenarios typically involve multiple types of sensors, unknown estimation parameters from off-the-shelf equipment (such as AIS), and a vast number of objects that need to be associated and fused within short processing times.

In [4], for example, various considerations are made regarding different approaches to implementing T2TF without feedback, partial feedback, and full feedback, each requiring different levels of communication between local trackers and the fusion center. The study also introduces T2TA, which utilizes track estimates within a sliding time window, and examines the impact of cross-covariance among track estimates.

The so-called "common process noise effect" has been investigated in several studies. In [1] it is emphasized that there is no way to capture these effects exactly in the case of heterogeneous state vectors. Differences in motion models across different sensors prevent the precise evaluation of the cross-covariance matrix. Although an approximate calculation technique for the cross-covariance matrix between two heterogeneous sensors has been proposed, the authors highlight that its accuracy is poor and that its results can be neglected.

In [5], the authors conclude that it is possible to achieve good results in target association even if the effects of cross-covariance are neglected.

In [6], the approach is based on the unscented transform, which does not require knowledge of the cross-covariance between the heterogeneous tracks.

In [7], within the field of heterogeneous tracks association, a passive sonar ranging algorithm is proposed based on the measurement of the noise power spectrum of a target, using a ship noise model and a sound absorption model in seawater. This is a possible next step for the system analyzed here.

III. METHODOLOGY

A. Architecture

To tackle the diverse data types and computational challenges described in the motivation section, we propose a multi-layered system architecture, which is currently under development and undergoing testing.

This architecture first fuses radar plot data, generating tracks in regions with overlapping radar coverage. This forms the plot fusion layer, which is performed by an IMM algorithm and generates the radar tracks. IMM is an estimation approach that employs multiple Kalman filters, each representing a possible target motion mode. The estimates generated by these filters are then combined and weighted according to their respective probabilities, resulting in a single, more accurate estimate of the target's state ([8]–[10]).

Subsequently, based on these radar tracks and tracks received from two other sources - AIS and an acoustic passive sensor (passive sonar) - a geo-contextualized mapping of areas of interest is produced to group tracks into spots (track-to-spot association), aiming to reduce the effort of the next layer, the T2TA, which investigates fewer pairwise associations.

The next step is T2TA of radar tracks and tracks from other sources whose estimators are unknown (and consequently, the covariance matrix of the estimation errors is unknown). This track association is performed in each spot identified in the previous step (track-to-spot) and calculated according to the type of information provided by the sensor about the track, whether homogeneous or heterogeneous.

The overall system architecture is illustrated in Figure 1, showing the three types of sensors, the generated tracks (with and without known estimators), and the sequential steps involved: radar plot fusion, radar track estimation, track-to-spot association (track clustering), T2TA of AIS and Radar tracks, track fusion of the associated tracks and, T2TA of passive sonar tracks and fused tracks (heterogeneous association).

B. Target simulation

The target motion is based on the Discrete White Noise Acceleration Model [10], which models the dynamics of a uniformly moving target with additional noise $\nu(k)$, assumed to be constant during the kth sampling period. The noise is interpreted as an acceleration that occurs during the sampling period and disturbs the velocity and position. The state equation for this modeling is (1), (2) and (3):

$$x(k+1) = Fx(k) + G\nu(k) \qquad (1)$$

$$F = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \qquad (2)$$

$$G = \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix} \qquad (3)$$

The noise covariance is given by (4). Changes in the target's velocity should obey the standard deviation $T\sigma_\nu$ and in the simulator, they are set to 10% of the target's maximum velocity.
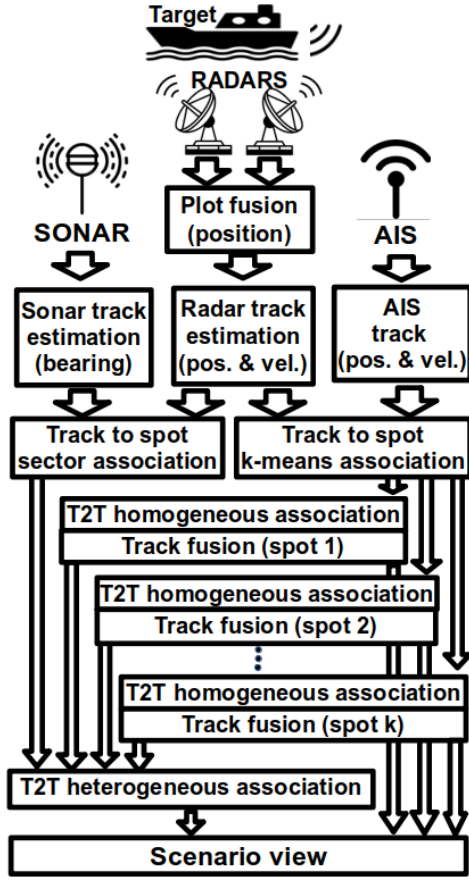
Fig. 1. System architecture and its various levels of association and fusion.

$$G = \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} \\ \frac{T^3}{2} & T^2 \end{bmatrix} \sigma_\nu^2. \tag{4}$$

For targets with acceleration, the reasoning is the same. The model used is the Discrete Wiener Process Acceleration Model and the standard deviation is σ_ν as explained in [10].

C. Plot fusion and radar track estimation (PF)

As explained in section III-A radar plot fusion model utilizes the Interacting Multiple Model (IMM) technique. In the context of a coastal surveillance system, plot fusion operates at a regional level, typically occurring when there is overlap between local radar coverage areas. This process handles data volumes that are not exceptionally large.

After regional centers estimate radar tracks using IMM models, these tracks are transmitted to high-level centers. This stage poses a significant computational challenge due to the large volume of tracks that need to be associated. At this level, radar tracks must be correlated with those generated by other sensors, such as AIS, external systems, and passive sonar. Notably, plot fusion and radar track estimation constitute a localized and inherently distributed process. However, this distribution does not mitigate the considerable computational burden imposed on the high-level center.

D. Track-to-spot association (T2SA)

The high-level fusion center faces a significant computational burden due to the processing of tracks originating from the entire National Jurisdictional Area or even a global scale. Since the subsequent T2TA step involves evaluating billions of potential pairwise relationships in real time, a clustering procedure is implemented to group tracks into spatially related clusters, which are referred to as "spots". This approach effectively reduces the T2TA problem to the set of tracks within each individual spot.

Track-to-spot association is performed in two distinct steps, utilizing different models depending on the type of association: homogeneous or heterogeneous. For homogeneous tracks, whose state vectors include position (e.g., radar and AIS tracks), the K-means algorithm — a widely used unsupervised clustering method — is applied. K-means partitions n tracks into k clusters (spots), assigning each track to the spot with the nearest mean (centroid), which serves as the representative prototype of the group. Essentially, the algorithm groups tracks into k distinct spots based on their inherent similarity, specifically geographical position.

The algorithm's iterative process begins with the random selection of k geographical points as initial spot centers (centroids). Subsequently, each track is assigned to the nearest centroid using Euclidean distance as a similarity metric. The centroids are then recalculated as the mean of the tracks assigned to each cluster. These assignment and update steps are repeated iteratively until the centroids converge to a stable position (i.e. when changes become insignificant) or until a predefined maximum number of iterations is reached.

K-means was selected for its conceptual simplicity and computational efficiency when dealing with a potentially large number of tracks. A relevant characteristic of the model is its sensitivity to initial centroid placement; therefore, the centroids are initialized using a random distribution. Although the algorithm implicitly assumes that clusters have approximately spherical shapes and similar sizes, certain known limitations of K-means, such as the need to specify k a priori, the difficulty in handling noise and outliers, and its inadequacy for clusters with non-convex shapes, do not significantly affect this specific track-to-spot application. Therefore, K-means remains a valuable choice due to its scalability and ease of implementation.

For heterogeneous tracks that provide only bearing information, clustering is performed based on the passive sonar position. Eight 45° sectors (octants), bounded by the passive sonar range, are created. All tracks within a given sector, regardless of their source (passive sonar or other), are clustered within that spot.

Each resulting spot and its associated list of tracks obtained in this track-to-spot association step, are then submitted to the homogeneous T2TA processes in the next processing stage.

E. Track-to-track association (T2TA)

Track association involves evaluating the hypothesis that two tracks from different sources correspond to the same real-

world object (target) and is carried out in two steps.

In the first step, homogeneous tracks (AIS and radar) within each K-means spot are associated, resulting in fused tracks. In the second step, these homogeneously fused tracks are then associated with passive sonar tracks within an octant of the passive sonar system (heterogeneous association).

Both steps evaluate the hypothesis that two estimates, $\hat{X}_1$ and $\hat{X}_2$, from independent sensors will exhibit a larger difference if they correspond to different targets and a smaller difference if they correspond to the same target.

A scalar variable $x = \Delta^T \cdot \Phi^{-1} \cdot \Delta$ is defined, where Δ represents the difference vector between the estimates $\hat{X}_1$ and $\hat{X}_2$ (estimation error) and Φ is the covariance matrix of this error. The variable x follows a Chi-Squared (χ^2) distribution and has unit variance [8], [10]. For the proposed test, the gains and covariance matrices of the estimators, P_1 and P_2, obtained from the recursions in the Kalman filters, as well as the cross-covariance P_{12} required to determine Φ have been extensively discussed in various studies [3], [8], [10].

In summary, we can calculate Φ as shown in (5) and P_{12} as shown in (6), both of which are defined in [8]:

$$\Phi = P_1(k|k) + P_2(k|k) - P_{12}(k|k) - P_{21}(k|k) \quad (5)$$

$$\begin{aligned} P_{12}(k|k) = & \\ & [I - K_1(k)H_1] \cdot \\ & [F(k-1)P_{12}(k-1|k-1)F^T(k-1) + Q] \cdot \\ & [I - K_2(k)H_2]^T \end{aligned} \quad (6)$$

P_{12} is recursively calculated, assuming dependence due to common process noise. However, when working with equipment that acts as a black box, this information is not available. Our approach assumes that by estimating the covariance matrix in the conventional way (using a sample of data within a time window of length p), as shown in (7), we can obtain a satisfactory value.

$$Cov\{W\} = \frac{1}{p-1} \sum_{i=1}^{p} (W_i - \bar{W}) \cdot (W_i - \bar{W})^T \quad (7)$$

Since the calculation of the Φ matrix lacks some necessary information, it is computed based on the above definition, using a sliding window containing the last p values of Δ. The window is utilized to estimate the covariance matrix for testing a single difference vector Δ, rather than all vectors within the window.

As a result, the χ_n^2 variable formed (neglecting temporal correlation) has n degrees of freedom, corresponding to the dimension of the state vector.

Once calculated, x is compared to a threshold from the χ^2 distribution table, corresponding to the same degrees of freedom and a desired confidence level. If x falls below the threshold, the test confirms that the estimates correspond to the same target.

The method requires synchronized sensors. Since, in practice, sensors are not synchronized, it is necessary to align them in time, considering their different acquisition rates. Thus, the most recent data from the sensor (usually the one with the highest acquisition rate) is extrapolated to align it with the other sensor [2].

In the first step, for tracks from homogeneous sensors and state vector as defined in (8), the hypothesis test takes this into account.

$$X = [p_x \ p_y \ v_x \ v_y]^T \quad (8)$$

Hence, for example, we defined Δ as shown in (9):

$$\begin{aligned} \Delta = \hat{X}_{AIS} - \hat{X}_{radar} = & \\ & \begin{bmatrix} \hat{p}_x & \hat{p}_y & \hat{\dot{p}}_x & \hat{\dot{p}}_y \end{bmatrix}^T_{AIS} - \\ & \begin{bmatrix} \hat{p}_x & \hat{p}_y & \hat{\dot{p}}_x & \hat{\dot{p}}_y \end{bmatrix}^T_{radar} \end{aligned} \quad (9)$$

In the second step, which involves the association of tracks from heterogeneous sensors, the state vector considered for the tracks consists of bearing and angular velocity, calculated from successive measurements over a given time interval (10):

$$X = [\theta \ \dot{\theta}]^T. \quad (10)$$

For passive sonar tracks, the bearing estimate $\hat{\theta}$ is directly measured, while the angular velocity $\hat{\dot{\theta}}$ is computed.

The same state variables can be derived for the fused tracks from homogeneous sensors in the first step, as passive sonar is land-based and its known position is used to calculate the bearing of these fused tracks. Therefore, the variable Δ is given by (11) :

$$\Delta = \hat{X}_{sonar} - \hat{X}_{radar} = \begin{bmatrix} \hat{\theta} & \hat{\dot{\theta}} \end{bmatrix}^T_{sonar} - \begin{bmatrix} \hat{\theta} & \hat{\dot{\theta}} \end{bmatrix}^T_{radar} \quad (11)$$

F. Track-to-track fusion (TF)

Once the tracks corresponding to the same target have been associated by T2TA, it is necessary to compute the best estimate for the fused track through the track fusion (TF) process. To achieve this, covariance matrices are once again required - some of which, such as those from AIS equipment, are not accessible. The formal approach to this calculation would be given by (12) and (13) outlined in [8].

$$\hat{X} = \hat{X}_1 + [P_1 - P_{12}] \cdot [P_1 + P_2 - P_{12} - P_{21}]^{-1} \cdot [\hat{X}_2 - \hat{X}_1] \quad (12)$$

$$P = P_1 - [P_1 - P_{12} \cdot [P_1 + P_2 - P_{12} - P_{21}] \cdot [P_1 - P_{12}] \quad (13)$$

To calculate the best estimate for the fused track without using the aforementioned formulas, two models were adopted and compared.

The first model applies a weighted average of positions provided by radar and AIS tracks. The weighting factors are determined based on the error estimates of the individual tracks (radar and AIS). The mean error of the radar track's

position estimate relative to the actual target position, as well as the mean error of the AIS estimate relative to the target position, was measured. The weighting factors were then assigned according to the relative estimation accuracy between the radar and AIS. For radar tracks, estimates are derived from the plot fusion process, where the covariance matrix is known. For AIS tracks, the accuracy information provided by the equipment manufacturer is used.

The second model is the same IMM filter used for radar plot fusion but now considering as inputs ("measurements") not the radar plots but the tracks estimations from PF model and from AIS equipment. In this second application of IMM model, process error and "measurement" errors were adjusted downwards, considering that both inputs are already providing filtered information.

IV. Experiments

A. Methodology

To validate the proposed model, an operational environment simulator was developed. This simulator generates contacts of interest in balanced quantities, representative of real-world scenarios. It also models the dynamic behavior of targets and sensors, incorporating their respective process and measurement noises.

The simulator provides an interface for creating targets and assigning them linear, accelerated, or circular motion. It also allows for the parametrization of Gaussian distributions for both target process errors and sensor measurement errors, creation of static or dynamic sensors, control over sensor operation (start/stop), and the generation of radar clutter. Within this simulated environment, the proposed models for track-to-spot association, T2TA, and track fusion were evaluated.

To assess the efficacy of the track-to-spot association model, high-density scenarios were generated to compare the performance of T2TA with and without the track-to-spot pre-processing step.

To evaluate T2TA performance, dense scenarios were created in which 50% of the targets were equipped with AIS transponders while 50% were not. Furthermore, 50% of the targets were within radar or passive sonar coverage areas, while the other 50% were outside these detection ranges. This set of tracks, with known correspondence (or non-correspondence) to the simulated targets, was then used to compute performance metrics including Precision, Accuracy, F1-Score, and Recall. These metrics were measured separately for tracks originating from homogeneous and heterogeneous sensor suites.

Finally, a table presenting the Root Mean Squared Error (RMSE) between the state vectors of the fused tracks and the simulated target states is included to evaluate the overall performance of the two proposed methodology for Track Fusion (TF) .

B. Validation scenarios

To validate the ensemble models illustrated in Figure 1, some system performance tests were conducted with a variable amount of targets: 100 (tiny), 200 (small), 500 (medium), and 1000 (large). The initial positions of the targets were uniformly distributed within a test area (10 NM x 10 NM). The targets moved randomly inside the test area in either linear or circular motion, with speeds ranging from 0 to 15 knots.

In each test, 50% of the targets were equipped with AIS. The radars have an intersection area, and the union of their coverage areas were equivalent to 50% of the test area. Thus, approximately 50% of the targets were not resulted in radar tracks (by PF model).

The same 50% rate applies to the passive sonar coverage area and the test area. The uniform distribution of targets within the coverage area, the proportional size of the radar coverage areas relatively to the test area, and the number of AIS tracks generated by the targets were chosen to create proportional quantities of tracks (created by PF model) to be associated and not to be associated (by T2TA model) and fused (by TF model). This strategy creates similar number of targets with only AIS tracks, targets with only radar tracks, targets with both AIS and radar tracks, targets with fused tracks from AIS tracks and radar tracks, targets with bearing only tracks (from sonar) and targets outside sensor's coverage areas with none tracks.

There are two radars with rotation frequencies of 15 RPM and 20 RPM, a range of 18 nautical miles and measurement errors with Gaussian distributions $N(0m, 10m)$ and $N(0°, 0.3°)$, for range and bearing respectively.

There is one AIS equipment with a position accuracy (provided by the manufacturer) of 10 m. Considering this 10 m as 3 x standard deviation (99.7% of estimations), this implies in an introduction of a Gaussian distribution error $N(0m, 3.33m)$ in the AIS track's position.

There is also a passive sonar with a bearing detection accuracy of 1 degree (as provided by the manufacturer), which results in the introduction of a Gaussian distribution error $N(0°, 1°)$ in the detected bearing of the passive sonar tracks.

C. Results

To validate the adopted models and the system as a whole, a simulation environment was built in which vessels (called targets) can be created, sensors (Radar, AIS, and passive sonar), and the models to be tested. The models are for plot fusion (PF), track-to-spot association (T2SA), track-to-track association of sensors that provide position like radar and AIS (T2TP), track fusion (TF) of the tracks associated by T2TP, and track-to-track association of fused tracks with tracks from passive sonar that provides only bearing (T2TB).

Figure 2 illustrates the simulation environment with three sensors, two radars (the blue and red circles centered on points 1 and 2 represent the coverage areas) and a passive sonar (the orange circle represents the coverage area). There are three targets represented by solid circle points. One target is only inside radar 2 area, another target is only inside the AIS area and the third one is inside all the three sensor coverage areas. The target that is only inside an AIS area generates an AIS track (green triangle). The target within a single radar coverage

area has a radar track (white square) generated by the PF model. The target that is inside all the coverage areas, has a radar track (white square) generated by the PF model and a fused track (red bow tie) generated by TF model. There are also two target's bearing lines (red dot lines) generated by the T2TB model that associated the bearing tracks that are inside the passive sonar area with the fused track generated by TF model and the AIS. In this image, the process error used in the PF model has a distribution $N(0, 0.1)$, the measurement errors have a distribution $N(0, 0.2)$ for radar 1, $N(0, 0.3)$ for radar 2, and $N(0, 3.3)$ for AIS. The range is 6 NM for the radars and for the passive sonar. The rotation frequency is 15 RPM (radar 1) and 20 RPM (radar 2). A uniform linear motion was assigned to the targets with a speed of 8 knots and with variation of motion as described in the section III-B.
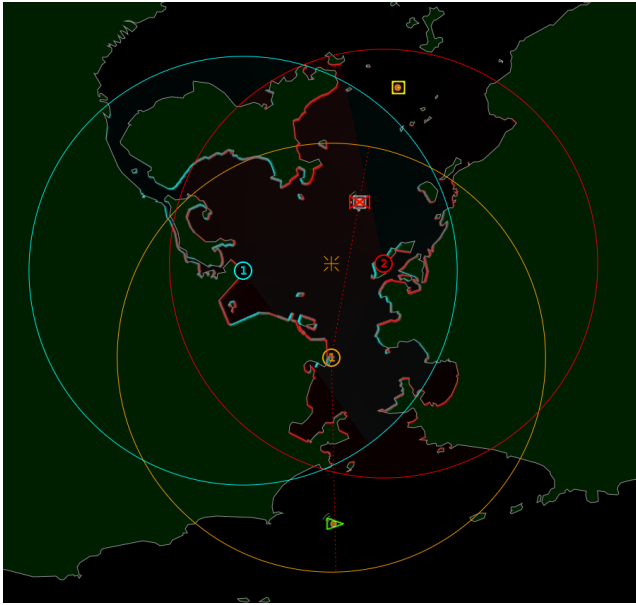

Fig. 2. Simulator environment screen aspect around the Guanabara Bay (Brazil).

The PF model is the IMM ([8]–[10]) and is not analyzed in this work since it is well known. To evaluate the T2SA model, which directly impacts the T2TP time, the relative execution times of T2TP for different numbers of clusters (spots) are presented in Figure 3. For each number of targets, the execution time of T2TP without applying T2SA (i.e., with only one cluster containing all tracks) as being 100%.

Each of the four groups of columns in Figure 3 was generated with a different number of targets (tiny, small, medium, or large). Each group shows times for 1, 5, 10, and 20 spots. The execution time of T2TP without applying T2SA (i.e., with only one cluster containing all tracks) represents 100%. The T2TP processing was done without any parallelization resources, such as the use of threads or separate processors, which theoretically could be employed to increase performance since the T2TP processes are independent between spots. The presented percentages times were obtained by sequentially processing the T2TP for each spot. The values in the graphs represent

the averages of 10 executions of the models under the same conditions of target density and number of spots.

As can be observed in Figure 3, the time spent by T2TP decreases significantly with an increasing number of spots at a similar rate for all amount of targets. For the highest amount of groups tested (20 groups) and the highest number of targets (1000 targets), the time spent was around 5% of time without T2SA, representing a strong argument for using it in real-world dense scenarios, where may be necessary to deal with 100,000 to 400,000 vessels.
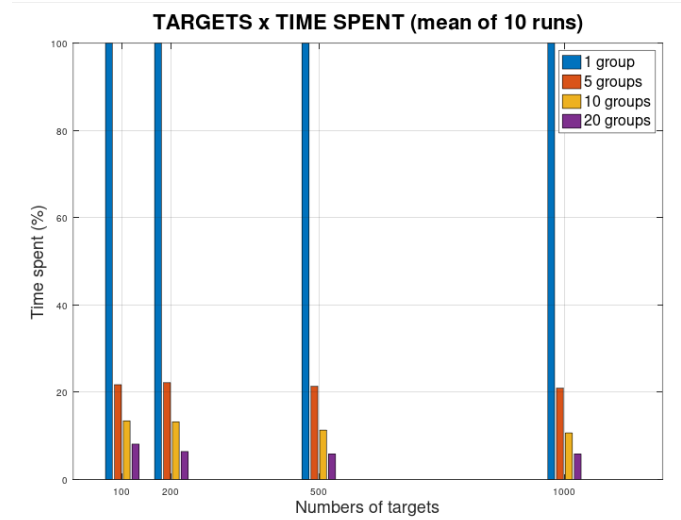

Fig. 3. Percentage of time spent by T2TP for different numbers of groups (1, 5, 10, or 20 spots) relative to the time required for a single group (100%), considering varying numbers of targets (100, 200, 500, and 1000).

The evaluation tests for the T2TP model were conducted considering the metrics of Precision, Accuracy, F1-Score and recall. These values were calculated within a continuous 2-minute time window of T2TP model execution for the maximum density scenario with 1000 targets. The association decisions between AIS and Radar tracks were computed within this period. Since the number of targets that generate tracks is known, it is possible to evaluate whether the association was correct, because tracks correspond to the same target (True Positive - TP), or incorrect, because tracks do not correspond to the same target (False Positive - FP). Similarly, it is possible to evaluate whether each non-association of tracks is correct, because tracks do not correspond to the same target (True Negative - TN), or incorrect, because tracks correspond to the same target (False Negative - FN).

The Table I displays the confusion matrix of the T2TP and the metrics associated with this matrix, considering the scenarios described in the section IV-B.

The matrix presents much larger numbers of associations and non-associations than the number of targets because they represent the average of 10 tests, with each test being a set of T2TP model hypothesis investigations taken every 3 seconds over 2 continuous minutes with 1000 moving targets (totalizing 40 sequential investigations).

TABLE I
T2TP CONFUSION MATRIX (RADAR TRACKS AND AIS TRACKS
ASSOCIATION).

Condition	Predicted	
	TP	**FN**
Actual	7980	38
	38	6884
	FP	**TN**

TABLE II
T2TP METRICS OBTAINED (RADAR TRACKS AND AIS TRACKS
ASSOCIATION).

Accuracy	Precision	Recall	F1-Score
99.48%	99.53%	99.53%	99.53%

TABLE V
OVERALL SYSTEM PERFORMANCE. THE RESULTS OF ROOT MEAN
SQUARED ERROR (RMSE) ARE THE AVERAGE OF 30 RUNS FOR UNIFORM
RECTILINEAR MOTION.

Weighted Average			IMM Filter		
Position	Course	Speed	Position	Course	Speed
10.56	5.38	0.93	5.79	4.37	0.74

TABLE VI
OVERALL SYSTEM PERFORMANCE. THE RESULTS OF ROOT MEAN
SQUARED ERROR (RMSE) ARE THE AVERAGE OF 30 RUNS FOR UNIFORM
CIRCULAR MOTION.

Weighted Average			IMM Filter		
Position	Course	Speed	Position	Course	Speed
11.06	1.09	0.97	6.11	3.47	0.15

As can be seen in Table I and II the T2TP model associates AIS and radar tracks with good performance for all the proposed metrics.

The Table III displays the confusion matrix of the T2TB method and the metrics associated with this matrix, considering a scenario with 100 targets (that is a number of targets compatible with sonar capability) in rectilinear motion. Around 50% of the targets generated passive sonar tracks because they were within the sonar coverage area. The T2TB method sought to associate passive sonar tracks with both tracks fused by the TF model (departing from tracks associated by the T2TP model) and tracks only from AIS, as long as the tracks of both types were within the sonar coverage area.

The numbers presented in Table III are much larger than the number of targets because they represent the average of 10 tests, with each test being a set of T2TB model hypothesis investigations taken each 3 seconds over 2 continuous minutes (totalizing 40 sequential investigations).

The results show that T2TB model associates bearing only tracks with position tracks less efficiently than T2TP model presented in Table II, but still with a reliability that is useful to increase the situational awareness in multi-sensor environments.

To test the performance of the last step, the TF model, the Root Mean Squared Error (RMSE) of the differences between the fused track position (from the TF model) and the actual target position (that is known by simulator), was calculated.

The Tables VI and V present RMSE for two test scenarios, one with 1000 targets in Uniform Rectilinear Motion (URM) and another with 1000 targets in Uniform Circular Motion (UCM). The units adopted for representation were: speed in knots (kn), course in degrees (º) and position in meters (m).

Both methods of TF (Weighted Average and IMM) were used to fuse AIS tracks and radar tracks.

In order to define the weights to be used in the Weighted Average method, the precision of radar and AIS tracks relative to the actual target position was first investigated.

First, the RMSE of the difference between the position estimate provided by the radar track and the actual position of the target was calculated. The same was done with the difference between the AIS track and the target.

From the results, we found that the mean error of the radar track position was four times greater than the error of the AIS position. Thus, the weighting factors for calculating the Weighted Average method were set to 0.2 for the radar track and 0.8 for the AIS track.

The results for both methods were presented separately, aiming to assess the robustness of each method for targets with and without acceleration.

From the Tables V and VI results, it is possible to conclude that IMM and Weighted Average have similar performance, with IMM being better in position and velocity estimations while Weighted Average being slightly better in course estimation for UCM motion.

Since the IMM method is significantly more computationally intensive than the Weighted Average method, the latter should be more suitable for dense scenarios involving hundreds of thousands of vessels to be fused.

Finally, since the overall objective of this work is to test a stack of models used in real-world surveillance systems and considering that these systems operate under varying sea conditions, a test was conducted to investigate the robustness of the T2TP and TF models under favorable, median, and unfavorable conditions. The sea state and its influence on vessel motion depend on several factors, most notably the vessel's mass. Considering the Beaufort scale of sea states, from 0

TABLE III
T2TB CONFUSION MATRIX FOR THE ASSOCIATION OF PASSIVE SONAR
TRACKS WITH POSITION TRACKS (RADAR + AIS).

Condition	Predicted	
	TP	**FN**
Actual	1747	79
	214	112
	FP	**TN**

TABLE IV
T2TP METRICS OBTAINED (RADAR TRACKS AND AIS TRACKS
ASSOCIATION).

Accuracy	Precision	Recall	F1-Score
86.41%	89.09%	95.67%	92.26%

TABLE VII
EVALUATION METRICS. 1000 TARGETS, UNIFORM RECTILINEAR MOTION
WITH THREE SEA CONDITIONS (σ). THE RESULTS ARE THE AVERAGE OF
10 RUNS FOR EACH MOTION AND σ.

Standard Deviation Process Noise	Weighted Average			IMM Filter		
	Position	Course	Speed	Position	Course	Speed
σ_{weak}	10.00	5.24	0.93	5.56	4.28	0.73
σ_{medium}	10.91	5.19	0.92	6.09	4.23	0.74
σ_{strong}	10.77	5.72	0.93	5.73	4.60	0.74

TABLE VIII
EVALUATION METRICS. 1000 TARGETS, CIRCULAR RECTILINEAR
MOTION WITH THREE SEA CONDITIONS (σ). THE RESULTS ARE THE
AVERAGE OF 10 RUNS FOR EACH MOTION AND σ.

Standard Deviation Process Noise	Weighted Average			IMM Filter		
	Position	Course	Speed	Position	Course	Speed
σ_{weak}	10.36	1.09	0.98	7.27	2.85	0.15
σ_{medium}	10.53	1.08	0.96	5.00	3.20	0.14
σ_{strong}	12.28	1.09	0.96	6.07	4.35	0.16

(calm sea) to 12 (hurricane), a speedboat may experience an interference in its motion in a sea state 1 which can be similar to that of an aircraft carrier in a sea state 5 (rough sea). Thus, due to the wide variety of vessels and the variations in sea state interference according to the mass and geometry of the vessels, three interference scenarios were created, varying the intensity of process noise in target motion: weak, medium, and strong. In these scenarios, standard deviations were used for the process noise that affects target motion as described in the section III-B, with values of 1/10 (weak), 2/10 (medium), and 3/10 (strong) of the target's maximum speed or target's acceleration, depending on the motion model of each target. The results shown in Tables VII and VIII demonstrate that both methods used for TF are robust against the variations of sea conditions, as long as they provide minimal changes in estimates with increasing process error.

V. CONCLUSION AND FUTURE WORKS

This paper proposes an evaluation of a stack of models used for track-to-track association (T2TA) and track fusion (TF) in real-world maritime surveillance systems. Some of these models are classical, while others incorporate innovations aimed at reducing computational costs and enhancing the system's feasibility in dense vessel and sensor scenarios. The stack begins with the classical IMM method, which is used to generate radar tracks via plot fusion (PF). Subsequently, a context-based method (T2SA) is applied to group vessels into geographic spots before performing T2TA. This grouping significantly reduces the processing time of T2TA by more than one order of magnitude in dense scenarios.

Furthermore, a novel application of IMM is proposed for the fusion of AIS and radar tracks, where radar and AIS estimations are treated as measurements. Variances are dynamically calculated within a time window, as the estimation covariances for AIS are unknown. The results of this IMM-based TF approach are compared with a simpler method that weights the estimates from both sensors based on their measured precision in a simulated environment.

Additionally, a track-to-track bearing association method (T2TB) is tested between bearing-only tracks (sonar) and position tracks (radar), using a state vector appropriate for bearing-only tracks. Although this method shows lower performance than T2TP, it is valuable for enhancing situational awareness in dense scenarios with multiple sensor types.

Finally, the complete system, composed of the stack of models, is tested under various simulated sea conditions. These conditions are modeled by the noise introduced in target simulation, creating real-world scenarios with challenging maritime environments. The results demonstrate that the system is robust to variations in sea conditions.

For future research, given the comparatively inferior results yielded by T2TB relative to T2TP, a potential avenue for improvement involves incorporating the noise power spectrum of the sonar track and the vessel's class metadata provided by the AIS. As long as the vessel's class can be associated with an acoustic signature, this integration could enhance the performance of the T2TB method.

REFERENCES

[1] T. Yuan, Y. Bar-Shalom, X. Tian, "Heterogeneous track-to-track fusion," Journal of Advances in Information Fusion, vol. 6, n° 2, pp. 133-148, December 2011.

[2] Y. Bar-Shalom, "On hierarchical tracking for the real world." IEEE Transactions on Aerospace and Electronic Systems, vol. 42, n° 3, pp. 846–850, July 2006.

[3] Y. Bar-Shalom, P. K. Willet, X. Tian, "Tracking and data fusion - a handbook of algorithms," 1st Ed., Storrs, CT, YBS Publishing, 2011.

[4] X. Tian, Y. Bar-Shalom, "Track-to-track fusion configurations and association in a sliding window," Journal of Advances in Information Fusion, vol. 4, n° 2, pp. 146-164, December 2009.

[5] B. L. Scala, A. Farina, "Effects of cross-covariance and resolution on track association," In Proceedings of the Third International Conference on Information Fusion, vol. 2, Paris, France, July 2000.

[6] C. Allig, G, Wanielik. "Heterogeneous track-to-track fusion using equivalent measurement and unscented transform." In 21st International Conference on Information Fusion, pp. 1952–1958. 2018.

[7] R. C. Gertosio, G. Gaonach, E. Beyna, L. Martin, A. Meyrat, "Passive Sonar Ranging and Range-Doppler-Bearing Target Motion Analysis," 27th International Conference on Information Fusion, FUSION 2024, Venice, Italy, July 8-11, 2024.

[8] Y. Bar-Shalom, X. R. Li, "Multitarget-Multisensor Tracking: Principles and Techniques." Storrs, CT, YBS Publishing, 1995.

[9] E. Mazor, A. Averbuch, Y. Bar-Shalom, J. Dayan, "Interacting multiple model methods in target tracking: a survey," IEEE Transactions on Aerospace and Electronic Systems, vol. 34, n° 1, pp. 103–123, January 1998.

[10] Y. Bar-Shalom, X. R. Li, T. Kirubarajan, "Estimation with applications to tracking and navigation - theory, algorithms and software," 1st Ed., New York, John Wiley & Sons Ltd, pp.57–59, 2001.

[11] A. D. Waite, "Sonar for practising engineers," 3rd Ed. Chichester, John Wiley & Sons Ltd, 2002, ch. 8, pp. 125—160.

[12] M. E. Liggins, D. L. Hall, J. Llinas, "Handbook of Multisensor data fusion theory and practice," 2nd Ed., Boca Raton, CRC Press, 2009.

[13] X. Tian, T. Yuan, Y. Bar-shalom. "Track-to-Track fusion in linear and non-linear systems." In Advances in Estimation, Navigation, pp. 21—41. 2015.