SCIENCE AND TECHNOLOGY ORGANIZATION
CENTRE FOR MARITIME RESEARCH AND EXPERIMENTATION

Reprint Series **CMRE-PR-2019-055**

Performance assessment of vessel dynamic models for long-term prediction using heterogeneous data

Gemine Vivone, Leonardo M. Millefiori, Paolo Braca, Peter Willett

June 2019

Originally published in:

IEEE Journal of Oceanic Engineering, volume 55, issue 11, November 2017, pp. 6533 – 6546, doi: 10.1109/TGRS.2017.2729622

About CMRE

The Centre for Maritime Research and Experimentation (CMRE) is a world-class NATO scientific research and experimentation facility located in La Spezia, Italy.

The CMRE was established by the North Atlantic Council on 1 July 2012 as part of the NATO Science & Technology Organization. The CMRE and its predecessors have served NATO for over 50 years as the SACLANT Anti-Submarine Warfare Centre, SACLANT Undersea Research Centre, NATO Undersea Research Centre (NURC) and now as part of the Science & Technology Organization.

CMRE conducts state-of-the-art scientific research and experimentation ranging from concept development to prototype demonstration in an operational environment and has produced leaders in ocean science, modelling and simulation, acoustics and other disciplines, as well as producing critical results and understanding that have been built into the operational concepts of NATO and the nations.

CMRE conducts hands-on scientific and engineering research for the direct benefit of its NATO Customers. It operates two research vessels that enable science and technology solutions to be explored and exploited at sea. The largest of these vessels, the NRV Alliance, is a global class vessel that is acoustically extremely quiet.

CMRE is a leading example of enabling nations to work more effectively and efficiently together by prioritizing national needs, focusing on research and technology challenges, both in and out of the maritime environment, through the collective Power of its world-class scientists, engineers, and specialized laboratories in collaboration with the many partners in and out of the scientific domain.

Copyright © IEEE, 2017. NATO member nations have unlimited rights to use, modify, reproduce, release, perform, display or disclose these materials, and to authorize others to do so for government purposes. Any reproductions marked with this legend must also reproduce these markings. All other rights and uses except those permitted by copyright law are reserved by the copyright owner.

NOTE: The CMRE Reprint series reprints papers and articles published by CMRE authors in the open literature as an effort to widely disseminate CMRE products. Users are encouraged to cite the original article where possible.

Performance Assessment of Vessel Dynamic Models for Long-Term Prediction Using Heterogeneous Data

Gemine Vivone, *Member, IEEE*, Leonardo M. Millefiori, *Member, IEEE*, Paolo Braca, *Senior Member, IEEE*, and Peter Willett, *Fellow, IEEE*

***Abstract*— Ship traffic monitoring is a foundation for many maritime security domains, and monitoring system specifications underscore the necessity to track vessels beyond territorial waters. However, vessels in open seas are seldom continuously observed. Thus, the problem of long-term vessel prediction becomes crucial. This paper focuses attention on the performance assessment of the Ornstein–Uhlenbeck (OU) model for long-term vessel prediction, compared with usual and well-established nearly constant velocity (NCV) model. Heterogeneous data, such as automatic identification system (AIS) data, high-frequency surface wave radar data, and synthetic aperture radar data, are exploited to this aim. Two different association procedures are also presented to cue dwells in case of gaps in the transmission of AIS messages. Suitable metrics have been introduced for the assessment. Considerable advantages of the OU model are pointed out with respect to the NCV model.**

***Index Terms*— Automatic identification system (AIS), high-frequency surface wave (HFSW) radar, long-term prediction, maritime surveillance, nearly constant velocity (NCV) model, Ornstein–Uhlenbeck (OU) process, synthetic aperture radar (SAR).**

I. Introduction

MODERN monitoring system specifications reflect the need for a continuous ability to track vessels at sea for traffic monitoring. Unfortunately, current sensor coverage does not allow a worldwide seamless observation of the maritime traffic. Usually, vessels sailing in open seas are in fact not continuously observed and there are often large gaps between consecutive observations of the same vessel. The lack of a capability to continuously monitor maritime traffic clearly increases the risk in terms of safety at sea, thus making the problem of long-term vessel prediction of crucial importance.

Many papers can be found in the scientific literature that address the problem of short-term target state prediction. As an instance, the nearly constant velocity (NCV) model [1], [2], which describes nonmaneuvering target dynamics with a velocity that is perturbed by a white noise process, is usually applied in the target tracking literature (see [1]–[4]), commonly with a prediction step only to the very near future. The long-term target state prediction is instead an overlooked problem, and only few items can be found in the literature that address this issue (see [5]–[8]). At the same time, by analyzing real-world automatic identification system (AIS) data, it turns out that a significant portion of maritime traffic tends to follow regular routes, maneuvering very seldom. NCV models have been proven useful in nonmaneuvering target short-term prediction [1], [2] and they have also been used in conjunction with AIS route information in [4]. In the case of long-term prediction, Ornstein–Uhlenbeck (OU) models can outperform the well-established NCV, improving the performance [8], [9].

The OU model is popular in several scientific fields (see [10]–[12]). It was first introduced in physics [10] to describe the velocity of a Brownian particle under the influence of friction, and can be seen as a modified Wiener process so that there is a tendency of the "walk" to move back toward a central location, with a greater attraction when it is further away from the center. In the tracking literature, the OU model has been discussed most notably in [13]–[17], where the stability of the OU and the so-called mixed OU (MOU) processes is studied. The main difference between OU and MOU is that the former includes a mean-reverting effect on the velocity component of the target state, while in the latter, the feedback term is applied on both the position and the velocity components of the target state. As time diverges, consequently, the OU model's velocity uncertainty is bounded (although position correspondingly unbounded), while both the position and velocity uncertainties are bounded for the MOU.

A key parameter for the long-term target state prediction [8] is the long-run mean velocity parameter of the OU model, which basically represents the desired velocity of the target. A sample mean estimator (SME) was recently proposed in [18] to estimate this parameter from the observations in the

Manuscript received January 17, 2017; revised June 20, 2017; accepted July 14, 2017. Date of publication August 7, 2017; date of current version October 25, 2017. The work of G. Vivone, L. M. Millefiori, and P. Braca was supported by NATO Allied Command Transformation under Project SAC000708—Data Knowledge Operational Effectiveness, and the work of P. Willett was supported by NPS through ONR under Contract N00244-16-1-0017. *(Corresponding author: Leonardo M. Millefiori.)*

G. Vivone, L. M. Millefiori, and P. Braca are with the North Atlantic Treaty Organization Science and Technology Organization Centre for Maritime Research and Experimentation, 19126 La Spezia, Italy (e-mail: gemine.vivone@cmre.nato.int; leonardo.millefiori@cmre.nato.int; paolo.braca@cmre.nato.int).

P. Willett is with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269 USA (e-mail: peter.willett@uconn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2017.2729622

presence of data samples that are unevenly spaced in time. The proposed SME is $\sqrt{n}$-consistent when the sampling time is random and asymptotically efficient when the sampling time is constant, and very close to the Cramér–Rao lower bound in the cases of practical interest for traffic monitoring applications.

The OU-based formulation reduces by orders of magnitude the uncertainty region of the predicted position with respect to the models available in the literature [8]. Many applications in maritime surveillance can benefit from this reduction of uncertainty.

1) *Search and Rescue Operations [8]:* In this case, a smaller uncertainty region implies a smaller search region, which can significantly improve the probability of success for search cases.
2) *Counter Piracy [8]:* Vessels can keep their AIS transmitters off for hours when sailing in risky areas in order not to be detected by pirates. The OU-based formulation can provide useful information even in the case of a lack of messaging from vessels for several hours.
3) *Association With Heterogeneous Sources of Information (e.g., Radar or Synthetic Aperture Radar (SAR) Detections):* The reduction of the uncertainty can lead to a reduction of the association ambiguity thanks to a narrower gating region, especially in open seas with a usually poor AIS coverage. Furthermore, a smaller uncertainty region can reduce the probability of missing a vessel of interest in a high-resolution radar acquisition, which acquisition is usually planned several hours in advance.

As opposed to [8], where the OU model was validated for real-world commercial maritime traffic, in this paper, we specifically address point 3) above by assessing the performance of the OU and the NCV models in terms of association with heterogeneous sources of information. To this aim, data from four different sources of information are considered in this paper: 1) AIS data; 2) high-frequency surface wave (HFSW) radar detections, recorded in the Ligurian Sea in 2009; 3) tracks, provided by a joint probability data association (JPDA) tracker fed by HFSW radar data; and 4) SAR images.

In order to assess the performance of long-term prediction models, gaps in the transmission of AIS data are simulated. This is the case, for instance, of ships sailing in open seas, when AIS messages can be only received by satellite receivers that often have a very limited temporal resolution (the typical revisit time is of the order of several hours). The prediction is performed using both the OU and the NCV models, along with the association with the other sources of information. Thanks to the simulated gaps in AIS data, the performance can be properly assessed by exploiting the availability of the ground truth (i.e., the correct association between the AIS and the other—secondary—source of information).

Regarding the association procedures among heterogeneous data, two different approaches are exploited in this paper to assess the performance of the long-term prediction models to be compared.

1) *Data Association Hypothesis Testing Procedure [2]:* Based on a well-known and a widely exploited approach in the tracking literature [2], a data association hypothesis test procedure is derived to evaluate the Type I error (missed association) and the Type II error (wrong association), in order to assess the association performance.
2) *Gating Procedure [2]:* This procedure is introduced to overcome the limitation of the above-mentioned hypothesis testing approach, which implicitly makes the assumption of clutter-free environment. The drawback of the proposed gating procedure is instead related to a lack of performance metrics that can be used in order to properly assess the association performance. Thus, in this paper, two performance metrics, which can be recast into the classical concepts of "detection rate" (DR) [or "misdetection rate" (MDR)] and "false alarm rate" [or "ambiguity rate" (AR)], are proposed to be exploited in conjunction with the gating procedure for performance assessment.

Extensive experimental analysis will show the advantages of using the OU model for long-term vessel prediction over the well-established NCV model. A more accurate estimation of the prediction covariance matrix given by the OU model will consequently enable us to properly set the Type I error in the hypothesis testing procedure. Furthermore, a considerable ambiguity reduction will also be shown during the gating procedure.

The rest of this paper is organized as follows. Section II describes the OU model and the NCV model for long-term prediction. The hypothesis testing and the gating procedures together with the metrics for performance assessment are detailed in Section III. Section IV is devoted to the description of the exploited heterogeneous data. The extensive experimental analysis is shown in Section V. Finally, concluding remarks are drawn in Section VI.

II. Vessel Dynamic Models: Ornstein–Uhlenbeck and Nearly Constant Velocity

The stochastic models of vessel (target) dynamics are introduced here. Let us indicate the target state at time k with

$$\boldsymbol{s}_k = [x_k, y_k, \dot{x}_k, \dot{y}_k]^T \tag{1}$$

where x_k and y_k are the 2-D Cartesian coordinates, also denoted by $\boldsymbol{x}_k = [x_k, y_k]^T$, $\dot{x}_k$ and $\dot{y}_k$ are the corresponding velocities, which are also indicated by $\dot{\boldsymbol{x}}_k = [\dot{x}_k, \dot{y}_k]^T$, and $[\cdot]^T$ is the transpose operator.

Remark: As in [8], this paper makes the following assumptions: 1) only nonmaneuvering vessels are considered and 2) the target positional state acquired by AIS sensors is noiseless. This differs from the target tracking literature, where the target state observation, usually acquired by radar systems, is corrupted by noise. Measurement noise of course still exists and is a factor, but it is negligible.[1] A less negligible source of error for AIS can be traced back to timing issues in multiple-receiver networks, and can be addressed with a dedicated methodology [19].

[1]The measurement noise of the AIS is that of the global positioning system, i.e., less than 10 m, and it is therefore negligible with respect to the distances involved in the experimental results (tens of kilometers).

A. Nearly Constant Velocity Model

A very popular target motion model, commonly exploited in the tracking literature, is the NCV model [1]. In this case, we have that the optimal prediction, given the target state $\boldsymbol{s}_{k-1}$, is as follows [2]:

$$\boldsymbol{s}_{k|k-1} = \mathbf{F}_k \boldsymbol{s}_{k-1} \tag{2}$$

where

$$\mathbf{F}_k = \begin{bmatrix} \mathbf{I} & T_k \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{3}$$

is usually referred to as the state transition matrix and T_k is the sampling time at time k.

Remark: For conventional tracking applications, the focus is short-term prediction, i.e., T_k is usually very small. Instead, the aim of the long-term prediction, which is the goal of this paper, is to characterize the predicted target state when T_k is high. Thus, the estimator is strongly dependent on the motion model described in [8], wherein it is also provided its covariance matrix

$$\mathrm{Cov}[\boldsymbol{s}_k|\boldsymbol{s}_{k-1}] = \begin{bmatrix} \frac{T_k^3}{3} & \frac{T_k^2}{2} \\ \frac{T_k^2}{2} & T_k \end{bmatrix} \otimes \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \tag{4}$$

where σ_x^2 is the variance along the coordinate x, σ_y^2 is the variance along the coordinate y, σ_{xy} is the cross covariance of x and y, and $\otimes$ is the Kronecker product.

Remark: A slightly different process noise assumption of the NCV ("direct discrete time," see [2]) would lead to the same estimator (2) but with a higher scaling law of the errors, i.e., proportional to $T_k^4/4$ instead of $T_k^3/3$ for the position and proportional to T_k^2 instead of T_k for the velocity. For that reason, two NCV models can be considered in this paper. The one with the cubic position error scaling law is hereafter called NCV3, whereas the other with scaling law raised to the power of four is named NCV4. The main difference between the NCV3 and NCV4 models is that the former comes out as the result of a continuous stochastic time integral, whereas the latter is the result of a stepwise constant noise process. As a consequence, the NCV4 can be seen as a low-fidelity approximation of the NCV3. Nonetheless, it is still taken into consideration in this paper, withstanding its widespread use in literature.

B. Ornstein–Uhlenbeck Model

For the OU model, we have that the optimal prediction, given the initial target state, is provided by [11], [20], and [8].

Thus, for the velocity

$$\dot{\boldsymbol{x}}_{k|k-1} = \boldsymbol{v} + \begin{bmatrix} e^{-\gamma_x T_k} & 0 \\ 0 & e^{-\gamma_y T_k} \end{bmatrix} (\dot{\boldsymbol{x}}_{k-1} - \boldsymbol{v}) \tag{5}$$

where $\boldsymbol{v} = [v_x, v_y]^T$ is the long-run mean velocity, which plays a key role in the proposed model representing the "typical" velocity in Cartesian coordinates of the target on the trajectory under consideration, and γ_x and γ_y are two parameters representing the mean reversion effect along the x and y components. Instead, for the target position that is an integrated OU process, we have

$$\boldsymbol{x}_{k|k-1} = \boldsymbol{x}_{k-1} + T_k \boldsymbol{v} + \begin{bmatrix} \frac{1 - e^{-\gamma_x T_k}}{\gamma_x} & 0 \\ 0 & \frac{1 - e^{-\gamma_y T_k}}{\gamma_y} \end{bmatrix} (\dot{\boldsymbol{x}}_{k-1} - \boldsymbol{v}). \tag{6}$$

Thus, the overall optimal prediction can be rearranged in matrix form as follows:

$$\boldsymbol{s}_{k|k-1} = \boldsymbol{\Phi}(T_k, \boldsymbol{\gamma})\, \boldsymbol{s}_{k-1} + \boldsymbol{\Psi}(T_k, \boldsymbol{\gamma})\, \boldsymbol{v} \tag{7}$$

where $\boldsymbol{\Phi}(T_k, \boldsymbol{\gamma})$ is the analog of the state transition matrix and $\boldsymbol{\Psi}(T_k, \boldsymbol{\gamma})\, \boldsymbol{v}$ is often called the control input function. Their definitions are provided in [8], and we report here only the variance terms

$$E[(x_{k|k-1} - x_k)^2 | \boldsymbol{s}_{k-1}] = \frac{\sigma_x^2}{\gamma_x^3} f(\gamma_x T_k) \tag{8}$$

$$E[(y_{k|k-1} - y_k)^2 | \boldsymbol{s}_{k-1}] = \frac{\sigma_y^2}{\gamma_y^3} f(\gamma_y T_k) \tag{9}$$

$$E[(\dot{x}_{k|k-1} - \dot{x}_k)^2 | \boldsymbol{s}_{k-1}] = \frac{\sigma_x^2}{\gamma_x} g(\gamma_x T_k) \tag{10}$$

$$E[(\dot{y}_{k|k-1} - \dot{y}_k)^2 | \boldsymbol{s}_{k-1}] = \frac{\sigma_y^2}{\gamma_y} g(\gamma_y T_k) \tag{11}$$

where $f(\cdot)$ and $g(\cdot)$ are the prediction position and velocity error normalized variances defined as follows:

$$f(\beta) = \frac{1}{2}(2\beta + 4e^{-\beta} - e^{-2\beta} - 3) \tag{12}$$

$$g(\beta) = \frac{1}{2}(1 - e^{-2\beta}). \tag{13}$$

Remark: The parameters in (4) and (9)–(11) and can be estimated, in a batch fashion, from the AIS data with the procedure proposed in [8, Sec. III].

III. Association Procedures

This section is devoted to the introduction of two association procedures to assess the performance of the long-term prediction models. The data association hypothesis testing procedure is presented first. Afterward, a gating procedure is detailed, together with two new performance metrics that rely upon the concepts of detection (or misdetection) rate and false alarm (or ambiguity) rate.

A. Data Association Hypothesis Test

The hypothesis testing procedure employed in this paper is based on a classical data association testing procedure [2], often used in the target tracking literature to associate tracks generated by two different sensors. However, a modified version of this procedure is applied in this paper to verify whether or not a generic measurement (not necessarily a track) originates from a confirmed target. In our case, the measurement is provided by the radar and the confirmed target is given by the AIS track.

In order to be directly comparable, the AIS track and the radar measurement should be simultaneous, but they are not. Therefore, a prediction procedure is required to propagate the AIS state at the time of the radar measurement. Our purpose is to characterize the association performance as function of the prediction time and model. The possible outcomes are that either the measurement was originated by the same target as the AIS track—association between prediction and measurement—or the measurement does not correspond to the AIS track—they come from separate targets.

Let us define the difference between two error-free target states at time k, namely, $\boldsymbol{x}_k^t$ for target t and $\boldsymbol{x}_k^s$ for target s

$$\Delta_k^{t,s} = \boldsymbol{x}_k^t - \boldsymbol{x}_k^s. \tag{14}$$

The *same target* ($t = s$) and *different targets* ($t \neq s$) statistical hypotheses, namely, $\mathcal{H}_0$ and $\mathcal{H}_1$, can be formulated as

$$\mathcal{H}_0 : \Delta_k^{t,s} = 0, \quad \mathcal{H}_1 : \Delta_k^{t,s} \neq 0. \tag{15}$$

We wish to understand if the radar measurement (track) s originated from target t observed via the AIS. To this aim, let us introduce now the difference between the prediction $\hat{\boldsymbol{x}}_{k|k-1}^t$ of the AIS target at the time k when the radar measurement (track) $\boldsymbol{\delta}_k^s$ was observed, given by [2]

$$\hat{\Delta}_k^{t,s} = \hat{\boldsymbol{x}}_{k|k-1}^t - \boldsymbol{\delta}_k^s. \tag{16}$$

It is noteworthy to mention that $\boldsymbol{\delta}_k^s$ can be either a measurement (i.e., radar detector output) or a track (i.e., radar tracker output), and $\hat{\boldsymbol{x}}_{k|k-1}^t$ is computed with the information collected up to time $k-1$, using the long-term prediction procedure described in Section II to align temporally the data. Under $\mathcal{H}_0$, the difference term in (16) is zero mean, and under the error independence assumption,[2] its covariance is

$$\Sigma_k^{t,s} = \mathbf{P}_{k|k-1}^t + \mathbf{R}_k^s \tag{17}$$

where $\mathbf{P}_{k|k-1}^t$ is the long-term prediction covariance (see Section II) and $\mathbf{R}_k^s$ is the measurement (track) covariance error; in the specific case of a radar measurement, this latter term is a function of the measurement $\boldsymbol{\delta}_k^s$.

Assuming the estimation errors to be Gaussian, the track association test is [2]

$$\left(\hat{\Delta}_k^{t,s}\right)^T \left(\Sigma_k^{t,s}\right)^{-1} \hat{\Delta}_k^{t,s} \leq \gamma_\alpha \tag{18}$$

to accept $\mathcal{H}_0$, where γ_α is the test threshold, such that

$$P\left\{\left(\hat{\Delta}_k^{t,s}\right)^T \left(\Sigma_k^{t,s}\right)^{-1} \hat{\Delta}_k^{t,s} > \gamma_\alpha \mid \mathcal{H}_0\right\} = \alpha. \tag{19}$$

This last probability is the so-called Type I error probability and α is the significance level of the test. The test threshold definition follows from the Gaussian assumption [2]:

$$\gamma_\alpha = \chi_{n_x}^2(1-\alpha) \tag{20}$$

where $\chi_{n_x}^2$ is the χ^2-distribution with n_x degree of freedom evaluated at the confidence level $1-\alpha$. In the present case, $n_x = 2$, meaning that only the positional coordinates on the Cartesian plane are considered for the test.

[2]Note that $\hat{\boldsymbol{x}}_{k|k-1}^t$ and $\boldsymbol{\delta}_k^s$ could be not independent [2]. A higher computational burden is required to take into account the dependence.

Remark: It is noteworthy to mention that the data association hypothesis testing procedure has the goal of identifying if two estimates are related either to the same target (i.e., under hypothesis $\mathcal{H}_0$) or to two different targets (hypothesis $\mathcal{H}_1$). Thus, if the measured difference $\hat{\Delta}_k^{t,s}$ between the prediction of the confirmed target $\hat{\mathbf{x}}_{k|k-1}^t$ and the measurement (track) $\boldsymbol{\delta}_k^s$ at time k, is *significant*—large enough with respect to the selected significance level of the test α—the decision is $\mathcal{H}_1$ (no association), i.e., the predicted state at time k, $\hat{\mathbf{x}}_{k|k-1}^t$, and the measurement (track) $\boldsymbol{\delta}_k^s$ belong to two different targets; otherwise, the hypothesis $\mathcal{H}_0$ is selected ($\hat{\mathbf{x}}_{k|k-1}^t$ and $\boldsymbol{\delta}_k^s$ originated from the same target). Assuming the estimation errors to be Gaussian, it is possible to derive test (18) and calculate the threshold γ_α, given the significance level of the test α. Whereas targets can be assumed to follow the Gaussian assumption, the clutter—usually uniformly distributed—invalidates the test (18). For this reason, in the experimental results using real data, the hypothesis $\mathcal{H}_1$ is simulated, and the sole validation is about the capability of the OU model to correctly set the Type I error (i.e., the model, which is followed by real targets, is OU; see Fig. 6).

B. Gating and Performance Assessment

The procedure described in the previous section implicitly makes the hypothesis of a clutter-free environment, i.e., the clutter is not taken into account using the hypothesis test approach. Thus, even though real data are exploited in this paper, the hypothesis $\mathcal{H}_1$ had to be simulated for the association test (18). To overcome this limitation, a multidimensional gate is exploited to associate heterogeneous data with the predicted position of the target. The advantage of this procedure is that it can be used in a real cluttered environment. However, suitable metrics that can be used for association performance assessment have to be defined. In this paper, two performance metrics are proposed in Section III-B2 that can be applied in conjunction with the gating procedure briefly described in Section III-B1 and widely adopted in the tracking literature [2].

1) Gating Procedure: In order to associate the heterogeneous data with the predicted position of the AIS track, a multidimensional gate is set up. A measurement within the gate, while not guaranteed to have originated from the target the gate pertains to, is a valid association candidate, and hence the name of *validation region* or *association region* [2].

We assume that the heterogeneous data $\boldsymbol{\delta}_k^t$ at time k coming from a target t (e.g., an HFSW radar contact/track or SAR detection) are Gaussian distributed with mean $\hat{\boldsymbol{x}}_{k|k-1}^t$ and covariance $\Sigma_{k|k-1}^t$. Thus, its gate at time k is defined as

$$\mathcal{V}_k^{t,\gamma} = \left\{\boldsymbol{\delta} : \left[\boldsymbol{\delta} - \hat{\boldsymbol{x}}_{k|k-1}^t\right]^T \left(\Sigma_{k|k-1}^t\right)^{-1} \left[\boldsymbol{\delta} - \hat{\boldsymbol{x}}_{k|k-1}^t\right] \leq \gamma\right\} \tag{21}$$

with a probability depending on the threshold γ. Under the above-mentioned hypothesis, this quadratic form is chi-square distributed with degrees of freedom equal to the dimension n_x of the positional components of the target state. The probability P_G that $\boldsymbol{\delta}_k^t$ is in $\mathcal{V}_k^{t,\gamma}$, i.e., $P_G = P\{\boldsymbol{\delta}_k^t \in \mathcal{V}_k^{t,\gamma}\}$, which depends on both n_x and γ, is reported in Table I. Finally, it is worthy to note that the square root g of the

TABLE I
GATE THRESHOLDS γ AND THE RELATED PROBABILITY MASS P_G FOR $n_x = 2$

γ	1	4	9	16	25
g	1	2	3	4	5
	0.3930	0.8650	0.9890	0.9997	1.0000

threshold γ, i.e., $g = \sqrt{\gamma}$, is often referred to as "number of sigmas" (standard deviations) of the gate.

Remark: When the association between the predicted AIS position and the radar detections is performed, the uncertainties of the measurements provided by the radar system have also to be taken into account (measurement noise has an impact, in particular, for short-term predictions). Thus, the noise covariance matrix for a generic measurement $\boldsymbol{\delta}_k$, $\mathbf{R}_k$, should be taken into account. Since the radar measurement noise is conditionally independent of the AIS prediction of target t position given the true starting target state, the sum of these two random variables is still a Gaussian random variable with mean $\hat{\boldsymbol{x}}^t_{k|k-1}$ and covariance matrix $\Sigma^t_{k|k-1}$ that is the sum of the AIS prediction covariance matrix $\mathbf{P}^t_{k|k-1}$ for target t at time k and the matrix $\mathbf{R}_k$

$$\Sigma^t_{k|k-1} = \mathbf{P}^t_{k|k-1} + \mathbf{R}_k \tag{22}$$

where $\mathbf{R}_k$ is a function of the measurement $\boldsymbol{\delta}_k$ in the specific case of a radar measurement.

2) Performance Metrics for Gating: This section is devoted to the introduction of some performance metrics to assess the suitability of the OU model for vessel long-term prediction. Such metrics should quantify the uncertainty in the gating association, which can generate false associations, and the ability to associate the predicted target with a measurement provided by another acquisition system.

Thus, denote $\mathcal{D}_k = \{\boldsymbol{\delta}^i_k\}_{i=1}^{N_k}$ to be the set of the N_k heterogeneous data at time k and also denote $\mathcal{D}^{t,\gamma}_k = \mathcal{D}_k \cap \mathcal{V}^{t,\gamma}_k$ as the subset of validated data from (21). The performance indices are as follows.

1) *AR:* The AR using a threshold γ for the target t at time k, i.e., $AR^{t,\gamma}_k$, is defined as

$$AR^{t,\gamma}_k = \frac{\left|\mathcal{D}^{t,\gamma}_k\right| - \mathbf{1}_{\mathcal{V}^{t,\gamma}_k}\left(\boldsymbol{\delta}^t_k\right)}{N_k - \mathbf{1}_{\mathcal{A}}\left(\boldsymbol{\delta}^t_k\right)} \tag{23}$$

where $|\cdot|$ indicates the set cardinality, $\mathbf{1}_{\mathcal{X}}(y)$ is the indicator function that is 1 if $y \in \mathcal{X}$ and 0 otherwise, $\mathcal{A}$ represents the whole surveillance area, and $\boldsymbol{\delta}^t_k$ is the measurement originated by target t at time k. That is, the numerator is the number of gated measurements irrelevant to the target, and the denominator is total number of measurements irrelevant to the target regardless of whether they have been gated or not. If N_k is equal to the total number of radar detections in the surveillance area at time k, we have the *radar AR* (RAR), $\mathrm{RAR}^{t,\gamma}_k$. Instead, if the normalization N_k is calculated considering all the tracker contacts in the surveillance area at time k, the index is called *tracker AR* (TAR), $\mathrm{TAR}^{t,\gamma}_k$.

The overall AR index is obtained by averaging $\mathrm{AR}^{t,\gamma}_k$ on the number of targets and time. Note that the overall AR index is only a function of γ (gate threshold). The overall indices for the radar and the tracker cases are indicated as RAR^{γ} and TAR^{γ}, respectively. Both the indices assume values in the range $[0, 1]$ and the ideal values are 0, i.e., no ambiguity is present.

2) *MDR:* The MDR using a threshold γ for a target t at time k, i.e., $\mathrm{MDR}^{t,\gamma}_k$, is defined as

$$\mathrm{MDR}^{t,\gamma}_k = 1 - \mathbf{1}_{\mathcal{V}^{t,\gamma}_k}(\boldsymbol{\delta}^t_k) \tag{24}$$

Namely, $\mathrm{MDR}^{t,\gamma}_k = 1$ if the measurement provided by target t at time k ($\boldsymbol{\delta}^t_k$) is in the set $\mathcal{V}^{t,\gamma}_k$, otherwise $\mathrm{MDR}^{t,\gamma}_k = 0$. Again, the overall MDR index is obtained by averaging $\mathrm{MDR}^{t,\gamma}_k$ on the number of targets and time. Note that the overall MDR index is only function of γ, it can assume values only in the range $[0, 1]$, and the ideal value is 0, i.e., the measurement originated by the target is always in the gate. Finally, the DR can be simply defined starting from MDR^{γ}, i.e., $\mathrm{DR}^{\gamma} = 1 - \mathrm{MDR}^{\gamma}$.

Remark: It is worthy to remark that AR $= 0$ and MDR $= 1$ can trivially be obtained when $\gamma = 0$, or, equivalently, AR $= 1$ and MDR $= 0$ can simply be reached when $\gamma = +\infty$. This is in line with what happens in the case of false alarm probability and detection (or misdetection) probability.

In order to give an idea of how these metrics work, an illustrative example using real data is described in Section V (see Fig. 7). This clarifies how the two proposed performance metrics reach the above-mentioned goals of measuring the association uncertainty together with the ability to identify the right association. The association ambiguity is quantified by the AR index, whereas the identification of the right association can be inferred via the MDR index. Thus, the greater the gating region, the greater the probability to retain the right measurement for the association. However, the other side of the coin is an increased ambiguity, i.e., more measurements can fall into the gating region. Thus, in the proposed example, it is apparent how the most accurate prediction model, i.e., OU, has a smaller gating region compared with the other models, i.e., a reduced uncertainty or ambiguity. At the same time, this gating region is large enough to retain the right association.

Remark: These indices can be employed in the performance assessment using HFSW radar detections or JPDA tracks. Instead, in the case of the presented SAR data sets, where there is no detection provided by other targets/clutter, the validation using the AR index cannot be properly performed (i.e., it practically gets the same values for all the compared models). Thus, in this case, the MDR index is evaluated for the different models and, instead of measuring the AR, which cannot be evaluated due to the absence of other measurements, the direct calculation of the gating volume provided by the compared models is provided. Thus, we have that the volume $V^{t,\gamma}_k$ of the gating region $\mathcal{V}^{t,\gamma}_k$ in (21) corresponding to the threshold $\gamma = g^2$ for target t at time k is

$$V^{t,\gamma}_k = c_{n_x}\left|\gamma\,\mathbf{P}^t_{k|k-1}\right|^{1/2} = c_{n_x} g^{n_x}\left|\mathbf{P}^t_{k|k-1}\right|^{1/2} \tag{25}$$

where $|\cdot|$ indicates the determinant

$$c_{n_x} = \frac{\pi^{n_x/2}}{\Gamma(n_x/2+1)} \tag{26}$$

and $\Gamma(\cdot)$ is the Gamma function. In the particular case where $n_x = 2$, $c_2 = \pi$.

IV. Heterogeneous Data Description

This section provides an overview of the types data used in this paper. AIS data are briefly described first. Afterward, some details are provided to the readers with regard to the HFSW radar data and the tracking approach exploited to track targets starting from the HFSW radar data. Finally, SAR data are briefly described.

A. Automatic Identification System

Vessels exceeding a given gross tonnage[3] are equipped with AIS transponders for position reporting, as mandated by the safety of life at sea convention [21].

The AIS contacts retain information related to the vessel dynamics. In particular, the position of the vessel in latitude and in longitude, the acquisition time, the course over ground, and the speed over ground are usually included into an AIS message broadcast by a vessel. Static information is also available, such as the type and the size of the vessel.

Whereas this system allows vessels to be aware and keep track of each other, coastal states are also able to receive, plot, and log these data via coastal base stations. AIS data have also been used as ground truth to assess performance in the tracking literature [3], [4]. Recently, low earth orbit satellite constellations have also been designed to receive AIS signals, providing a global monitoring capability for all AIS-equipped vessels.

B. HFSW Radar Data

The Wellen radar (WERA) HFSW radar system was installed by the North Atlantic Treaty Organization Science and Technology Organization Center for Maritime Research and Experimentation on the Palmaria Island (44° 2′ 30″ N, 9° 50′ 36″ E) approximately 200 m above the sea level, and it was operated from May to December 2009. Fig. 1 depicts its location and its field of view.

The system operating frequency was $\approx$ 12.5 MHz (corresponding to a wavelength of $\lambda \approx 25$ m). The WERA setup was composed by decoupled transmitting and receiving antenna arrays. The transmit array utilized four antennas arranged in a rectangular shape, whereas the receiving array had 16 antennas along a line perpendicular to the look direction. Electronic control of the arrays was adopted to sweep a 120° angular sector depending on the bandwidth, while Doppler resolution was achieved using continuous wave signals [22]. The system used a linearly frequency-modulated continuous wave, that is, a linear chirp with a configurable range resolution between 0.3 and 1.5 km, when the chirp bandwidth is 100 kHz. The azimuth resolution is 1°. This system is able to reach a maximum range of about 130 km assuring a large surveillance area at low cost in terms of energy consumption (this system operated at approximately 35 W on average). Data were recorded by each antenna element in complex samples along the range cells. The remaining 206 s was used to select a new free HFSW channel (between 12.2 and 12.6 MHz) and the available bandwidth according to spectrum crowding. The system used the same operating frequency, but the modulating waveforms were orthogonal to each other to avoid coupling interference.

[3]The AIS is required for all the ships exceeding 300 gross tons and engaged on international voyages, for all cargo ships of 500 gross tons, not engaged on international voyages, and all passenger ships. On average, a gross weight of 300 tons corresponds to a length of about 25 m.

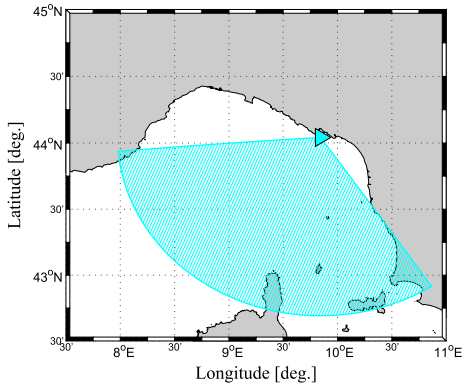


Fig. 1. Setup of the WERA system located in Palmaria, Italy. The radar field of view is depicted in cyan, whereas the radar location is pointed out using a cyan triangle marker.

Data from the antennas were beamformed to retrieve the azimuth information. Afterward, the target detection was performed in the Fourier domain by a 3-D ordered-statistic constant false alarm rate algorithm [23]. An accurate description of the detector that was employed is provided in [23].

C. JPDA Tracker for HFSW Radar Data

Another kind of data that can be exploited is represented by the outcomes of a tracking algorithm applied to the HFSW radar data. The tracker used in this paper relies upon the joint probabilistic data association (JPDA) algorithm [2], [24]. An NCV model is used as dynamic model, whereas the measurement model takes into account the measurements acquired by the HFSW radar system (i.e., range, azimuth, and range rate). The measurement-to-track association is based on the JPDA paradigm, which is a Bayesian approach that associates all the validated measurements to the tracks by probabilistic weights. The filtering stage is performed using the unscented Kalman filter [25]. Finally, the track management is based on the popular M/N logic [2]. Table II summarizes the set of parameters, already defined in [3], used to track targets exploiting the JPDA tracker applied to HFSW radar data [3]. Interested readers can refer to [3] and [26] for further details.

D. SAR Data

The SAR data herein employed were acquired by Sentinel-1A and provided freely by European Space Agency (ESA) through the Sentinels Scientific Data Hub.

TABLE II

PARAMETER SETTINGS FOR THE JPDA TRACKER USING HFSW RADAR DATA [3]

Parameter	Value	Specification
T_k	16.64/33.28 s	Radar sampling time
σ_v	10^{-2} m s^{-2}	Std. dev. process noise
σ_r	150 m	Std. dev. noise range
σ_b	1.5°	Std. dev. noise bearing (azimuth)
$\sigma_{\dot{r}}$	10^{-1} m s^{-1}	Std. dev. noise range rate
γ	5^2	Gate threshold
P_D	0.35	Detection probability
λ	10^{-9} m^{-2}	Clutter density
v_{max}	20 m s^{-1}	Maximum velocity
M/N	5/6	Track initialization logic
M^*/N^*	10/10	Track termination logic

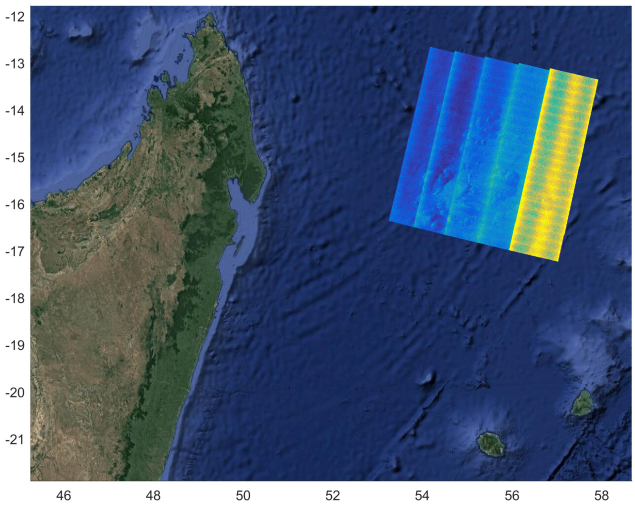


Fig. 2. SAR image used in the experimental analysis, spanning an area of about 400 × 400 km^2. The image was acquired by Sentinel-1A on May 31, 2016 in the EW swath operating mode and is located off the coast of Antananarivo, Madagascar. The complete figure spans an area of approximatively 1400 × 1100 km^2 (produced from ESA remote sensing data).

Sentinel-1A is the first satellite in the new fleet of ESA satellites, and carries a C-band SAR system that builds on ESA's and Canada's heritages on European Remote Sensing (ERS)-1, ERS-2, Envisat, and Radarsat.

Specifically, the SAR image used in this paper was acquired off the coast of Antananarivo, Madagascar, about 350 nmi north to the Mauritius island, and is shown in Fig. 2. The acquisition time was May 31, from 01:52:51 to 01:53:55 UTC, on a descending orbit, with a central operating frequency of 5.405 GHz and a range bandwidth of 22.2 MHz.

The operation mode of the SAR instrument at the time of the acquisition of this image was the extra wide (EW) swath mode, which employs the Terrain Observation with Progressive Scans SAR (TOPSAR)[4] technique [27] to acquire data over a wide

[4]The TOPSAR technique is a form of ScanSAR imaging, where data are acquired in bursts by switching the antenna beam between multiple adjacent subswaths.

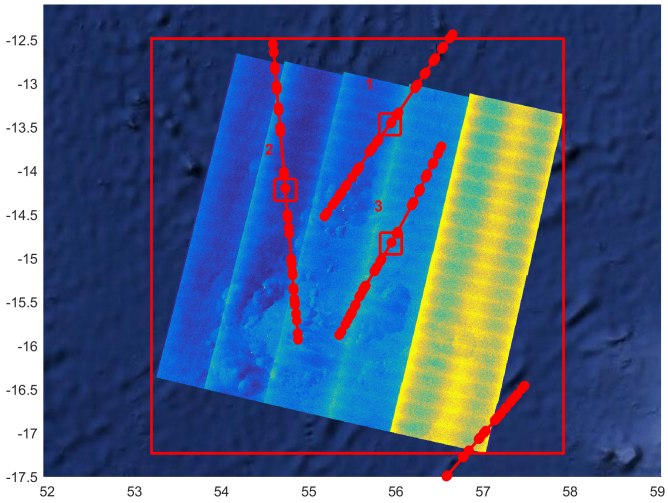


Fig. 3. SAR image, spanning an area of 400 × 400 km^2, used in the experimental analysis with AIS trajectories overlaid, which have been recorded in the area of the acquisition 6 h prior and posterior to the acquisition time. Red circles denote the AIS contacts. The three red squares are centered in the positions of the targets detected in the SAR image. A fourth AIS trajectory yields no association with any SAR target, being the acquisition times not compatible. The complete figure spans an area of approximatively 740 × 600 km^2 (produced from ESA remote sensing data).

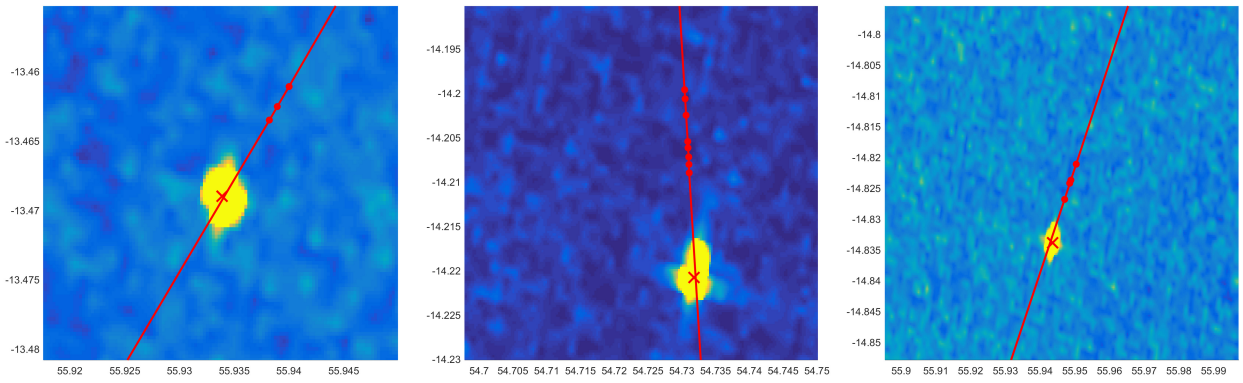
area using five subswaths. Fig. 2 shows a medium resolution ground range detected product, which consists of focused SAR data that have been detected, multilooked, and projected to ground range using an earth ellipsoid model such as WGS84. The range-azimuth pixel spacing of this product is 40×40 m^2. Fig. 3 shows the same SAR image with overlaid the AIS trajectories recorded in the area 6 h prior and posterior to the SAR acquisition time.

In the postprocessing phase, a number of other steps were undertaken, in order to obtain the ship detections from the SAR image, using the routines made available by ESA through the SNAP desktop application. First of all, a radiometric calibration of the image was performed [28], followed by a multilook despeckling step that averaged the received power over three looks in range and three looks in azimuth, and then, a land-sea masking step, followed by a CFAR detection routine with $P_{fa} = 10^{-8.5}$, target window size set to 75 m, guard window size set to 400 m, and background window size of 4 km.

These steps leave us with three detections[5] of cargo[6] ships, highlighted with rectangular boxes in Fig. 3, and also shown in Fig. 4 at a much closer zoom level. The three detections will be used afterward for the experimental analysis. Each of them perfectly matches with one of the AIS trajectories recorded in the area that are also shown in Figs. 3 and 4.

[5]Thanks to the perfect match with AIS trajectories, a full-fledged detector is not really required for our analyses with SAR data, as the targets could be easily identified by visual inspection. This detection step should be intended only as a *completeness* exercise.

[6]The ship category was inferred from the AIS information and can be assumed correct, given the perfect match of the SAR detections with the corresponding AIS trajectories.



(a) (b) (c)

Fig. 4. SAR images of the three cargo ships used in the experimental analysis, matched with the corresponding AIS trajectories, overlaid in red. The category of the ship has been inferred from the AIS information. Red crosses denote the SAR detection, which is shown to be aligned with the corresponding AIS track in all the three cases (produced from ESA remote sensing data). (a) *Test 1*: 3.5×2.8 km close-up. (b) *Test 2*: 5.6×4.4 km close-up. (c) *Test 3*: 10.9×6.3 km close-up.

V. Experimental Results

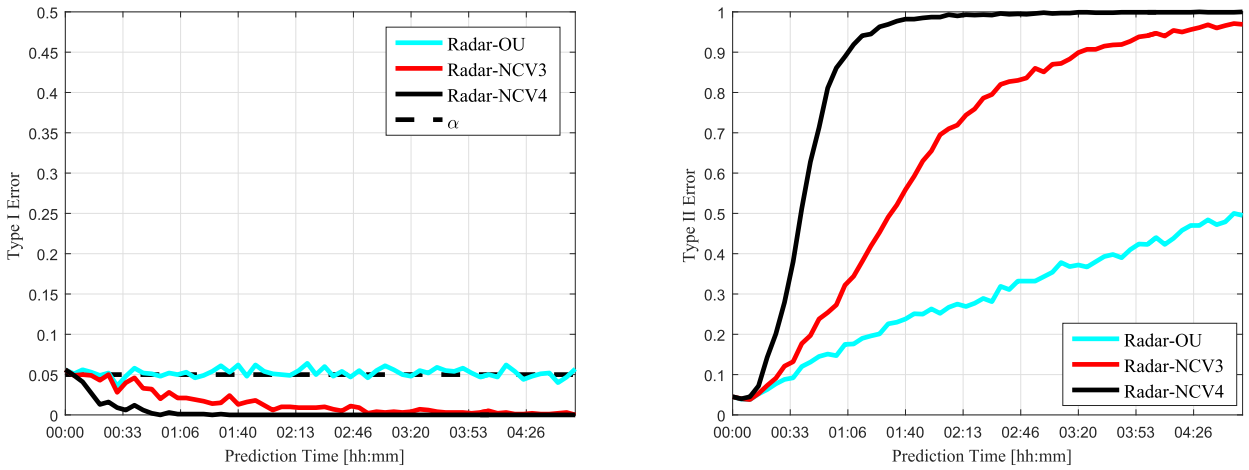
In this section, the experimental results obtained using different data sets comparing the three stochastic models (i.e., OU, NCV3, and NCV4) are described. The first analysis exploits the data association hypothesis testing procedure on both simulated and real data acquired by the HFSW radar system located in Palmaria, Italy, used in conjunction with AIS data. On the other hand, the second analysis is devoted to the gating procedure used to associate heterogeneous data. A first testbed is related to the association of AIS data with either HFSW radar detections or tracks obtained by the application of the JPDA tracker to HFSW radar data. The second test regards the assessment using SAR images.

A. Data Association Hypothesis Testing Procedure Using HFSW Data

A simulated environment is created first. AIS data provided by a generic target t are simulated following an OU model with parameters randomly selected in the set of the ones estimated using the real AIS tracks in the proposed experimental analysis. Tracks under the null hypothesis $\mathcal{H}_0$ are sampled from a Gaussian distribution, where the mean is the position in Cartesian coordinates of the target t and the covariance is defined in polar coordinates with standard deviations $\sigma_r = 150$ m, for the range, and $\sigma_b = 1.5°$, for the azimuth, in order to take into account the HFSW radar instrumental noise. The alternative hypothesis $\mathcal{H}_1$ is built by shifting the position of the target t by a quantity $\Delta_s = 3$ km (in both x and y coordinates) in order to simulate the acquisition by the radar of a different target. Again, this value is used as the mean of a Gaussian distribution with the above-mentioned covariance to take into account the HFSW radar instrumental noise. The parameters for the prediction models are estimated with the procedure indicated in Section II (see also [8] for details). The gaps in AIS data are simulated removing AIS data (i.e., comparing the performance of the models simulating different prediction times). Thus, the test in (18) is implemented and the Type I and Type II errors are evaluated (see Fig. 5). It is easy to see that using the OU model, it is possible to effectively set the Type I error (missed association) to a desired value $\alpha = 0.05$. This is not possible using NCV models, because the estimated covariances are not comparable with the simulated ones and the mismatch does not allow to set the Type I error using the test (18). Furthermore, it is also shown that the Type II error for the OU model is always lower than the ones obtained by the compared NCV models.

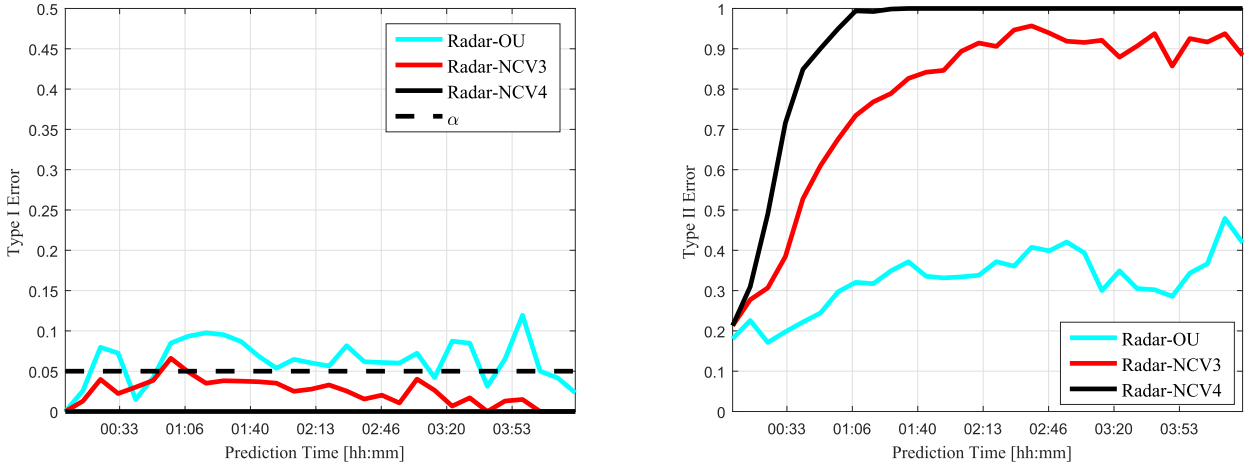
The simulated results are corroborated by real data. In this case, AIS data are represented by real AIS contacts that can be used to assess the performance of the three long-term prediction models (i.e., OU, NCV3, and NCV4). The error-free association between the real HFSW radar detections and the AIS data is possible thanks to the availability of the full AIS track (without the simulated intermittence).

The data set under the null (same target) hypothesis $\mathcal{H}_0$ comes from the association of the (nonpredicted) AIS positions with their closest (in space and time) radar contacts. It is worthy to note that this is correct in most of the cases, with some exceptions, because of clutter or close targets. Targets under $\mathcal{H}_1$ are instead simulated as in the previous case, i.e., by shifting the AIS target position of $\Delta_s = 3$ km in both x and y coordinates. The parameters for the prediction models are estimated as described in Section II. Thus, the test in (18) is implemented and the Type I and Type II errors are evaluated. The outcomes, averaged on 3034 real trajectories acquired by the HFSW radar located in Palmaria, Italy, are depicted in Fig. 6. We can conclude that the process generating the real AIS data better matches the OU process than the NCV one [8]. As a result, this allows us to set the Type I error to a desired value thanks to a proper estimation of the covariance matrix used in the test (18). Instead, the two compared NCV-based prediction models are not able to achieve it, and as a consequence, their Type I error decreases over the prediction time (i.e., an overestimation of the NCV covariance matrices with respect to the one obtained by the OU model can be pointed out). Furthermore, the Type II error of the OU model is again lower than the ones obtained by the compared NCV models corroborating the simulated outcomes.



(a) (b)

Fig. 5. Simulated results using the data association hypothesis testing procedure averaged on 10^3 Monte Carlo trials. (a) Type I error. (b) Type II error. While NCVs' decrease in Type I (missed association) error may seem appealing, it is caused by an overestimation of the estimation uncertainty whose effect is an intolerable increase in Type II error (incorrect association). The OU models' controlled Type I error and manageable Type II error are far preferable.



(a) (b)

Fig. 6. Results on real HFSW radar detections using the data association hypothesis testing procedure averaged on 3034 trajectories. (a) Type I error. (b) Type II error.

B. Gating Procedure Using Heterogeneous Data

In this section, a performance assessment is provided, which exploits a gating procedure for the three prediction models based on the NCV models and the OU process, using data acquired by either the HFSW radar system or the JPDA tracker or the SAR system.

1) Gating Procedure Using HFSW Radar and Tracking Data: A qualitative analysis is performed first. Fig. 7 depicts the benefits in terms of target localization with a relevant reduction of the uncertainty area after approximately 2 h and 30 min of prediction using both the OU and the NCV4 models. It is straightforward to see that the OU estimation of the target position (cyan cross) almost overlaps the true AIS position (black triangle), whereas NCV provides an estimation (red cross), which is several kilometers off the true position. Furthermore, the related 100%-confidence prediction covariance ellipses (plotted in black, NCV4, and cyan, OU) are very different in size. Indeed, the size of the OU prediction covariance is considerably smaller than that of the NCV4 prediction covariance. This difference is well captured by the performance metrics introduced in Section III-B2. Indeed, the RAR can be seen as an indirect measurement of how large the NCV4 gate is with respect to the OU gate (or, equivalently, how much uncertainty there is in the NCV4 case with respect to the OU case). On the one hand, the MDR is 0 (ideal value) in both the cases. This is due to the fact that both the compared models include in their 100%-confidence prediction covariance ellipse the HFSW detection provided by the predicted target. On the other hand, the AR in the OU case is 0 (ideal value), i.e., its 100%-confidence prediction covariance ellipse includes only one detection that is the one originated by the predicted target, whereas it is about 0.17 (i.e., six detections in the gate) in the case of NCV4. Thus, in this real example, the advantages given by the OU model for associating the HFSW detections

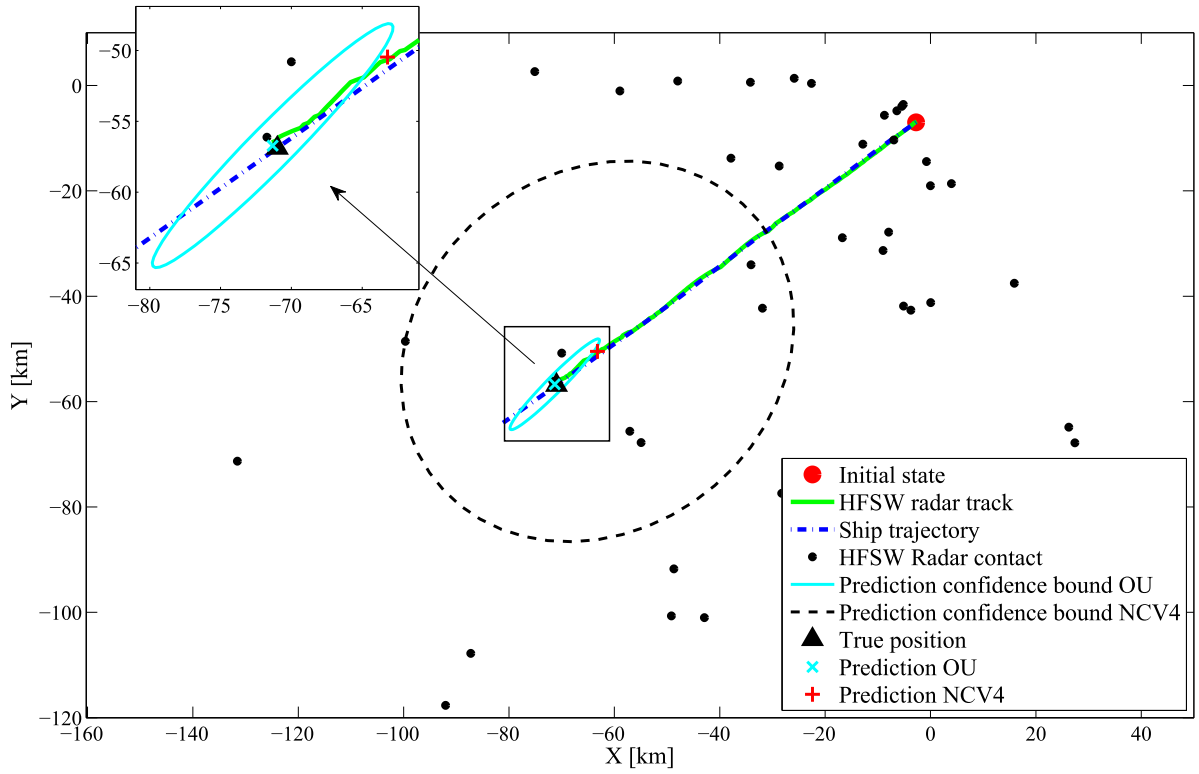


Fig. 7. Portion of one of the trajectories acquired by the HFSW radar system (converted into projected coordinates) is depicted. The predictions of target position after approximately 2 h and 30 min from the initial state are indicated with cross markers. The related 100%-confidence prediction covariance ellipses are plotted with black dashed line (NCV4) and cyan continuous line (OU). The true position is indicated with a black triangle, the HFSW radar contacts are plotted with black dots, whereas the HFSW radar track is shown with a green line.

with the predicted AIS target are clear and properly assessed using the proposed metrics.

A quantitative assessment is performed using one month of data acquired by the HFSW radar located in Palmaria, Italy. The evaluated metrics are the MDR, MDR (or, equivalently, the DR), and the AR, presented in Section III-B2. To assess the performance, 3034 trajectories, following the nonmaneuvering hypothesis under the use of the OU model, have been selected. The validation is performed considering simulated gaps in AIS data, which define the x-axis (prediction time) in Fig. 8. This procedure enables us to properly identify the correct association between AIS and HFSW data (defining a ground truth, see the discussion in the previous section) and consequently evaluate properly the performance of the three compared prediction models.

The first quantitative analysis is related to the understanding of the performance metric trend over the prediction time. This is obtained by fixing the gate threshold γ and averaging the indices for all the predicted targets in the surveillance area on contiguous prediction intervals. The prediction time is bounded to 4.5 h. The selected gate threshold γ is 25 (i.e., $g = 5$ sigmas and $P_G = 1$). Fig. 8 shows the obtained outcomes, with both the radar and TARs exhibiting the same behavior. A significant increment of the ambiguity is observed over the prediction time. This is much more apparent for the NCV models [i.e., NCV3 and NCV4; see Fig. 8(a) and (c)]. The increment is related to the growth of the uncertainties of the compared models when the prediction time increases. Thus, the AR values obtained with the OU model are always better than those achieved with the NCV models. On the other side, the better outcomes provided by the OU on the AR metrics are paid by an increased MDR [or, equivalently, a reduction of the DR; see Fig. 8(b) and (d)]. Therefore, further analyses are required to quantify the advantages in using the OU model for vessel long-term prediction.

Such analyses are related to the variation of the gate threshold γ in the range $[0.05^2,\ 20.05^2]$ to depict the behavior of MDR against AR averaged over prediction time. Fig. 9 clearly shows the advantages of the OU model for both the cases of HFSW radar detections and the JPDA tracks. Table III summarizes the values of the RAR and TAR metrics for some fixed values of the MDR using both radar and tracker data for the three compared models. For very high values of the MDR, the models perform similarly (see Fig. 9). The greatest advantage in using the OU model over the NCV ones is shown for MDR values in the range $[10^{-3}, 10^{-1}]$. For instance, fixing MDR to 10^{-1}, the reduction of the RAR for the OU model is about the 100% and about the 180% with respect to the NCV3 model and the NCV4 model, respectively. In the case of MDR equal to 10^{-2}, the gain obtained by the OU model with respect to the NCV3 model is almost the same as before, whereas the RAR reduction with respect to the NCV4 is about the 160%.

A similar analysis is performed using the JPDA tracker's outcomes as inputs for the association. In this case, the AR is measured using the TAR index. Again, the advantages of

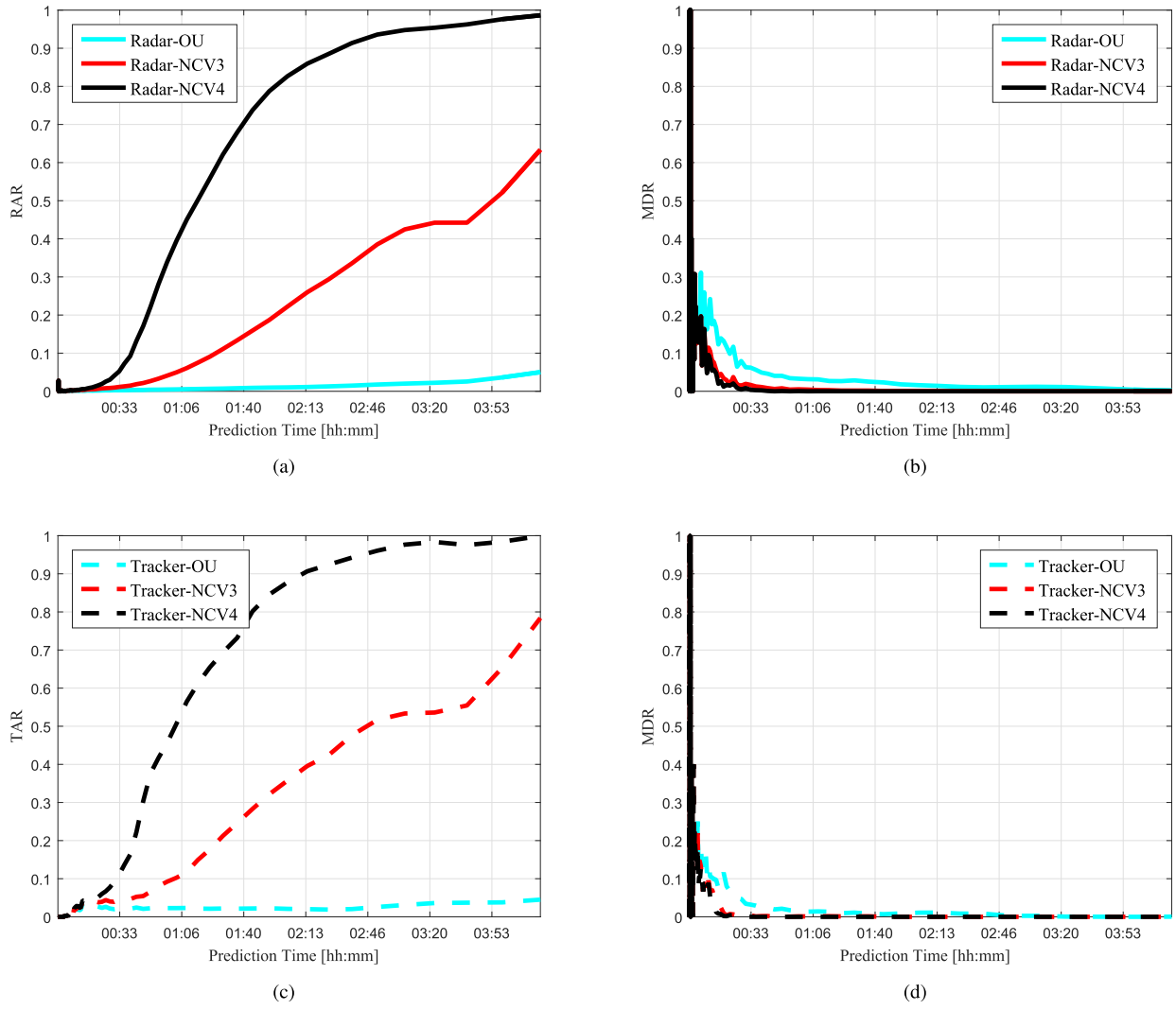


Fig. 8. Performance over prediction time by fixing the gating threshold $\gamma = 25$ (i.e., $g = 5$ sigmas and $P_G = 1$) for HFSW radar and tracker data. (a) RAR. (b) MDR for HFSW radar detections. (c) TAR. (d) MDR for tracker data. The outcomes are obtained averaging on 3034 real trajectories.

TABLE III

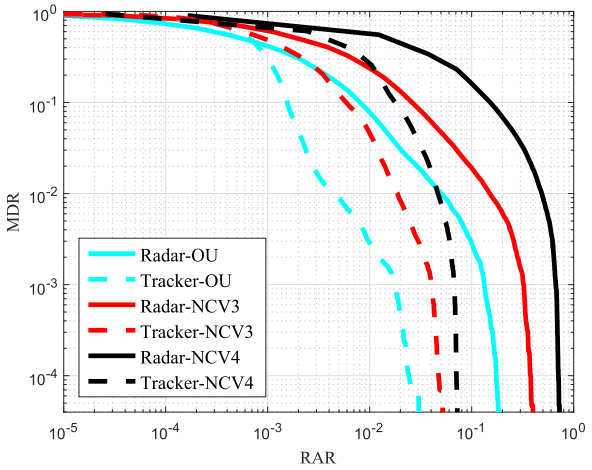
METRIC VALUES FOR SOME FIXED MDRS USING HFSW RADAR DETECTIONS AND JPDA TRACKER DATA. THE BEST RESULTS PER MDR VALUE ARE POINTED OUT IN BOLDFACE

	Metrics	Models	MDR		
			10^{-1}	10^{-2}	10^{-3}
Detections	RAR	OU	**0.008**	**0.051**	**0.135**
		NCV3	0.027	0.150	0.325
		NCV4	0.150	0.488	0.667
Tracks	RAR	OU	**0.001**	**0.004**	**0.018**
		NCV3	0.006	0.019	0.040
		NCV4	0.018	0.046	0.067
	TAR	OU	**0.014**	**0.041**	**0.176**
		NCV3	0.056	0.185	0.441
		NCV4	0.178	0.507	0.733

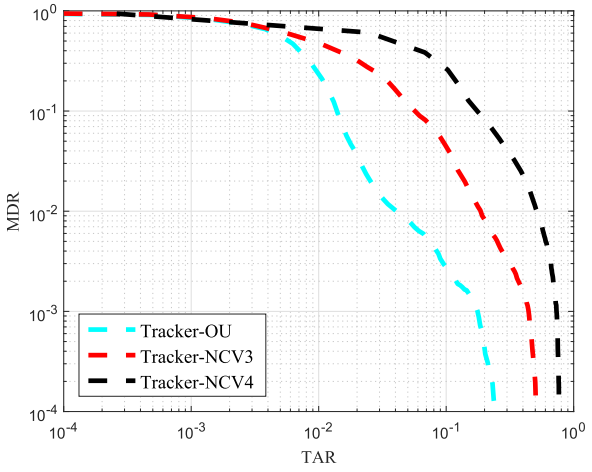
the OU model are apparent with respect to the NCV models [see Fig. 9(b)]. For instance, by fixing the MDR to 10^{-1} (see Table III), the reduction of the TAR for the OU model is about the 120% and about the 170% with respect to the NCV3 model and the NCV4 model, respectively.

Finally, the last analysis is focused on the comparison between the results obtained directly using the HFSW detections and those reached by exploiting the outcomes provided by the JPDA tracker. In order to compare them, we evaluate the RAR even in the case of tracker data (i.e., N_k is, in this case, the total number of HFSW detections in the surveillance area at time k). The results are shown in Fig. 9(a). It can be easily seen that the JPDA tracker's outcomes (dashed lines) surpass those reached by the HFSW data (continuous lines) for all the prediction models. For instance, by fixing MDR to 10^{-1}, all the compared models show a reduction of the RAR around the 140% (see Table III). This reduction is more evident in the case of the NCV4, where it is about the 160%.

2) Gating Procedure Using SAR Data: The assessment using SAR data is described in this section. Three different test cases, described in Section IV-D and denoted by *Test 1*, *Test 2*, and *Test 3*, are provided using three cargo vessels. Due to the reduced number of vessels acquired by the SAR system in these cases and the clutter-free environment, only the MDR

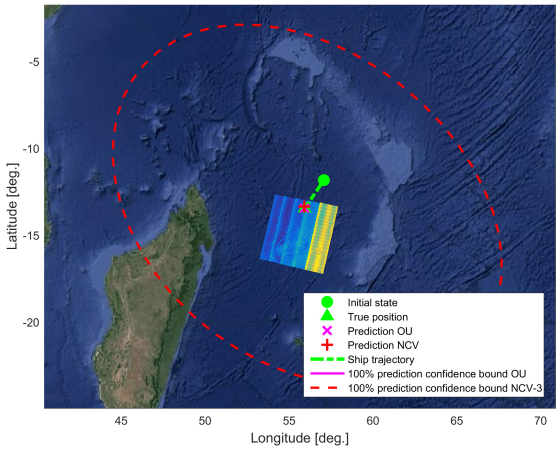


(a)

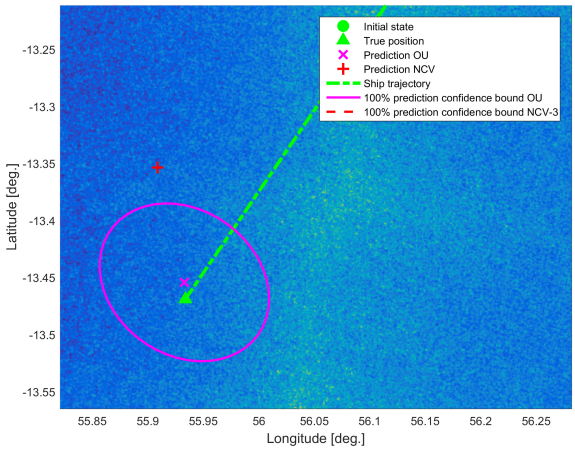


(b)

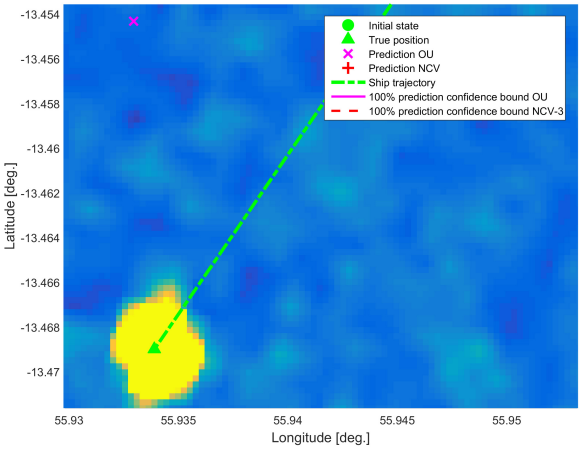
Fig. 9. Misdetection and ARs varying the gating threshold γ (from 0.05^2 to 20.05^2). (a) RAR versus MDR. (b) TAR versus MDR. The outcomes are obtained averaging on 3034 real trajectories.



(a)



(b)



(c)

Fig. 10. SAR test case 1. (a) 1000×800 km^2 overview. (b) 50×40 km^2 close-up. (c) 2.5×2 km close-up. The predictions of the target position after approximately 9 h from the initial state (red dot) are indicated with cross markers. The related 100%-confidence prediction covariance ellipses are plotted with red dashed line (NCV3) and magenta continuous line (OU). The true position is indicated with a green triangle, whereas the AIS ship trajectory is depicted with a green dashed line (produced from ESA remote sensing data).

can be evaluated. Thus, the AR is substituted by the volume of the gate in (25) provided for $\gamma = 1$.

A qualitative analysis is performed first. Fig. 10 depicts the benefits in the target localization with a relevant reduction of the uncertainty area after approximately 9 h of prediction using both the OU and the NCV3 models. It is easy to see that the OU estimation of the target position (magenta cross) almost overlaps the true AIS position (green triangle), whereas NCV3 provides an estimation (red cross), which is several kilometers off the true position. Furthermore, the related 100%-confidence prediction covariance ellipses (plotted in red, NCV3, and magenta, OU) are very different in size (in this qualitative analysis, NCV4 is not even considered, as it would result in an infeasible covariance ellipse, i.e., its size larger than the earth). The OU ellipse size is in fact considerably smaller than that of NCV3. This difference is well captured by the performance metrics introduced in Section III-B2. The gate volume is a measurement of how large the NCV3 gate is with respect to the OU gate (or, equivalently, how much association uncertainty there is in the NCV3 case with respect to the OU case). On the one hand, the MDR is 0 (ideal value) in both the cases. This is due to the fact that both the compared models include in their 100%-confidence prediction covariance

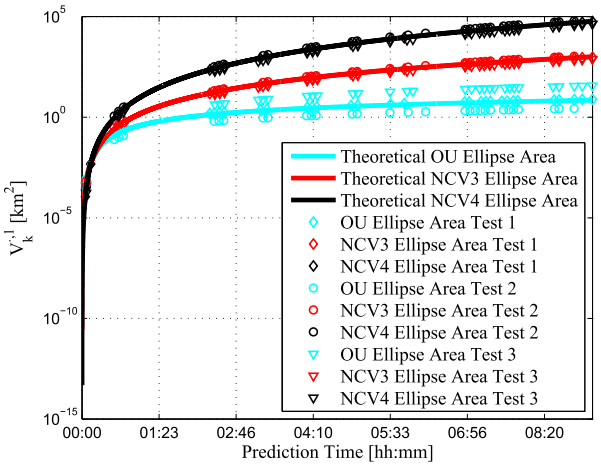


Fig. 11. Ellipse volumes over prediction time fixing $\gamma = 1$ and theoretical curves using average model parameters per ship category using SAR data.

ellipse the SAR detection provided by the predicted target. On the other hand, the gate volume in the OU case is strongly reduced with respect to NCV3. Thus, in this real example, the advantages given by the OU model in the association of the SAR detection with the predicted AIS target are clear and properly assessed using the proposed metrics.

A quantitative assessment is also performed. It is worth pointing out that for the three compared models and for all the three test cases, the MDR is always equal to 0. Thus, the unique feature to discriminate the performance of the three models is the volume of the gate. Fig. 11 depicts the outcomes related to the three test cases and all the compared models. The three different markers define the measured volumes for the three different test cases. The three colors indicate the three compared prediction models (i.e., OU in cyan, NCV3 in red, and NCV4 in black). The solid lines are instead obtained using the theoretical models [8] calculating the volumes for the three long-term prediction models. Average parameters for the cargo category are used to feed the theoretical models and draw the curves. The advantages in using the OU model are clear with a reduction of the gate volume and, thus, a reduction of the uncertainty with respect to the NCV models. The benefits are more and more relevant for high prediction times (long-term prediction).

VI. Conclusion

The association, exploiting both a gating and a data association hypothesis testing procedures, of AIS data with HFSW radar detections, JPDA tracker data, and SAR images has been investigated in order to assess the performance of the OU model against two well-established NCV models for vessel long-term prediction. Performance metrics, which quantify the ability of the gate in including the detection of the predicted target (MDR) without ambiguity (AR), have been proposed and used for performance assessment. The Type I and Type II errors have instead been exploited in the case of the data association hypothesis testing procedure.

Extensive experimental results using real data have demonstrated advantages in using the OU model for vessel long-term prediction instead of the well-established NCV models. Thus, the proper estimation of the predicted covariance matrix given by the OU model enables us to set the Type I error exploiting the data association hypothesis testing procedure. Furthermore, ambiguity reductions up to 180%, with same values of MDRs, have been observed in comparison with the use of the NCV models. Besides, the use of a tracker for HFSW data has been demonstrated helpful to improve the performance. Reductions of the AR up to 160% have been shown, comparing the outcomes with and without the tracker applied to HFSW data. Finally, a considerable reduction of the gating volume provided by the OU model with respect to the NCV models (with the same MDRs) has been observed in the three experiments using SAR data.

References

[1] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.

[2] Y. Bar-Shalom, P. Willett, and X. Tian, *Tracking Data Fusion: A Handbook Algorithms*. Storrs, CT, USA: YBS Publishing, Apr. 2011.

[3] S. Maresca, P. Braca, J. Horstmann, and R. Grasso, "Maritime surveillance using multiple high-frequency surface-wave radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5056–5071, Aug. 2014.

[4] G. Vivone, P. Braca, and J. Horstmann, "Knowledge-based multitarget ship tracking for HF surface wave radar systems," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3931–3949, Jul. 2015.

[5] B. Ristic, B. L. Scala, M. Morelande, and N. Gordon, "Statistical analysis of motion patterns in AIS data: Anomaly detection and motion prediction," in *Proc. 11th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2008, pp. 1–7.

[6] G. Pallotta, M. Vespe, and K. Bryan, "Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction," *Entropy*, vol. 15, no. 6, pp. 2218–2245, Jun. 2013.

[7] M. Vespe, I. Visentini, K. Bryan, and P. Braca, "Unsupervised learning of maritime traffic patterns for anomaly detection," in *Proc. 9th IET Data Fusion Target Tracking Conf. (DFTT)*, May 2012, p. 14.

[8] L. M. Millefiori, P. Braca, K. Bryan, and P. Willett, "Modeling vessel kinematics using a stochastic mean-reverting process for long-term prediction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 5, pp. 2313–2330, Oct. 2016.

[9] L. M. Millefiori, P. Braca, K. Bryan, and P. Willett, "Long-term vessel kinematics prediction exploiting mean-reverting processes," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2016, pp. 232–239.

[10] A. Einstein, "On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat," *Ann. der Phys.*, vol. 17, pp. 549–560, May 1905.

[11] O. E. Barndorff-Nielsen and N. Shephard, "Integrated OU processes and non-Gaussian OU-based stochastic volatility models," *Scandinavian J. Statist.*, vol. 30, no. 2, pp. 277–295, Jun. 2003.

[12] P. G. Blackwell, "Random diffusion models for animal movement," *Ecol. Model.*, vol. 100, nos. 1–3, pp. 87–102, Dec. 1997.

[13] S. Coraluppi and C. Carthel, "A hierarchical MHT approach to ESM-radar fusion," in *Proc. 15th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2012, pp. 677–683.

[14] S. Coraluppi and C. Carthel, "Stochastic data association in multi-target filtering," *Proc. SPIE*, vol. 8393, pp. 83930Q-1–83930Q-7, Mar. 2012.

[15] S. Coraluppi and C. Carthel, "Stability and stationarity in target kinematic modeling," in *Proc. IEEE Aerosp. Conf.*, Mar. 2015, pp. 1–8.

[16] S. Coraluppi and C. Carthel, "Advances in multi-target filtering of evasive targets," in *Proc. IEEE Aerosp. Conf.*, Mar. 2015, pp. 1–8.

[17] S. Coraluppi, C. Carthel, P. Braca, and L. Millefiori, "The mixed Ornstein-Uhlenbeck process and context exploitation in multi-target tracking," in *Proc. 19th Int. Conf. Inf. Fusion (FUSION)*, 2016, pp. 217–224.

[18] L. M. Millefiori, P. Braca, and P. Willett, "Consistent estimation of randomly sampled Ornstein–Uhlenbeck process long-run mean for long-term target state prediction," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1562–1566, Nov. 2016.

[19] L. M. Millefiori, P. Braca, K. Bryan, and P. Willett, "Adaptive filtering of imprecisely time-stamped measurements with application to AIS networks," in *Proc. 18th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2015, pp. 359–365.
[20] D. Gillespie, "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 54, no. 2, pp. 2084–2091, Aug. 1996.
[21] *International Maritime Organization Safety of Life at Sea (SOLAS) Convention*. IMO, ch. 5, Regulation 19, 2014.
[22] K.-W. Gurgel, Y. Barbin, and T. Schlick, "Radio frequency interference suppression techniques in FMCW modulated HF radars," in *Proc. Oceans Eur.*, Aberdeen, Scotland, Jun. 2007, pp. 1–4.
[23] A. Dzvonkovskaya, K.-W. Gurgel, H. Rohling, and T. Schlick, "Low power high frequency surface wave radar application for ship detection and tracking," in *Proc. IEEE Radar Conf.*, Vilnius, Lithuania, Sep. 2010, pp. 627–632.
[24] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," *IEEE Control Syst.*, vol. 29, no. 6, pp. 82–100, Dec. 2009.
[25] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.
[26] P. Braca, S. Maresca, J. Horstmann, R. Grasso, and K. Bryan, "Maritime surveillance with multiple over-the-horizon HFSW radars: An overview of recent experimentation," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 30, no. 12, pp. 4–18, Dec. 2015.
[27] F. D. Zan and A. Monti Guarnieri, "TOPSAR: Terrain observation by progressive scans," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2352–2360, Sep. 2006.
[28] H. Laur, P. Bally, P. Meadows, J. Sanchez, B. Schaettler, E. Lopinto, and D. Esteban, *Derivation of the Backscattering Coefficient $\sigma 0$ in ESA ERS SAR PRI Products*, no. 2, document ES-TN-RE-PM-HL09ESA, ESA, Noordjiwk, The Netherlands, 2002. [Online]. Available: https://earth.esa.int/documents/10174/13019/ers_sar_calibration_issue2_ 5f.pdf

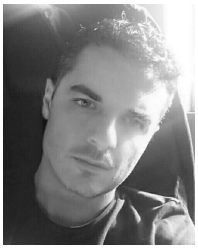
Gemine Vivone (M'17) received the B.Sc. (*cum laude*), the M.Sc. (*cum laude*), and the Ph.D. degrees in information engineering from the University of Salerno, Salerno, Italy, in 2008, 2011, and 2014, respectively.

In 2012, he was a Visiting Researcher with the NATO Undersea Research Center, La Spezia, Italy. In 2013, he was a Visiting Scholar with the Grenoble Institute of Technology, Grenoble, France, conducting his research at the Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique GIPSA-Lab. In 2014, he joined the NATO Science and Technology Organization Center for Maritime Research and Experimentation, La Spezia, as a Scientist. He is currently with the Department of Information Engineering, Electrical Engineering and Applied Mathematics, University of Salerno. His research interests include statistical signal processing, detection of remotely sensed images, data fusion, and tracking algorithms.

Dr. Vivone serves as a referee for several journals, such as the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He is currently a Lead Guest Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS Special Stream "Advances in image processing for very high resolution optical remote sensing images." He was a recipient of the Symposium Best Paper Award at the 2015 IEEE International Geoscience and Remote Sensing Symposium.

Leonardo M. Millefiori (S'12–M'14) received the B.Sc. degree in aerospace information engineering and the M.Sc. degree (*summa cum laude*) in communication engineering with a focus on radar systems and remote sensing from the Sapienza University of Rome, Rome, Italy, in 2010 and 2013, respectively.

In 2013, he was a Visiting Researcher at the NATO Science and Technology Organization Center for Maritime Research and Experimentation, La Spezia, where he joined the Research Department under the Maritime Security Program as a Scientist in 2014. His research interests include target motion modeling, statistical signal processing, target tracking, and data fusion and radar systems.



Paolo Braca (M'14–SM'17) received the Laurea degree (*summa cum laude*) in electronic engineering and the Ph.D. degree (highest rank) in information engineering from the University of Salerno, Salerno, Italy, in 2006 and 2010, respectively.

In 2009, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT, USA. From 2010 to 2011, he was a Post-Doctoral Associate with the University of Salerno. In 2011, he joined the NATO Science and Technology Organization Center for Maritime Research and Experimentation, La Spezia, as a Scientist with the Research Department. He has co-authored about 100 publications in international scientific journals and conference proceedings. His research interests include statistical signal processing with an emphasis on detection and estimation theory, wireless sensor network, multiagent algorithms, target tracking and data fusion, adaptation and learning over graphs, and distributed radar (sonar) processing.

Dr. Braca was awarded the National Scientific Qualification (Abilitazione Scientifica Nazionale) to function as an Associate Professor in Italian universities by a qualification committee of professors, the Italian Ministry of Education, Universities and Research, in 2017. He was a recipient of the Best Student Paper Award (first runner-up) at the 12th International Conference on Information Fusion in 2009. He serves as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, the *ISIF Journal of Advances in Information Fusion*, and the *EURASIP Journal on Advances in Signal Processing*. In 2017, he was a Lead Guest Editor of the Special Issue on "Sonar Multi-Sensor Applications and Techniques" in *IET Radar Sonar and Navigation*. He served as an Associate Editor of the IEEE SIGNAL PROCESSING MAGAZINE (E-Newsletter) from 2014 to 2016. He is in the Technical Committee of the major international conferences in the field of signal processing and data fusion.

Peter Willett (F'03) received the B.A.Sc. degree in engineering science from the University of Toronto, Toronto, ON, Canada, in 1982, and the Ph.D. degree from Princeton University, Princeton, NJ, USA, in 1986.

Since 1986, he has been a Faculty Member at the University of Connecticut, Storrs, CT, USA, where he has been a Professor since 1998. He has authored in the areas of change/abnormality detection, optical pattern recognition, communications, and industrial/security condition monitoring. His research interests include statistical signal processing, detection, machine learning, data fusion, and tracking.

Dr. Willett was the Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS (2014–2016) and the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS (2006–2011). He was the Vice President for publications for AESS (2012–2014). He was the Chair of the IEEE SPS SAM-TC and is an AESS BoG Member. He is the General Co-Chair of the IEEE/ISIF Fusion Conference in Florence in 2006, in Cologne in 2008, and in Chicago in 2011, and the IEEE SPS SAM in Sheffield in 2018.

Document Data Sheet

Security Classification		Project No.
Document Serial No.	**Date of Issue**	**Total Pages**
CMRE-PR-2019-055	June 2019	14 pp.

Author(s)

Gemine Vivone, Leonardo M. Millefiori, Paolo Braca, Peter Willett

Title

Performance assessment of vessel dynamic models for long-term prediction using heterogeneous data

Abstract

Ship traffic monitoring is a foundation for many maritime security domains, and monitoring system specifications underscore the necessity to track vessels beyond territorial waters. However, vessels in open seas are seldom continuously observed. Thus, the problem of long-term vessel prediction becomes crucial. This paper focuses attention on the performance assessment of the Ornstein-Uhlenbeck (OU) model for long-term vessel prediction, compared with usual and well-established nearly constant velocity (NCV) model. Heterogeneous data, such as automatic identification system (AIS) data, high-frequency surface wave radar data, and synthetic aperture radar data, are exploited to this aim. Two different association procedures are also presented to cue dwells in case of gaps in the transmission of AIS messages. Suitable metrics have been introduced for the assessment. Considerable advantages of the OU model are pointed out with respect to the NCV model.

Keywords

Automatic identification system (AIS), high-frequency surface wave (HFSW) radar, long-term prediction, maritime surveillance, nearly constant velocity (NCV) model, Ornstein–Uhlenbeck (OU) process, synthetic aperture radar (SAR)

Issuing Organization

NATO Science and Technology Organization
Centre for Maritime Research and Experimentation

Viale San Bartolomeo 400, 19126 La Spezia, Italy

[From N. America:
STO CMRE
Unit 31318, Box 19, APO AE 09613-1318]

Tel: +39 0187 527 361
Fax:+39 0187 527 700

E-mail: library@cmre.nato.int