SCIENCE AND TECHNOLOGY ORGANIZATION
CENTRE FOR MARITIME RESEARCH AND EXPERIMENTATION

Reprint Series **CMRE-PR-2019-065**

Model performance assessment for long-term vessel prediction using HFSW radar data

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

June 2019

Originally published in:

2017 IEEE Radar Conference (RadarConf), 08-12 May 2017, Seattle, WA, USA, pp. 243-247, doi: 10.1109/RADAR.2017.7944205

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.

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.

Model Performance Assessment for Long-Term Vessel Prediction Using HFSW Radar Data

Gemine Vivone*, Leonardo M. Millefiori*, Paolo Braca* and Peter Willett†
* NATO STO Centre for Maritime Research and Experimentation, La Spezia, Italy
†Department of Electrical and Computer Engineering, University of Connecticut, Storrs, USA

***Abstract*—Vessels in open seas are seldom continuously observed. Thus, the problem of long-term vessel prediction becomes crucial. This paper focuses its attention on the performance assessment of the Ornstein-Uhlenbeck target motion model comparing it with the well-established nearly constant velocity model. A gating association procedure and proper performance metrics are introduced to assess the performance using automatic identification system and high-frequency surface wave radar data.**

I. Introduction

Ship traffic monitoring is a foundation for several maritime security domains and modern monitoring system specifications reflect the need for a continuous ability to track vessels at sea. However, vessels in open seas are seldom continuously observed, and these coverage gaps represent a high risk in terms of safety at sea.

Thus, the problem of long-term vessel state estimate and prediction becomes crucial, but only few items in literature can be found that address this issue, see e.g. [1]–[4]. The nearly constant velocity (NCV) model [5], [6], which describes non-maneuvering target dynamics with a velocity that is perturbed by a white noise process, is usually exploited in the target tracking literature, see e.g. [7], with a prediction step that usually refers to the very near future. On the contrary, in the case of non-maneuvering targets, as is often the case for the maritime traffic, long-term prediction can be tackled in a better way exploiting Ornstein-Uhlenbeck (OU) models [4].

The OU model is popular in various and heterogeneous scientific fields, see e.g. [8]–[10]. It was firstly introduced in physics [8] 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 towards a central location, with a greater attraction when the process is further away from the center. In the tracking literature the OU model has been discussed mostly notably by Coraluppi et al., see e.g. [11].

In this paper, a gating procedure for associating automatic identification system (AIS) and high-frequency surface wave (HFSW) radar data is presented. Two metrics are proposed to assess the performance in conjunction with the gating procedure. The OU and NCV model are compared. The association advantages, measured by the proposed performance metrics, of the OU model for vessel long-term prediction are pointed out by leveraging one month of data recorded in the Ligurian Sea by a HFSW radar located in Palmaria, Italy, during 2009. An extensive experimental analysis shows a considerable reduction in ambiguity—with an unchanging misdetection rate—when the OU model is used instead of the nearly constant velocity one. This analysis demonstrates the validity and the better accuracy of the OU model with respect to the NCV model on the specific problem of association among heterogeneous sources of information when long-term target prediction is performed.

The rest of the paper is organized as follows. Sect. II introduces the gating and the performance metrics for the assessment. Experimental results are presented in Sect. III. Finally, conclusions are drawn in Sect. IV.

II. Performance Assessment for Long-term Prediction Models

This section is devoted to the introduction of the gating procedure for HFSW radar data. Furthermore, the performance metrics used for the assessment of the long-term prediction models are presented.

A. Gating of HFSW Radar Data

A multidimensional gating procedure is used for association purposes between the predicted AIS and the HFSW radar data. A measurement in the gate, while not guaranteed to have originated from the target the gate pertains to, is a valid association candidate. Hence, the name of *validation region* or *association region*.

We assume that the HFSW radar data $\boldsymbol{\delta}_k$ at frame k coming from a target t is Gaussian distributed with mean $\hat{\boldsymbol{x}}^t_{k|k-1}$ and covariance $\mathbf{P}^t_{k|k-1}$. Thus, its gate $\mathcal{V}^{t,\gamma}_k$ at frame k is defined as

$$\mathcal{V}^{t,\gamma}_k = \left\{ \boldsymbol{\delta} : \left[\boldsymbol{\delta} - \hat{\boldsymbol{x}}^t_{k|k-1}\right]^{\mathrm{T}} \left(\mathbf{P}^t_{k|k-1}\right)^{-1} \left[\boldsymbol{\delta} - \hat{\boldsymbol{x}}^t_{k|k-1}\right] \leq \gamma \right\}, \quad (1)$$

with a probability depending on the threshold γ. The distance metric in (1) is also called Mahalanobis distance, which is a multi-dimensional generalization of the idea of measuring how many standard deviations away a point $\boldsymbol{\delta}$ is from the mean $\hat{\boldsymbol{x}}^t_{k|k-1}$ of the related Gaussian distribution. Under the above-mentioned hypothesis, this quadratic form is chi-square distributed with n_x degrees of freedom, n_x being the dimension of the positional part of the target state. Thus, the probability P_G that $\boldsymbol{\delta}_k$ is in $\mathcal{V}^{t,\gamma}_k$ is defined as $P\left\{\boldsymbol{\delta}_k \in \mathcal{V}^{t,\gamma}_k\right\}$, which depends on both n_x and γ [6]. Finally, it is worth remarking

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

Remark: Remember that when the association between the predicted AIS position and the HFSW radar data is performed, the uncertainties of the measurements provided by the radar systems have to also be taken into account (it has an impact, in particular, on short-term predictions). Thus, considering that the errors in range and bearing (azimuth) are zero-mean Gaussian distributed and considering the conversion between polar and Cartesian coordinates, we have that the noise covariance matrix, $\mathbf{R}(\boldsymbol{\zeta}_k)$, for a measurement $\boldsymbol{\zeta}_k$ at frame k in polar coordinates can be defined as in [12, Sect. II-C]. Since the radar measurement noises are conditionally independent with respect to the AIS prediction of the 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 $\mathbf{P}^t_{k|k-1} = \Sigma^t_{k|k-1} + \mathbf{R}(\boldsymbol{\zeta}_k)$, where $\Sigma^t_{k|k-1}$ is the prediction covariance matrix for target t at frame k.

Remark: $\hat{\boldsymbol{x}}^t_{k|k-1}$ and $\Sigma^t_{k|k-1}$, for the compared dynamic models (i.e. OU and NCV), are given by the prediction formulas in [4]. The model parameters are estimated in a batch way, starting from the acquired AIS data, using the procedure proposed in [4, Sect. III].

B. Performance Metrics

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 its measurement provided by another acquisition system.

Thus, denote $\mathcal{D}_k = \left\{\boldsymbol{\delta}^i_k\right\}_{i=1}^{N_k}$ the set of the N_k heterogeneous data at frame k and denote also $\mathcal{D}^{t,\gamma}_k = \mathcal{D}_k \cap \mathcal{V}^{t,\gamma}_k$ the subset of $\mathcal{D}_k$ containing only the heterogeneous data at frame k validated using the gating $\mathcal{V}^{t,\gamma}_k$ in (1). The performance indexes are as follows

- **Radar Ambiguity Rate** (RAR). The RAR using a threshold γ for the target t at frame k, i.e. $RAR^{t,\gamma}_k$, is defined as

$$RAR^{t,\gamma}_k = \frac{\left|\mathcal{D}^{t,\gamma}_k\right| - \mathbf{1}_{\mathcal{V}^{t,\gamma}_k}(\boldsymbol{\delta}^t_k)}{N_k - \mathbf{1}_{\mathcal{A}}(\boldsymbol{\delta}^t_k)}, \tag{2}$$

  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, N_k is the total number of radar detections in $\mathcal{A}$ at frame k, and $\boldsymbol{\delta}^t_k$ is the measurement originated by target t at frame 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.
  The overall RAR index is obtained by averaging $RAR^{t,\gamma}_k$ on the number of targets and frames (RAR^{γ}). Note that the overall RAR index is only function of γ (gate threshold). This index assumes values in the range $[0, 1]$ and the ideal value is 0, i.e. no ambiguity is present.
- **Misdetection rate** (MDR). The MDR using a threshold γ for a target t at frame k, i.e. $MDR^{t,\gamma}_k$, is defined as

$$MDR^{t,\gamma}_k = 1 - \mathbf{1}_{\mathcal{V}^{t,\gamma}_k}(\boldsymbol{\delta}^t_k), \tag{3}$$

  namely $MDR^{t,\gamma}_k = 1$ if the measurement provided by target t at frame k ($\boldsymbol{\delta}^t_k$) is in the set $\mathcal{V}^{t,\gamma}_k$, otherwise $MDR^{t,\gamma}_k = 0$.
  Again, the overall MDR index is obtained by averaging $MDR^{t,\gamma}_k$ on the number of targets and frames, MDR^{γ}. Note that the overall MDR index is only function of γ (gate threshold). The index can assume values in the range $[0, 1]$, where the ideal value is 0, i.e. the measurement originated by the target is always in the defined gate.

Remark: It is worth remarking that RAR = 0 and MDR = 1 can trivially be obtained when $\gamma = 0$, or, equivalently, RAR = 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 misdetection probability.

In order to give an idea of how these metrics work, an illustrative example using real data is described in Sect. III, see Fig. 1. This clarifies how the two proposed performance metrics reach the above-mentioned goals of measuring the association uncertainty together with the ability of identifying the correct association. The association ambiguity is quantified by the RAR index, whereas the identification of the correct association is considered using the MDR index. Thus, the greater the gating region, the greater the probability to retain the correct 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 to the NCV model, i.e. a reduced uncertainty or ambiguity. At the same time, this gating region is large enough to retain the correct association.

III. Experimental Results

The assessment of the long-term prediction models is provided in this section using one month of AIS data and detections acquired by a HFSW radar system located in Palmaria, Italy, in 2009.

A qualitative analysis is performed first. Fig. 1 depicts the benefits in the target localization with a relevant reduction of the uncertainty area after approximately 2 hours and 30 minutes of prediction using both the OU and the nearly constant velocity 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, NCV, and cyan, OU) are very different in size. Indeed, the OU size

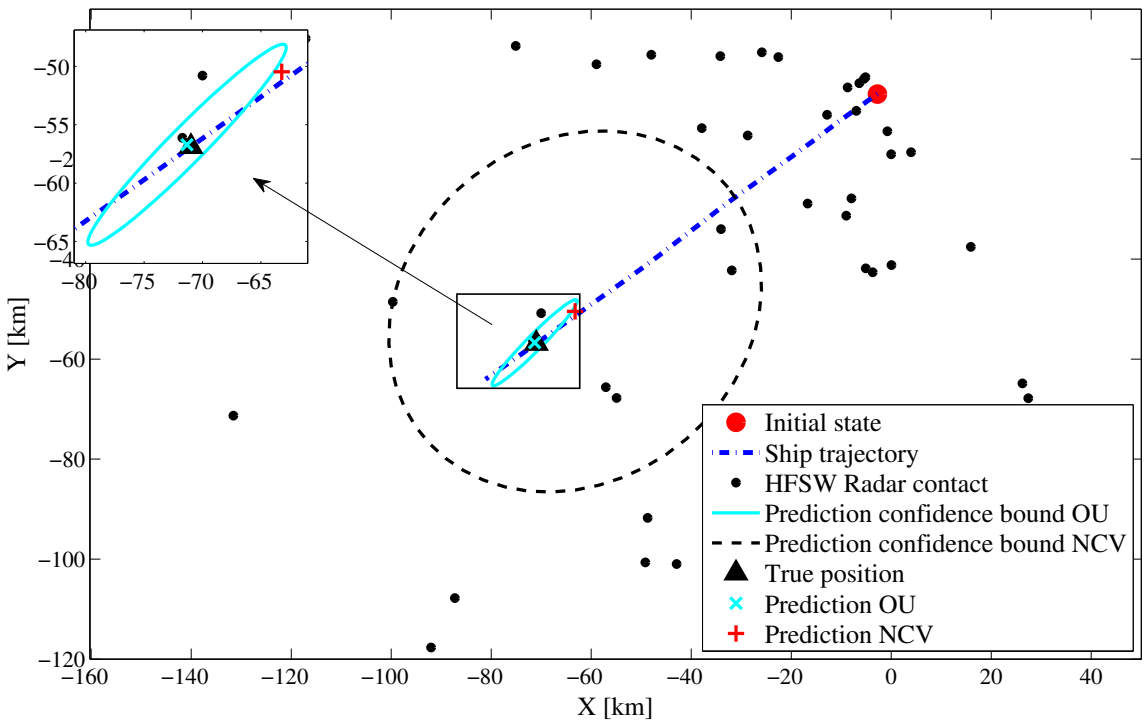
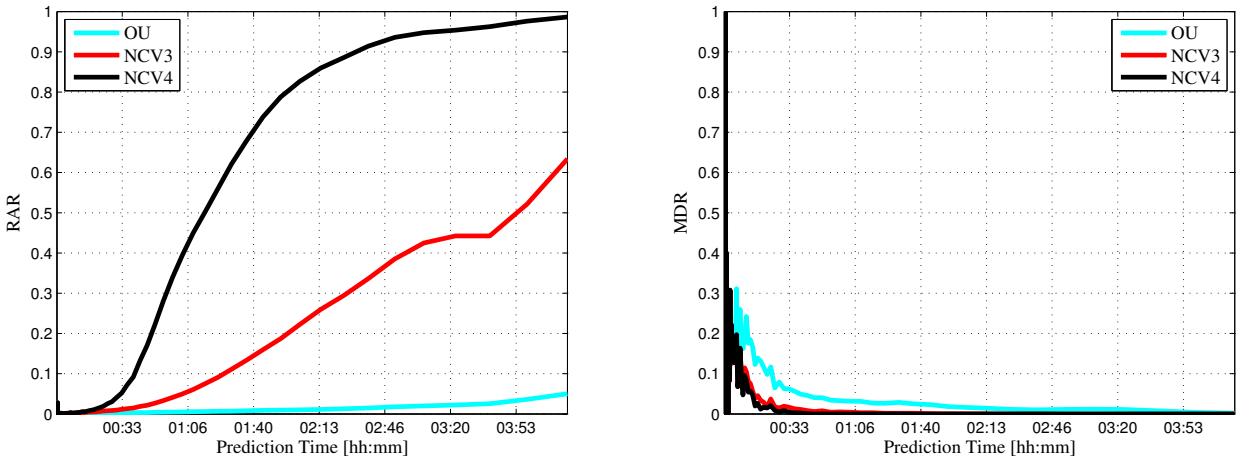


Fig. 1. A portion of one of the trajectories acquired by the Palmaria HFSW radar (converted into projected coordinates) is depicted. The predictions of the target position after approximately 2 hours and 30 minutes from the initial state are indicated with cross markers. The related 100%-confidence prediction covariance ellipses are plotted with black dashed line (NCV) and cyan continuous line (OU). The true position is indicated with a black triangle, whereas the HFSW radar contacts are plotted with black dots.

is considerably smaller than that of NCV. This difference is well captured by the performance metrics introduced in Sect. II-B. In other words, the RAR can be seen as an indirect measurement of how large the NCV gate is with respect to the OU gate (or, equivalently, how much more uncertainty there is in the NCV case with respect to the OU case). On one hand, the misdetection rate is 0 (ideal value) in either case. 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 other hand, the ambiguity rate 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. 6 detections in the gate) in the case of NCV. Thus, in this simple real example, the advantages given by the OU model in associating the HFSW detections with the predicted AIS target are clear and properly assessed using the proposed metrics.

A quantitative assessment is also performed using all the data (one month) acquired by the HFSW radar located in Palmaria. The evaluated metrics are the misdetection rate, MDR, and the radar ambiguity rate, RAR, proposed in Sect. II-B. In order to assess the performance, 3034 trajectories are considered that can be considered satisfying the non-maneuvering hypothesis under the use of the OU model. The OU model [4] has been compared with two well-established nearly constant velocity models, i.e. the so-called NCV3 and NCV4 discussed in [4], [5], widely used in the literature for target prediction [5]. It is worth to remark that different process noise assumptions bring to different NCV models. Thus, NCV4 differs from NCV3 for the scaling law of the errors (i.e. the former has a position error that scales proportionally to k^4, whereas the latter as k^3). The validation is performed considering simulated gaps in AIS data, which define the x-axis (prediction time) in Fig. 2. This procedure enables us to properly identify the true association between AIS and HFSW data (defining a ground-truth) 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 prediction time. This is obtained by fixing the gate threshold γ and averaging the indexes for all the predicted targets in the surveillance area on contiguous prediction intervals. The selected gate threshold γ is 25 (i.e. $g = 5$ sigmas, $P_G = 1$), a typical value used for gating radar data. Fig. 2 shows a significant increase of the RAR over prediction time. This is much more apparent for the nearly constant velocity models, see Fig. 2(a). This increase is attributable to the growing of the uncertainties of the compared models when the prediction time increases. Thus, the RAR values obtained with the OU model are always better than those reached by the nearly constant velocity models. At the same time, the best outcome provided by the OU on the RAR comes at a cost of an increase of the MDR, see Fig. 2(b), because NCV models have a very large uncertainty



(a) (b)

Fig. 2. Performance over prediction time by fixing the gating threshold $\gamma = 25$ (i.e. $g = 5$ sigmas, $P_G = 1$): (a) Radar ambiguity rate; (b) Misdetection rate.

TABLE I
RAR VALUES FOR SOME FIXED MDRS. THE BEST RESULTS PER MDR VALUE ARE POINTED OUT IN BOLDFACE.

Models	MDR 10^{-1}	10^{-2}	10^{-3}
OU	**0.008**	**0.051**	**0.135**
NCV3	0.027	0.150	0.325
NCV4	0.150	0.488	0.667

and detections are seldom missed. Thus, a further analysis is required to better point out and quantify the advantages of using the OU model for vessel long-term prediction.

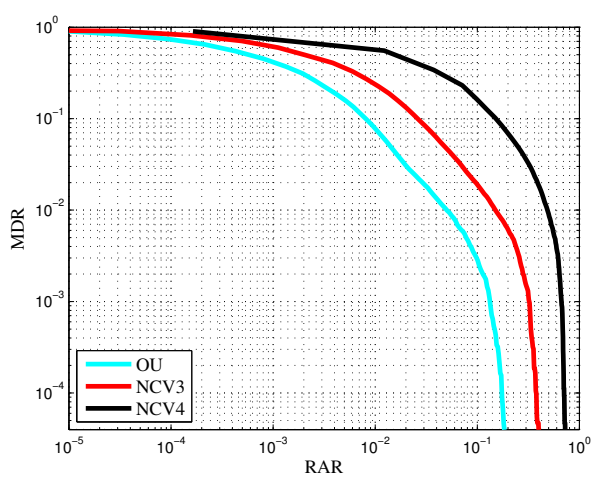


Fig. 3. Misdetection rate and radar ambiguity rate varying the gating threshold γ (from 0.05^2 to 20.05^2).

This further analysis is related to the variation of the gate threshold γ to depict the MDR against the RAR averaged over time. The range of variation for γ is $[0.05^2,\ 20.05^2]$. Fig. 3 clearly points out the advantages of the OU model with respect to the nearly constant velocity models. Tab. I summarizes the values of the RAR metric for some fixed values of the MDR for the three compared models. For very high values of the MDR, the models perform similarly, see Fig. 3. 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 70% and about the 95% 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 90%.

IV. CONCLUSION

The association (exploiting a gating procedure) of AIS data with HFSW radar data has been investigated in order to assess the performance of the OU model against two well-established NCV models for vessel long-term prediction. Two performance metrics, which quantify the ability of the gate of including the detection of the predicted target (misdetection rate) without ambiguity (ambiguity rate), have been proposed and used for performance assessment. The experimental results using real AIS and HFSW data have demonstrated advantages of the OU model in comparison with the two NCV models. Ambiguity reductions of the OU model with respect to the NCV ones up to the 95%, with comparable values of misdetection rates, have been observed.

REFERENCES

[1] 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. of the* 11th *Intern. Conf. on Inform. Fusion (FUSION)*, Jul. 2008.

[2] 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.

[3] M. Vespe, I. Visentini, K. Bryan, and P. Braca, "Unsupervised learning of maritime traffic patterns for anomaly detection," in *Data Fusion Target Tracking Conference (DF TT 2012): Algorithms Applications, 9th IET*, May 2012.

[4] 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. (in Press)*, 2016.

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

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

[7] 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.

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

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

[10] P. Blackwell, "Random diffusion models for animal movement," *Ecological Modelling*, vol. 100, no. 1, pp. 87–102, Dec. 1997.

[11] S. Coraluppi and C. Carthel, "A hierarchical MHT approach to ESM-radar fusion," in *Proc. of the* 15th *Intern. Conf. on Inform. Fusion (FUSION)*, Jul. 2012.

[12] G. Vivone, P. Braca, K. Granström, A. Natale, and J. Chanussot, "Converted measurements random matrix approach to extended target tracking using X-band marine radar data," in *Proc. of the* 18th *Intern. Conf. on Inform. Fusion (FUSION)*, Jul. 2015.

Document Data Sheet

Security Classification		Project No.
Document Serial No.	**Date of Issue**	**Total Pages**
CMRE-PR-2019-065	June 2019	5 pp.

Author(s)

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

Title

Model performance assessment for long-term vessel prediction using HFSW radar data

Abstract

Vessels in open seas are seldom continuously observed. Thus, the problem of long-term vessel prediction becomes crucial. This paper focuses its attention on the performance assessment of the Ornstein-Uhlenbeck target motion model comparing it with the well-established nearly constant velocity model. A gating association procedure and proper performance metrics are introduced to assess the performance using automatic identification system and high-frequency surface wave radar data.

Keywords

Radar, predictive models, measurement, logic gates, indexes, uncertainty

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