Association of SAR Measurements in Coastal Regions using Existing Tracks of Marine Vessels

Alexander M. Sjøberg* , Brita H. H. Gade*, Carina Vooren*, Morten Kloster*
*Norwegian Defence Research Establishment (FFI), Kjeller, Norway
Alexander-Meyer.Sjoberg@ffi.no, Brita.Gade@ffi.no, Carina-Norberg.Vooren@ffi.no, Morten.Kloster@ffi.no

***Abstract*—This paper considers a measurement-to-track association (MTA) problem of marine surface vessels in coastal regions. In particular, we consider a case where a set of non-cooperative measurements is associated with a set of existing tracks. We perform a case study involving marine vessels situated in a coastal environment dominated by islands, fjords and peninsulas. Euclidean distances, or more generally, Mahalanobis distances, between measurements and predicted positions may be an attractive approach due to its simplicity of implementation and low computational complexity. However, such metrics may not be feasible for association in coastal regions, as opposed to open waters. We propose an algorithm where numerical methods are used to produce a probability density map for evaluation of a set of non-cooperative measurements, obtained at a time between two positions belonging to a particular track. The proposed associator provides a list of hypotheses intended for a multiple hypothesis tracker (MHT), where each hypothesis is evaluated according to the deviation of expected arrival time.**

***Index Terms*—Data association, FMM, WGS 84, NRD, SAR**

I. Introduction

The ocean covers over two thirds of the Earth's surface, and fishing, transportation, and other maritime activities play a significant role in the global economy. Consequently, the maritime situational awareness must be high to ensure that marine vessels act in agreement with international conventions and national laws. Since 2002, the International Maritime Organization (IMO) has requested large vessels such as fishing, passenger and cargo ships, to be equipped with an AIS (automatic identification system) transponder onboard [12]. The intention was to improve the maritime safety, efficiency of navigation, and protection of the marine environment. AIS and synthetic aperture radars (SAR) are examples of tools which can be used for surveillance of the maritime domain [2]. In the following, a brief overview of relevant sensor technologies used in maritime surveillance will be discussed, followed by the problem statement motivating this paper.

A. Sensor Overview

1) Cooperative Reporting Systems (CRS): AIS, which is described in [5], is an example of a CRS. Each vessel actively broadcasts information to be detected by an AIS receiver, such as its maritime mobile service identity (MMSI), speed over ground (SOG), course over ground (COG), and additional information such as call sign and type of ship. Since each measurement in the AIS data is assigned a unique MMSI, the association problem can be resolved immediately, either with an existing track b_{MMSI}, or by creating a new track. The structure of an AIS measurement is provided in Table I.

TABLE I
AIS DATA OBTAINED AT TIME k. THE MMSIS CAN BE USED TO ASSOCIATE A MEASUREMENT WITH A TRACK.

MMSI	Lat.	Long.	COG	SOG
MMSI_1	φ_k^1	λ_k^1	ψ_k^1	v_k^1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
MMSI_n	φ_k^n	λ_k^n	ψ_k^n	v_k^n

AIS surveillance units are typically land-based or mounted on satellites, which allows surveillance over local waters as well as large areas in the ocean. Land-based AIS has a high sampling rate, but due to its limited field of view (FOV), this configuration is generally suited in coastal areas. Satellite-borne AIS can surveil large areas in the ocean, but has a much lower sampling rate due to the orbital nature of satellites. The ship tracking performance of the current satellite-borne AIS sensors operated by Norway was studied in [11].

2) Non-Cooperative Reporting Systems (NCRS): Three NCRS will be discussed in this section, namely SAR, electro-optical (EO) sensors, and navigation radar detector (NRD).

SAR is an active sensor which provides the output in the form of an image. SAR images are generated by emitting radar waves onto a surface, and the output is based on the reflected portion of the emitted signal [13]. The emitted radar waves are not absorbed by clouds, and the sensor up-time does not depend on the weather in the surveilled area. Finally, the SAR images are processed by detection algorithms, and the output is a set of measured positions.

EO is a passive, optical sensor, and consequently is weather dependent, as it requires clear conditions [7]. Image processing techniques are used to convert EO measurements to a set of measured positions.

SAR and EO sensors are typically satellite-borne in the context of maritime surveillance. This ensures that the sensors cover large areas from several satellite passes, at the cost of low sampling rates. Measurements from SAR and EO sensors are structured as in Table II. The assigned index of a measured position is arbitrary, and an association method is required to relate a position measurement to an existing track.

A third type of sensor belonging to NCRS is the NRD. The NRD is a passive sensor detecting the pulses from active navigation radars on marine vessels. While the navigation

TABLE II
DATA FROM SAR OR EO SENSORS RETURN POSITIONS AT TIME t_k.

Index	Lat.	Long.
1	φ_k^1	λ_k^1
$\vdots$	$\vdots$	$\vdots$
n	φ_k^n	λ_k^n

TABLE III
NRD SENSORS MEASURE POSITIONS AT TIME t_k WHICH CAN BE ASSOCIATED WITH OLDER MEASUREMENTS OF THE SAME VESSEL TO CREATE TRACKS.

Track	Lat.	Long.
NRD_1	φ_k^1	λ_k^1
$\vdots$	$\vdots$	$\vdots$
NRD_n	φ_k^n	λ_k^n

radar can be turned off, it is viewed as a necessity for all ships navigating in coastal waters to avoid colliding with other vessels or land. The NorSat-3 satellite, which was launched in 2021 [3], is a microsatellite intended for maritime traffic monitoring and is equipped with both AIS and an experimental NRD.

NRD measurements are distinct in comparison to SAR/EO: As different navigation radars have different properties, it is possible to distinguish emissions from different radars. Therefore, each measurement can be associated with a track b_{NRD}. In this paper, we omit such detailed radar properties, and an NRD measurement only contain the position of a vessel, as shown in Table III.

B. Problem Statement

AIS only functions as intended during the up-time of the AIS transponder located on each vessel. As full up-time cannot be guaranteed, AIS may occasionally fail as a reporting system for some period of time, regardless of whether satellite-borne or land-based installations are utilized. In such cases, other satellite-borne sensor technologies may be utilized to measure ship positions.

Consider two measurements from a sensor able to correctly associate a measurement to a track, such as AIS or NRD, obtained at the time instances t_1 and t_3. If SAR or EO data was obtained at the intermediate time t_2, which measured position does most likely belong to the correct track? The need for a robust associator is visualized in Figure 1 where the position measurements from three time instances are plotted.

It is noted that similar problem statements have been studied previously. In [10], the problem was one-sided in the sense that AIS-measurements obtained at t_1 were associated with SAR-measurements obtained at t_2. A similar problem was addressed in [6], where the time instances of AIS- and SAR-measurements were near-contemporaneous.

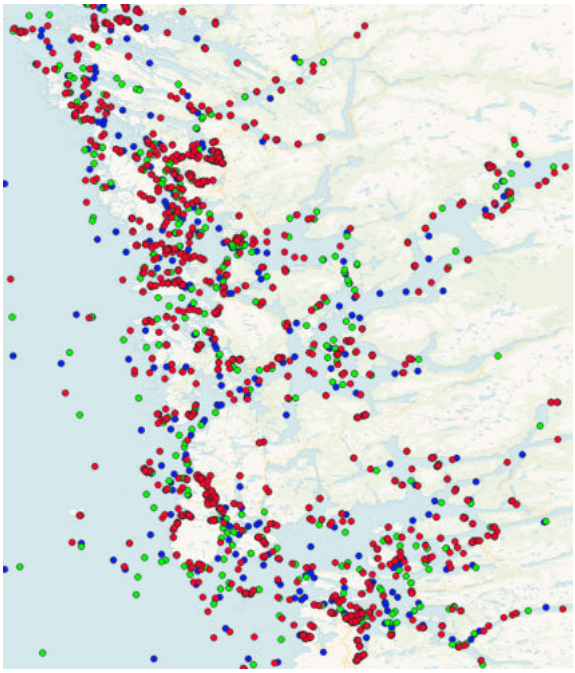

Fig. 1. The red, green and blue dots are position measurements obtained at time instances t_1, t_2 and t_3, respectively.

C. Paper Overview

The paper is organized as follows: Section II provides the preliminaries of this paper. Section III describes the association methods to be compared. The experimental part and the results is presented in Section IV, and Section V concludes the paper with a discussion.

II. PRELIMINARIES

Let $\boldsymbol{x} = \begin{bmatrix}\lambda & \varphi\end{bmatrix}^{\mathsf{T}}$ denote a position in a chart, where λ and φ are the components of the position, denoting longitude and latitude, respectively. A map segment is denoted as $\mathcal{M} \in [\lambda_{\min}, \lambda_{\max}] \times [\varphi_{\min}, \varphi_{\max}]$, i.e., a bounded, rectangular area in the longitude-latitude chart.

A. WGS 84 and Distance Metrics

The oblate WGS 84 spheroid is used in this paper as an approximation of the shape of the earth at sea level. While a sphere is fully described by a single parameter, i.e., the radius r, a spheroid requires two parameters, which will be denoted r_a and r_b. The parameters r_a and r_b define the length of the spheroid's semi-major and semi-minor axis, respectively. In the case of WGS 84, the parameters are given as $r_a = 6\,378\,137\,\text{m}$ and $r_b = 6\,356\,752.3\,\text{m}$, i.e., the eccentricity is $e = \sqrt{1-(r_b/r_a)^2} = 0.0818$. The metric tensor [4]

$$\boldsymbol{G} = \begin{bmatrix} g_{11} & 0 \\ 0 & g_{22} \end{bmatrix}. \tag{1}$$

allows computation of distances between positions on WGS 84 expressed in the longitude-latitude chart. The non-zero components in (1) are given as $g_{11} = r_a^2 \cos^2(\varphi)/d(\varphi)$ and

$g_{22} = r_a^2(1-e^2)^2/d^3(\varphi)$, where $d(\varphi) = 1 - e^2\sin^2(\varphi)$. It is noted that g_{11} and g_{22} are functions depending on the latitude, and not the longitude.

Let $\boldsymbol{x} = \begin{bmatrix}\lambda & \varphi\end{bmatrix}^\mathsf{T}$ and $\boldsymbol{y} = \begin{bmatrix}\lambda + d\lambda & \varphi + d\varphi\end{bmatrix}^\mathsf{T}$ denote two positions on WGS 84 where $d\lambda$ and $d\varphi$ are infinitesimal displacements, and let $d\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{x}$. The squared distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is given as

$$\begin{aligned} ds^2 =& d\boldsymbol{x}^\mathsf{T}\boldsymbol{G}d\boldsymbol{x} \\ =& g_{11}d\lambda^2 + g_{22}d\varphi^2 \end{aligned} \tag{2}$$

By utilizing the form in (2), we can approximate the geodesic distance between two points on WGS 84 as

$$d_\mathrm{E}(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{g_{11}\Delta\lambda^2 + g_{22}\Delta\varphi^2} \geq 0 \tag{3}$$

where $\Delta\lambda$ and $\Delta\varphi$ represent the angular displacements between two positions on WGS 84.

B. Fast Marching Method

The fast marching method (FMM) is an algorithm that was introduced in [9]. It was developed to solve the Eikonal equation, which is a partial differential equation that describes a wave front propagating in an admissible region. It can be formulated as

$$\|\nabla\tau(\boldsymbol{x})\|v(\boldsymbol{x}) = 1, \quad \boldsymbol{x} \in \mathcal{A} \subset \mathcal{M} \tag{4}$$

where the function $\tau : \mathbb{R}^2 \to \mathbb{R}$ denotes the time at which the propagating wave front crosses the point $\boldsymbol{x}$, and v denotes the speed of the wave front in the admissible region. In this article, the solution τ will be used to compute propability maps defined in $\mathcal{A}$. These probability maps will be used to compute the likelihood that a measured position belongs to a particular track.

The FMM requires a discretized representation of $\mathcal{M}$, and the subset $\mathcal{A}$ is the region where the wave front can propagate. The computational cost of FMM is $\mathcal{O}(n\log n)$ [8] where n is the number of pixels constituting the map segment to be evaluated. The algorithm provides the solution as a set of values associated with each grid point, i.e., the arrival time of the wave front. The function τ will in the following be referred to as the time-of-arrival (TOA) map. A TOA map evaluated from an initial point $\boldsymbol{x}' \in \mathcal{A}$ and with constant propagation speed $v > 0$, will be denoted $\tau_{\boldsymbol{x}',v}$, where $\tau_{\boldsymbol{x}',v}(\boldsymbol{x}') = 0$. In the special case where $v(\boldsymbol{x}) = 1$, the solution τ is a function assigning a distance value to a point $\boldsymbol{x}$, i.e., a distance map. A distance map will be denoted $d_{\boldsymbol{x}'}$, where $d_{\boldsymbol{x}'}(\boldsymbol{x}') = 0$. For further details regarding the implementation of FMM, we refer to [1].

C. Expected Arrival Time

Consider a TOA map $\tau_{\boldsymbol{x}',v}(\boldsymbol{x})$. In accordance with [10], each point $\boldsymbol{x} \in \mathcal{A}$ can be evaluated by the function $p : \mathbb{R}^2 \to \mathbb{R}$, which in this paper is given as

$$p(\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}\left(\frac{\tau_{\boldsymbol{x}',v}(\boldsymbol{x}) - \mu_T}{\sigma_T}\right)^2\right). \tag{5}$$

The expected arrival time is denoted μ_T, and the standard deviation $\sigma_T > 0$ assigns the uncertainty of the arrival time.

III. Methods

In this section, we will present two associators whose outputs are sets of weigths, denoted $\mathcal{W} = \{w^{(1)}, w^{(2)}, \ldots, w^{(n)}\}$ where the entries are sorted in descending order. The first associator sorts the weigths based on the evaluated Euclidean distance relative to a predicted position, while the second sorts the weigths based on the evaluated arrival time relative to the expected arrival time. A position measurement which will be evaluated with respect to a track, based on its computed weight, will be referred to as a hypothesis. Each associator will require a track $b = \{\boldsymbol{x}_1, \boldsymbol{x}_3\}$ where $\boldsymbol{x}_1 = \boldsymbol{x}(t_1)$ and $\boldsymbol{x}_3 = \boldsymbol{x}(t_3)$ are positions of the same vessel, and $\mathcal{X} = \{\boldsymbol{x}_2^1, \ldots, \boldsymbol{x}_2^n\}$ is a set of positions at time t_2. The N vessels in $\mathcal{X}$ with the highest weights will be deemed the most likely hypotheses to belong to b.

A. Distance Associator

The first associator to be considered ranks a set of hypotheses with respect to the Euclidean distances relative to a predicted state, in a descending order. The predicted state is the output of a constant velocity model (CVM), which assumes the kinematic differential equation

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{v} \\ \dot{\boldsymbol{v}} &= \boldsymbol{0}, \end{aligned} \tag{6}$$

where $\dot{\boldsymbol{x}} = \mathrm{d}\boldsymbol{x}/\mathrm{d}t$, and $\boldsymbol{v}$ is a constant velocity vector. The solution of (6) between two given positions is a trajectory which can be parameterized as

$$\bar{\boldsymbol{x}}(s) = (1-s)\boldsymbol{x}(0) + s\boldsymbol{x}(1), \tag{7}$$

where $s \in [0, 1]$, $\boldsymbol{x}(0)$ is the initial position, $\boldsymbol{x}(1)$ is the final position, and $\bar{\boldsymbol{x}}$ is the predicted position. The weight of a position $\boldsymbol{x}^i$ relative to $\bar{\boldsymbol{x}}$ when using the CVM, is computed as

$$w_\mathrm{CV}^{(i)} = \exp\left(-\frac{1}{2}\frac{d_\mathrm{E}^2(\boldsymbol{x}^i, \bar{\boldsymbol{x}})}{\sigma_\mathrm{CV}^2}\right), \tag{8}$$

where $\sigma_\mathrm{CV} > 0$ is the chosen standard deviation given in terms of a distance unit, which enables the use of validation gating. However, the use of Equation (8) may be prone to floating-point errors. To avoid such issues, the weights for the sorting algorithm are computed using the logarithm

$$\begin{aligned} \gamma_\mathrm{CV}^{(i)} =& \log\left(w_\mathrm{CV}^{(i)}\right) \\ =& -\frac{1}{2}\frac{d_\mathrm{E}^2(\boldsymbol{x}^i, \bar{\boldsymbol{x}})}{\sigma_\mathrm{CV}^2} \\ =& -\beta_\mathrm{CV} d_\mathrm{E}^2(\boldsymbol{x}^i, \bar{\boldsymbol{x}}) \end{aligned} \tag{9}$$

where $\beta_\mathrm{CV} = \frac{1}{2\sigma_\mathrm{CV}^2}$ is a scaling factor which does not change the order of the sorted weights. The weight to be normalized and used by the associator is then found by

$$w_\mathrm{CV}^{(i)} = \exp\left(\gamma_\mathrm{CV}^{(i)}\right). \tag{10}$$

The distance associator is described in Algorithm 1, where the notation $\boldsymbol{z}$ is used to emphasize that the evaluated points are measurements.

Algorithm 1 Constant velocity model

Require: $b = \{\boldsymbol{z}_1, \boldsymbol{z}_3\}$, $\mathcal{Z} = \{\boldsymbol{z}_2^1, \ldots, \boldsymbol{z}_2^n\}$
  CV path: $\bar{\boldsymbol{z}}_2 = (1-s)\boldsymbol{z}_1 + s\boldsymbol{z}_3$ where $s \in [0,1]$
  for each $\boldsymbol{z}_2 \in \mathcal{Z}$ **do**
    Compute $\gamma_{\text{CV}}^{(k)}$ as in Eq. (9) where $k = 1, \ldots, n$
  end for
  $\Gamma_{\text{CV}} = \{\gamma_{\text{CV}}^{(1)}, \gamma_{\text{CV}}^{(2)}, \ldots, \gamma_{\text{CV}}^{(n)}\}$ (sorted)
  $\mathcal{W}_{\text{CV}} = \exp(\Gamma_{\text{CV}})$ (element-wise)
  Normalize: $w_{\text{CV}}^{(k)} =: w_{\text{CV}}^{(k)} / \sum_{i=1}^{n} w_{\text{CV}}^{(i)}$
  return $\mathcal{W}_{\text{CV}} = \{w_{\text{CV}}^{(1)}, w_{\text{CV}}^{(2)}, \ldots, w_{\text{CV}}^{(n)}\}$

B. Arrival Time Associator

In this section, we propose an associator which requires one distance map and two TOA maps. If the speed is assumed to be constant in the admissible region, then these maps can be obtained by using the FMM two times. The distance map $d_{\boldsymbol{x}_1}(\boldsymbol{x})$ is obtained by using the FMM according to the special case described in Section II-B. Provided a track $b = \{\boldsymbol{x}_1, \boldsymbol{x}_3\}$ where $\boldsymbol{x}_1 = \boldsymbol{x}(t_1)$ and $\boldsymbol{x}_3 = \boldsymbol{x}(t_3)$, we can compute the distance $d_{1,3} = d_{\boldsymbol{x}_1}(\boldsymbol{x}_3)$, that is, the length of the shortest traversable path connecting the two end-points of track b [10]. Next, we compute the average speed $\bar{v} = \frac{d_{1,3}}{t_3 - t_1}$. Recall that function p in (5) is a dissimilarity function evaluating the difference between a given arrival time and an expected arrival time. In order to compute the desired dissimilarity function, we will require the following ToA maps:

$$\tau_{\boldsymbol{x}_1,\bar{v}}(\boldsymbol{x}) = \frac{d_{\boldsymbol{x}_1}(\boldsymbol{x})}{\bar{v}}, \quad \tau_{\boldsymbol{x}_3,\bar{v}}(\boldsymbol{x}) = \frac{d_{\boldsymbol{x}_3}(\boldsymbol{x})}{\bar{v}}$$

Function $\tau_{\boldsymbol{x}_1,\bar{v}}$ is the usual, forward propagated wave front, initiated from the position $\boldsymbol{x}_1$. This function is found by scaling the already available distance map $d_{\boldsymbol{x}_1}$ by the reciprocal of the computed average speed $\bar{v}$. Function $\tau_{\boldsymbol{x}_3,\bar{v}}$ is, on the other hand, propagated from $\boldsymbol{x}_3$, and may be interpreted as a wave front propagated backwards in time. This map is obtained by using the FMM a second time. These functions are then used to generate the following probability densities:

$$p_1(\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}\left(\frac{\tau_1(\boldsymbol{x}) - \mu_1}{\sigma_{\text{CS}}}\right)^2\right) \tag{11}$$

$$p_3(\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}\left(\frac{\tau_3(\boldsymbol{x}) - \mu_3}{\sigma_{\text{CS}}}\right)^2\right) \tag{12}$$

where the shorthand notation $\tau_1 = \tau_{\boldsymbol{x}_1,\bar{v}}$ and $\tau_3 = \tau_{\boldsymbol{x}_3,\bar{v}}$ is used, and where $\mu_1 = t_2 - t_1$, $\mu_3 = t_3 - t_2$, and $\sigma_{\text{CS}} > 0$ is the chosen standard deviation. Finally, we obtain the probability map

$$p_{2|1,3}(\boldsymbol{x}) \propto p_1(\boldsymbol{x}) p_3(\boldsymbol{x}), \tag{13}$$

and we refer to this procedure as the constant speed model (CSM). The weight of a position $\boldsymbol{x}^i$, while using the CSM, is computed as

$$w_{\text{CS}}^{(i)} = p_{2|1,3}(\boldsymbol{x}^i). \tag{14}$$

The floating-point issues addressed in Section III-B also apply to Equation (14). Therefore, the weights for the sorting algorithm are computed using the logarithm

$$\begin{aligned}\gamma_{\text{CS}}^{(i)} &= \log\left(w_{\text{CS}}^{(i)}\right) \\ &= -\frac{1}{2}\left(\frac{\tau_1(\boldsymbol{x}) - \mu_1}{\sigma_{\text{CS}}}\right)^2 - \frac{1}{2}\left(\frac{\tau_3(\boldsymbol{x}) - \mu_3}{\sigma_{\text{CS}}}\right)^2 \\ &= -\beta_{\text{CS}}\left((\tau_1(\boldsymbol{x}) - \mu_1)^2 + (\tau_3(\boldsymbol{x}) - \mu_3)^2\right),\end{aligned} \tag{15}$$

where $\beta_{\text{CS}} = \frac{1}{2\sigma_{\text{CS}}^2}$ is a scaling factor which does not change the order of the sorted weights. Finally, the weights to be sorted and used by the associatior are found according to

$$w_{\text{CS}}^{(i)} = \exp\left(\gamma_{\text{CS}}^{(i)}\right). \tag{16}$$

An example of how the densities p_1 and p_3 are formed in a coastal landscape, as well as the joint distribution in (13), are shown in Figure 2. The arrival time associator is summarized in Algorithm 2. Again, $\boldsymbol{z}$ is used to emphasize that the inputs are measurements.

Algorithm 2 Constant speed model

Require: $b = \{\boldsymbol{z}_1, \boldsymbol{z}_3\}$, $\mathcal{Z} = \{\boldsymbol{z}_2^1, \ldots, \boldsymbol{z}_2^n\}$
  Compute $d_{\boldsymbol{z}_1}(\boldsymbol{z})$ and $d_{\boldsymbol{z}_3}(\boldsymbol{z})$
  $d_{1,3} = d_{\boldsymbol{z}_1}(\boldsymbol{z}_3)$
  $\bar{v} = \frac{d_{1,3}}{t_{1,3}}$ where $t_{1,3} = t_3 - t_1$
  Compute $\tau_1(\boldsymbol{x})$ and $\tau_3(\boldsymbol{x})$
  for each $\boldsymbol{z}_2 \in \mathcal{Z}$ **do**
    Compute $\gamma_{\text{CS}}^{(k)}$ as in Eq. (15), where $k = 1, \ldots, n$
  end for
  $\Gamma_{\text{CS}} = \{\gamma_{\text{CS}}^{(1)}, \gamma_{\text{CS}}^{(2)}, \ldots, \gamma_{\text{CS}}^{(n)}\}$ (sorted)
  $\mathcal{W}_{\text{CS}} = \exp(\Gamma_{\text{CS}})$ (element-wise)
  Normalize: $w_{\text{CS}}^{(k)} =: w_{\text{CS}}^{(k)} / \sum_{i=1}^{n} w_{\text{CS}}^{(i)}$
  return $\mathcal{W}_{\text{CS}} = \{w_{\text{CS}}^{(1)}, w_{\text{CS}}^{(2)}, \ldots, w_{\text{CS}}^{(n)}\}$

C. Motivating example

We provide a motivating example on why the computationally demanding CSM should be considered over the CVM in certain cases. Figure 3 illustrates a case where $t_3 - t_1 = 120\,\text{min}$, and $t_2 - t_1 = 60\,\text{min}$. The true trajectory of the tracked vessel deviated significantly from a linear trajectory generated by the CVM. In open waters, ships will commonly follow a straight or nearly straight trajectory, while in coastal areas, ships will often be forced to follow less direct paths.

IV. Experiment

In the following, we will consider a case without AIS receivers, and where all the data was obtained from satellite-borne sensors. The structure of Algorithm 3 was used in the experimental part of this paper. For each evaluated track b, the linear distance between the measured positions $\boldsymbol{z}_1$ and $\boldsymbol{z}_3$ was computed. If the linear distance was less than or equal to a predefined threshold distance ε, then only CVM described in Algorithm 1 would be invoked. However, if the distance was greater than ε, then both Algorithm 1 and 2 would be used, and the results would be compared.

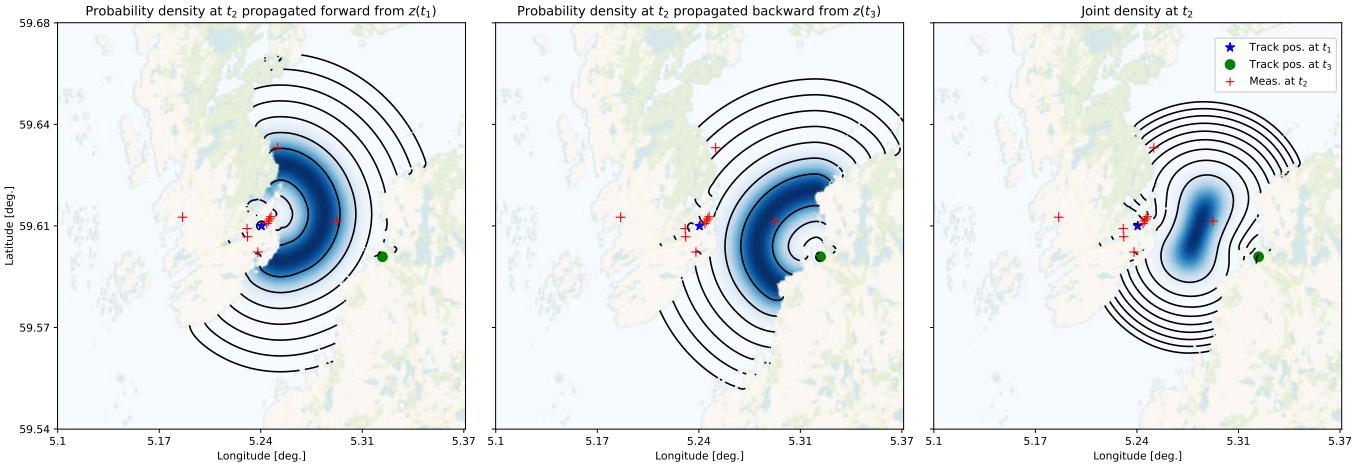

Fig. 2. The plots to the left and in the middle represents the forward and backward propagated probability distributions, given in (11) and (12), respectively. The joint distribution to the right is the product given in (13). Contours of each function are plotted in the logarithmic scale to indicate how the topography affects the distributions.

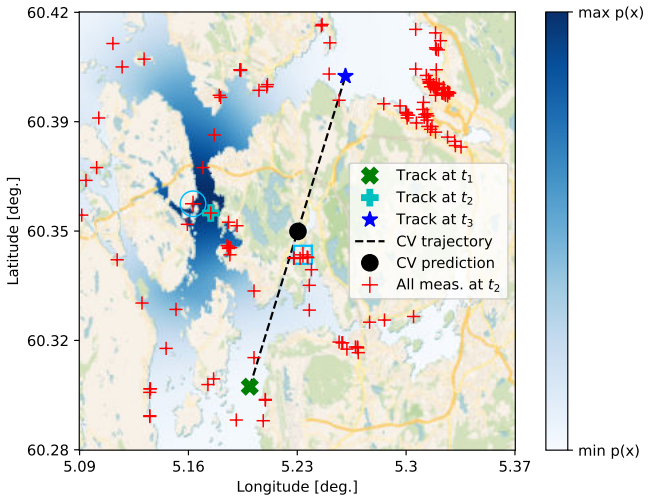

Fig. 3. The square marker indicates which measurement was assigned the highest probability by CVM, while the true measurement was ranked as the 19th most likely. The circle marker indicates the most probable hypothesis provided by the CSM, while the track at t_2 was ranked as the 2nd most likely. This illustrates how the CVM, for a sufficiently large Δt, may have to provide a large set of hypotheses to ensure that the true position measurement is included.

Algorithm 3 Association Routine for Experimental Data

Require: $b = \{\boldsymbol{z}_1, \boldsymbol{z}_3\}$, $\mathcal{Z} = \{\boldsymbol{z}_2^1, \ldots, \boldsymbol{z}_2^n\}$, $\varepsilon > 0$
  if $d_{\mathrm{E}}(\boldsymbol{z}_1, \boldsymbol{z}_3) \leq \varepsilon$ **then**
    Evaluate $\mathcal{Z}$ using Algorithm 1
  else
    Evaluate $\mathcal{Z}$ by comparing the outputs of Alg. 1 and 2
  end if

A. Data

The data was provided by Kongsberg Satellite Services, and consisted of AIS measurements of over 8000 vessels over a time span of 6 hours. Three sets of measurements were provided with the time stamps t_1, $t_2 = t_1 + \Delta t$ and $t_3 = t_1 + 2\Delta t$, for varying time steps Δt. The AIS data at t_1 and t_3 were simulated as NRD data, i.e., structured as in Table III. The AIS data at t_2 was simulated as SAR/EO data, i.e., structured as in Table II. In the following, the SAR/EO data will be referred to as SAR data for brevity. The use of AIS data for validation of the methods is considered to be a necessity in order to compare the findings with a ground truth.

All of the simulated NRD data, and therefore all the tracks, was obtained from the map segment $\mathcal{M}_{\mathrm{NRD}} = [4.5^\circ, 5.5^\circ] \times [59.75^\circ, 61.25^\circ]$. To account for the possiblity that the trajectories of tracks near the border of $\mathcal{M}_{\mathrm{NRD}}$ could leave the map segment before returning, the SAR measurements at t_2 were the vessels within the extended region

$$\begin{aligned}\mathcal{M}_{\mathrm{SAR}} =& [4.5^\circ - \Delta\lambda^W, 5.5^\circ + \Delta\lambda^E] \\ &\times [59.75^\circ - \Delta\varphi^S, 61.25^\circ + \Delta\varphi^N],\end{aligned}$$

where

$$\Delta\lambda = \frac{v_{\max}\Delta t}{\sqrt{g_{11}}}, \quad \Delta\varphi = \frac{v_{\max}\Delta t}{\sqrt{g_{22}}}.$$

The metric tensor components of $\Delta\varphi^N$ and $\Delta\varphi^S$ were evaluated at the latitudes 61.25° and 59.75° north, respectively, while $\Delta\lambda^W$ and $\Delta\lambda^E$ were evaluated at $\varphi = 60.5^\circ$ north.

B. Parameters

Each pixel in the rasterized map had a width and height equal to 30 m. Since the choice of standard deviation does not change the order of the sorted weights, the scaling factors β_{CV} and β_{CS} were set equal to 1. Algorithm 1 only assigned a non-zero weight to a measurement $\boldsymbol{z}$ if $d_{\mathrm{E}}(\boldsymbol{z}, \bar{\boldsymbol{z}}) \leq v_{\max}\Delta t$, where

the selected maximum speed $v_{\max} = 25\,\text{kn}$. Both distance maps used to obtain the TOA maps for Algorithm 2 were evaluated such that $d(\boldsymbol{x}) \leq v_{\max}\Delta t$. Consequently, a non-zero weight was assigned to a measurement $\boldsymbol{z}$ if, and only if, the position of a measurement was inside the overlapping area of the distance maps. Finally, the threshold parameter ε in Algorithm 3 was set equal to the distance traversed at a speed of $5\,\text{kn}$ over the evaluated time step Δt.

C. Results

The results of the experiment is given in Table IV, and shown in Figure 4. Table IV show how often the track at t_2 was assigned the highest weight, and how often it was included among the 5, 10, 15 and 20 most likely hypotheses, at various time steps Δt.

For the two smallest time steps, the performances of CVM and CSM were similar. For $\Delta t = 30\,\text{min}$, the CSM was able to associate correctly 6.7% more often than the CVM. However, by increasing the number of hypotheses, the CVM was performing slightly better, i.e., the true position was classified as a hypothesis by using fewer candidates than the CSM at the same time step. For $\Delta t \geq 40\,\text{min}$, the CSM performed better than the CVM when the number hypotheses was small, e.g., ≤ 10.

TABLE IV
COMPARISON OF TRACK DETECTION AT t_2 USING CVM AND CSM, SHORTENED TO CV AND CS, RESPECTFULLY, AND EVALUATED AT DIFFERENT TIME STEPS AND FOR VARYING NUMBERS OF HYPOTHESES.

	Top 1		Top 5		Top 10		Top 15		Top 20	
Δt	CV	CS	CV	CS	CV	CS	CV	CS	CV	CS
10	77.6	74.1	95.3	92.9	97.6	96.5	97.6	97.6	97.6	97.6
20	53.6	57.1	82.1	83.3	89.3	89.3	92.9	91.7	92.9	94.0
30	39.2	45.9	74.3	70.3	86.5	86.5	93.2	90.5	94.6	93.2
40	35.3	42.6	61.8	72.1	73.5	79.4	82.4	83.8	86.8	88.2
50	26.7	41.7	60.0	66.7	68.3	78.3	75.0	81.7	81.7	85.0
60	21.9	29.7	50.0	56.3	67.2	67.2	75.0	70.3	81.3	75.0

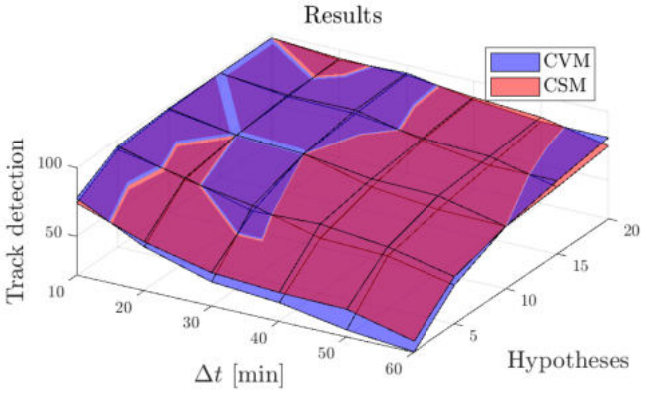

Fig. 4. The results in Table IV plotted as functions of time step $\Delta t = t_2 - t_1$ and in terms of how often the correct vessel was included in a cumulative set of hypotheses sorted by likelihood.

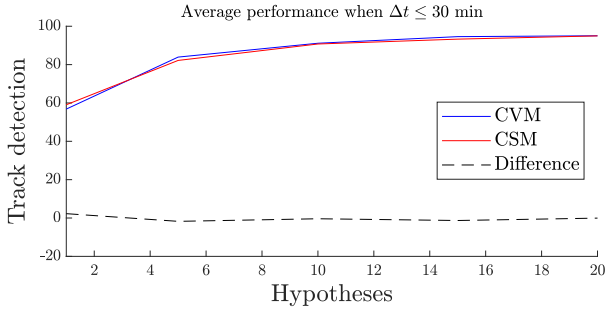

Fig. 5. Averaging the values in Table IV when the time steps $\Delta t \leq 30\,\text{min}$ shows that the outputs were similar. The plotted difference ranges from -1.74% to 2.27%.

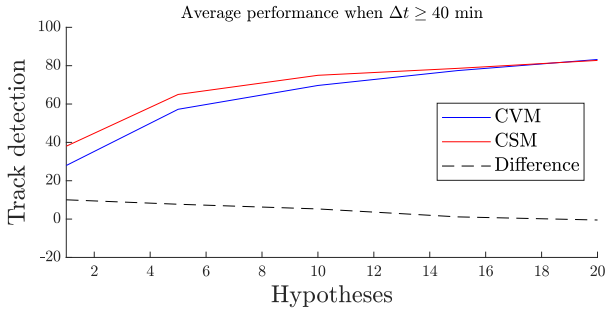

Fig. 6. Averaging the values in Table IV when the time steps $\Delta t \geq 40\,\text{min}$ shows that the CSM performed better than the CVM. The plotted difference ranges from -0.48% to 10.06%

V. DISCUSSION AND FUTURE WORK

Figure 5 show that the CSM generated similar results to the CVM when $\Delta t \leq 30\,\text{min}$. However, while the latter sorted each entry by the weight computed by a simple distance metric, the CSM required two FMMs for each track, which makes the CSM more time consuming when compared to the CVM. Therefore, the CVM seems to be a more reasonable choice than the CSM for small time steps.

Figure 6 show that the CSM performed better than the CVM when $\Delta t \geq 40\,\text{min}$ and the number of hypotheses was small. The CSM was able to assign the highest probability to the correct position measurement over 10% more often, in average, than the CVM. The difference between the evaluated associators decreased with increasing number of hypotheses, but the CSM outperformed the CVM by 7.7% when including 5 hypotheses, and 5.3% when including 10 hypotheses. The additional computational effort required by the CSM may be beneficial, as better hypotheses could improve the performance of the MHT.

Overall, the CSM seems to be a reasonable choice when the length of the traversed trajectory is long and the average speed is high, i.e., $\geq 5\,\text{kn}$. This may be because the underlying assumption of the CSM, namely constant speed along a trajectory moving away from the starting point, appears to be valid in such cases. It is also seen in Figure 7 that the CVM seems to be a reasonable associator in cases where the average speed is small, in this case $< 5\,\text{kn}$.

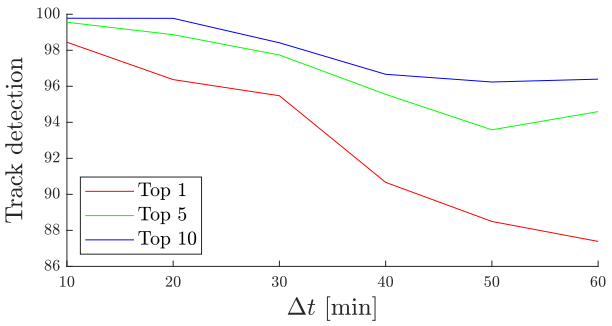

Fig. 7. The plot show how well the CVM performed when $d_{\mathrm{E}}(\boldsymbol{z}_1, \boldsymbol{z}_3) \leq \varepsilon$, and it is seen that associator assigned highest weight to the correct track in 87% of the cases. The correct measurement was included among the 5 most likely hypotheses in 93.5% of the cases, and 96% when evaluating the 10 most likely hypotheses.

If the resolution of the map used to evaluate the data is low, then position measurements near a coast line may be located on land due to round-off errors. This is a problem for the CSM which cannot evaluate such measurements, while the CVM is not affected by this issue. One ad hoc solution is to move a position measurement on land to the nearest body of water. However, it was decided to keep the raw data untinkered in this work. Measurements on land were classified as not found and assigned a weight equal to zero.

Future work includes further development of the CSM by including speed restricted zones, and including shipping lanes and ferry routes to improve the predicted probability density map. Another approach would be to adapt the wave propagation model to include the curvature of the trajectories of the moving vessels.

ACKNOWLEDGEMENT

We want to thank Kongsberg Satellite Services in Tromsø for providing data for this paper.

REFERENCES

[1] Jakob Andreas Bærentzen. *On the implementation of fast marching methods for 3D lattices*. 2001.

[2] Craig Carthel, Stefano Coraluppi, and Patrick Grignan. Multisensor tracking and fusion for maritime surveillance. pages 1 – 6, 08 2007.

[3] Norwegian Defence Research Establishment (FFI). Norsat-3, ship surveillance with a navigation radar detector. *Facts*, 2021.

[4] Brita Gade, Morten Kloster, Knut Landmark, and Carina N. Vooren. Gating for target tracking among islands and fjords. In *Global Oceans 2020: Singapore – U.S. Gulf Coast*, pages 1–6, 2020.

[5] International Telecommunication Union (ITU). ITU-R M.1372-5 – technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile band. *Recommendation, ITUR*, 02 2014.

[6] Maximilian Rodger and Raffaella Guida. Data Association Techniques for Near-Contemporaneous SAR and AIS Datasets from NovaSAR-1. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 700–703, 2019.

[7] Scott Scheff. State of the Industry: UAS Sensor Review. Technical report, NASA: Washington, DC, 2021.

[8] J. A. Sethian. Fast Marching Methods. *SIAM Review*, 41(2):199–235, 1999.

[9] James A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences of the United States of America*, 93 4:1591–5, 1996.

[10] Alexander M. Sjøberg and Brita H. H. Gade. Probability Distributions in Coastal Regions for Association of Naval Vessels. In *2023 26th International Conference on Information Fusion (FUSION)*, pages 1–6, 2023.

[11] Andreas Nordmo Skauen. Ship tracking results from state-of-the-art space-based AIS receiver systems for maritime surveillance. *CEAS Space J*, 11:301–316, 2019.

[12] Giovanni Soldi, Domenico Gaglione, Nicola Forti, Alessio Di Simone, Filippo Cristian Daffinà, Gianfausto Bottini, Dino Quattrociocchi, Leonardo M. Millefiori, Paolo Braca, Sandro Carniel, Peter Willett, Antonio Iodice, Daniele Riccio, and Alfonso Farina. Space-based global maritime surveillance. part i: Satellite technologies. *IEEE Aerospace and Electronic Systems Magazine*, 36(9):8–28, 2021.

[13] Arsenios Tsokas, Maciej Rysz, Panos M. Pardalos, and Kathleen Dipple. SAR data applications in earth observation: An overview. *Expert Systems with Applications*, 205:117342, 2022.