Incorporating Aircraft Kinematics and Radar Cross Section into the Performance Prediction of Air Surveillance

Juha Jylhä, Marja Ruotsalainen, Minna Väilä, Henna Perälä
Computing Sciences
Tampere University
Tampere, Finland
juha.jylha@tuni.fi, marja.ruotsalainen@tuni.fi, minna.vaila@tuni.fi, henna.perala@tuni.fi

***Abstract*—The evolution of modern radar is heading toward a networked, multifunctional, adaptive, and cognitive system. The network of software-controllable fast-adapting radars follows a highly complex control and operation logic. It is not straightforward to assess its instantaneous capability to detect, track, and recognize targets. To be able to predict or optimize the system performance, one has to understand its behavior not only on a general level, but also in various operating conditions and considering the target behavior and properties accurately. In this paper, we propose the fusion of radar and tracker recordings with an extensive database of cooperative aircraft navigation recordings and radar cross section data to assess and learn the performance measures for the air surveillance. The main contribution of this paper is the incorporation of the aircraft kinematics, orientation, and radar cross section into an automated measurement-based analysis. We consider the employment of the measurement-based metrics and machine learning in the performance prediction. Simulations and experiments with real-life data demonstrate the feasibility and potential of the proposed concept.**

***Index Terms*—radar, radar cross sections, artificial intelligence, machine learning, system analysis and design, systems modeling**

I. Introduction

The performance of multi-radar-based air surveillance is influenced by the operating conditions, which include several factors. They can be divided into the following categories:

1) the parameters of the radar system, which describe the general capability of the radars [1]–[3],
2) the description of the sensor fusion system, which provides the general capability to employ the information collected by the radar network [4]–[6],
3) the properties of the environment [1], [4], and
4) the properties of the targets [7]–[11].

Predicting the air surveillance performance incorporates all these four categories of factors. The concept of the operating conditions in performance modeling, as well as the subdivision and terminology, are discussed in more detail in [12].

Our previous work covers several aspects of the operating conditions related to the air surveillance. The radar cross section (RCS) describes the electromagnetic scattering of a radar target. Our RCS simulation procedure that utilizes three-dimensional surface models of the targets as the source information is discussed in [7]. We have also studied the estimation of the RCS distribution for the noncooperative target recognition [8], [9] and the radar performance analysis [9]–[11]. For modeling the air surveillance performance, we have proposed a machine learning framework [3], [6], which divides the performance analysis into two phases. The radar model [3] provides several performance measures (PM) that characterize the radars. The radar PMs include, for example, the probability of detection and the measurement accuracies in azimuth and range that have been estimated from measurement data. In our framework [3], the radar PMs are utilized by the tracker model [6] to provide the quality levels of the sensor fusion, also in before-unseen situations (to be used e.g. in real-time performance prediction); the tracker quality levels are defined using discretized tracker PMs. The tracker model learning is implemented by using a rule learner algorithm to find the relationship between the radar PMs and the tracker quality levels; the tracker model is a rule-based classifier that outputs the quality levels.

In this paper, we extend our machine learning framework and propose a concept to use target kinematics and RCS information for more detailed assessment of the air surveillance performance based on a massive amount of recorded data. With the term "assessment" we refer to the calculation of the PMs from the measurements, and with the term "prediction" we refer to the use of the models to rapidly predict the PMs for various needs. The assessment is needed as the ground truth information for learning and verifying the prediction model. Our PM assessment produces the PMs at different discretized spatial locations (volumetric cells, voxels) during the time covered by the data recordings. An example of the produced PM is shown in Fig. 1. It illustrates the basic principles of the procedure proposed in this paper for assessing the detection threshold as the minimum detectable RCS. The main contribution of this paper is the improved target description for the performance assessment, which is achieved by the assessment of the target flight state and the adoption of the RCS processing concepts similar to ones described in [7]–[11]. A secondary contribution presented in this paper is the enhanced modeling and prediction of the performance, as the

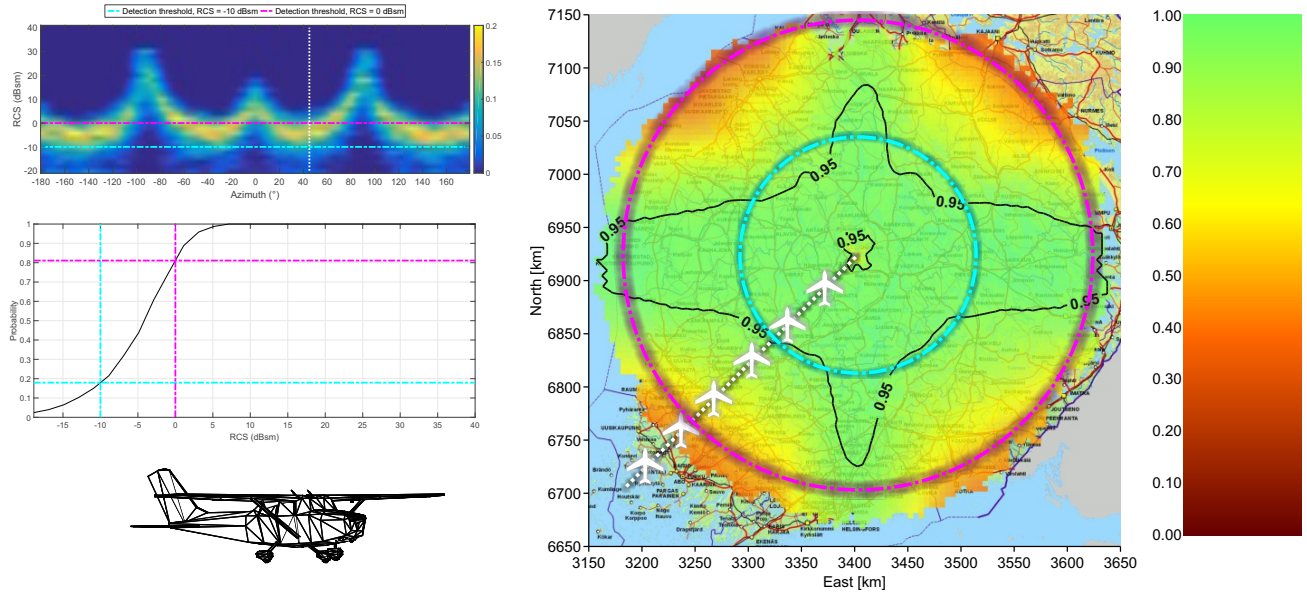


Fig. 1. As an example of the radar PM, the simulated probability of detection of a Cessna 172 aircraft is visualized by color in the picture on the right-hand side. The radar is located in the middle of the picture and the Cessna 172 is heading north throughout the picture. The effect of the target RCS is visible in the probability of detection: e.g. the detection range in the nose aspect is longer than in the tail aspect. The simulated RCS (in decibel square meter, dBsm) is compiled in a histogram form visualized in the top left picture at the target elevation angle of $0°$ which corresponds to the horizontal plane of the target. Color represents the RCS likelihood. In the horizontal axis, the target azimuth of $0°$ represents the nose aspect. The target appearance at the azimuth of $45°$ is marked with the white dashed lines. The Cessna 172 at this aspect is illustrated in the bottom left corner of this figure. The cumulative density function of the RCS likelihood at the azimuth of $45°$ is presented in the lower left diagram. The probability of detection can be coarsely determined as a complement of the value at the point of the detection threshold: 0.82 at the RCS threshold of –10 dBsm and 0.19 at the threshold of 0 dBsm. The colored lines indicate the thresholds in all the three pictures.

more detailed description of the target improves the tracker model learning [3], [6]. As indicated by Fig. 1, we aim at assessing and predicting the PMs as a function of the spatial domain, including also separate altitude layers. This allows for instance the analysis of the multi-radar configuration as well as the terrain and atmosphere, the influence of which varies in the spatial domain. Potential applications for the proposed concepts include the decision-making related to the use of the air surveillance system, real-time algorithms related to adaptive or cognitive radar, performance-driven sensing [12], adaptation to the varying target RCS in tracking, and noncooperative target recognition.

The structure of this paper is as follows. Section II presents the different sources of information, their fusion in the calculation of the PMs, and the use of the target kinematics and the RCS information. Section III considers the learning of the tracker model. Section IV presents the experiments we have conducted with real data. Finally, Section V is the discussion and Section VI gives our conclusions.

II. Performance Assessment Using Measured Data and Library of Targets

We perform the assessment of the PMs using essentially the same approach we previously proposed in the context of our machine learning framework [3], [6]. In this section, we describe our further developments in the assessment of the PMs: the separate performance assessments for different aircraft flight states and the use of the RCS information for estimating the detection threshold. The target information is precomputed and stored in a target library. The library should include simulated or measured RCS information for each aircraft type essential in the analysis. As well the information concerning kinematics that is characteristic to different aircraft types can be included in the library. We first introduce the general use of the measured data, and then present the estimation of the aircraft flight states and the procedure for the estimation of the detection threshold.

A. Fusion of the Different Data Recordings

Our fundamental idea [3] is to measure both, the ground truth "representation" of what has happened in the air surveillance scene, and the "representation" of how the air surveillance system acquires the scene. The sophisticated comparison between these two representations of the same phenomena provides an assessment for the performance of the air surveillance system. We include the following three types of measured data to establish the assessment:

1) aircraft navigation recordings (flight trajectories),
2) radar recordings (observations), and
3) sensor fusion recordings (tracks).

Vast amounts of data are needed to produce an adequate and reliable assessment of the performance. The cause-and-effect relationship of the different data recordings are visualized in Fig. 2: the air surveillance scene affects the radar network performance which affects the sensor fusion performance. The navigation recordings provide the flight paths of the aircraft

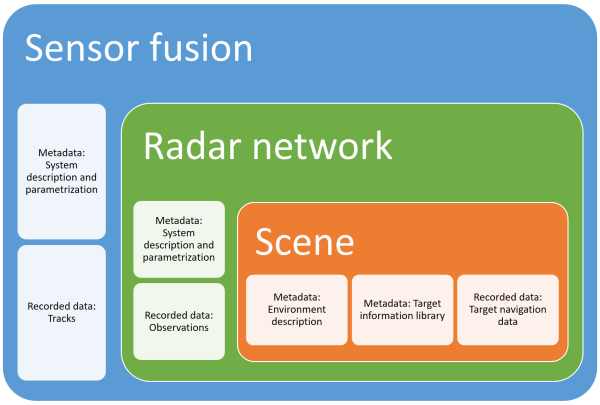


Fig. 2. The data sources for assessing the performance of the radar network and the sensor fusion. The data flow and the PMs are illustrated in Fig. 3

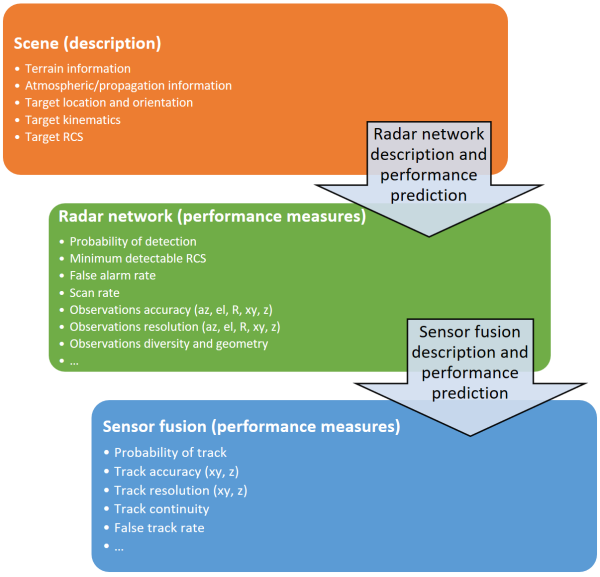


Fig. 3. Some typical PMs to be assessed or predicted. The radar model predicts the radar PMs based on the description of the scene and the radar network. The sensor fusion model predicts the quality levels of the sensor fusion by using the radar PMs as input; in the sensor fusion model used in this paper, the quality levels have been defined as classes discretized based on the sensor fusion PMs.

and information about their kinematic behavior. They determine the ground truth about the aircraft position (and possibly orientation) as a function of time. The radar recordings allow the analysis of the realized likelihoods and accuracies of the radar observations, i.e. the radar PMs. Correspondingly, the sensor fusion recordings provide information on the realized tracks for determining the tracking PMs.

In addition to the three data sources we listed, the performance assessment requires the specification of some additional information for the analysis: the measurement setup, the environment, and the targets. They are indicated as the metadata in Fig. 2. The measurement setup describes the air surveillance system including the sensor configuration (i.e. the locations and the operating modes of the sensors) and the parametrization of the sensor fusion. The environment includes the information concerning the surrounding terrain and the atmosphere such as the weather. For instance, when studying the influence of the weather on the radar PMs, the weather needs to be characterized adequately and the radar recordings need to include different weather conditions. Besides the navigation recordings, the description of the targets requires information about RCS.

For the assessment and prediction of the radar PMs, the RCS data in the target library has a significant role. The measured RCS of the radar observations should correspond to the reference RCS distributions adequately; the reference RCS data in the library needs to be of high quality in order to prevent biased and erroneous prediction of the PMs. We propose formulating the RCS histograms dependent on the aspect angle, illustrated in Fig. 1, using the concepts of [8]–[11]. All the RCS simulations in this paper have been conducted based on physical optics and the physical theory of diffraction using three-dimensional surface models of the targets as source information.

The fusion of the surveillance system data with the scene data is similar to our previous work described in [3]. It is based on the calculation of the similarity measures between the aircraft trajectories (the aircraft navigation data) and the tracks provided by the sensor fusion. When a track and a trajectory are similar enough, we consider them a pair and calculate the PMs using the particular track and its underlying observations. The probability of detection P_{d} is a radar PM that nicely combines several factors, such as the radar transmitting power and sensitivity or target's radar cross section, behind the ability of the radar to detect the target. It can be calculated through

$$P_{\mathrm{d}} = N/(N+K) \tag{1}$$

where N is the number of detections and K is the number of missed detections during the analyzed part of the trajectory. We determine N and K by comparing the observations data to the trajectory. We compute the probability of track following a corresponding principle:

$$P_{\mathrm{track}} = M/(M+L) \tag{2}$$

where M is the number of tracked samples of the trajectory and L is the number of samples where the track is missing within the analyzed part of the trajectory; uniform sampling is assumed for the navigation data (for the trajectory). The accuracies are computed as a straightforward subtraction or an Euclidean distance between the locations of the trajectory samples and the locations of the track samples or the observations.

We divide the airspace into voxels that each represents a volumetric cell in a Cartesian grid. Accordingly, we divide and assign the trajectories into the voxels to assess the PMs in the spatial domain, as a function of the East and North coordinates and the altitude. We define all the PMs as average values over the samples assigned to each voxel. In order to obtain a statistically significant result, there should be plenty

of trajectories going through each voxel. The number of the trajectory samples should be at least tens but preferable hundreds in one voxel. Thereby to calculate one PM value in one voxel, several information sources are fused and a large number of individual PM values is calculated (over each trajectory) and averaged. Some typical PMs are listed in Fig 3. The main difference in the PM assessment, compared to [3], is the novel use of the flight state estimation and the RCS information, which we discuss next.

B. Estimation of the Flight States

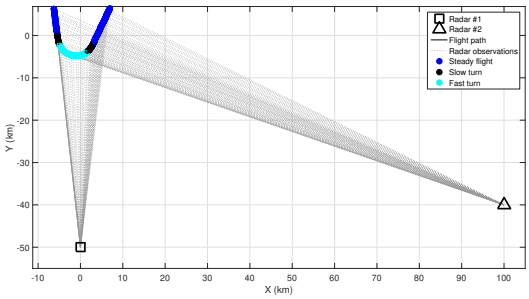
The machine learning framework we presented in [3] lacked the processing of the aircraft kinematics in the prediction of the tracking performance. However, the behavior of the aircraft during flight—i.e. its velocity, angular velocities, accelerations, and maneuvering—strongly affects the tracker performance, especially its accuracy. To overcome this issue, we introduce the analysis of the aircraft kinematics based on its trajectory. Unless the navigation data include the aircraft orientation (the yaw, pitch, and roll angles), we employ elementary flight dynamics [13] to estimate the orientation to obtain the trajectories with six degrees of freedom and further the flight states during the trajectories. We define the flight states by computing several quantities: velocities, accelerations, angles, and angular velocities. In this paper, we consider the following three flight states.

1) Steady flight: flying straight with almost constant yaw, tightly limited roll angle, and limited accelerations.
2) Slow turn: low yaw angle rate with limited roll angle and limited roll rate.
3) Fast turn: high yaw angle rate with moderately limited roll angle and limited roll angle rate.

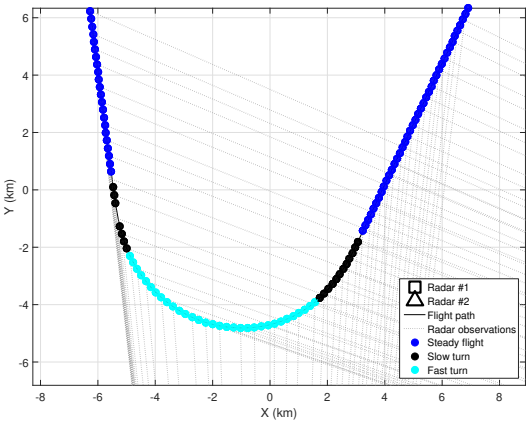
We divide the flight trajectories into segments based on the flight state (see Fig. 4) and assess the air surveillance PMs for each flight state separately. The assessment includes the PMs for both the radar and the sensor fusion, and based on it, we produce the tracker model using the machine learning approach. This procedure is described in more detail in Section III and in Section IV.

C. Estimation of the Detection Threshold

Next we describe how we extend the radar performance assessment of [3] to include the detection threshold by a parallel analysis of the P_{d} and the RCS individually for each voxel. The inclusion of the RCS calls for the knowledge about the aspect angle of the target which we compute based on the aircraft trajectories with six degrees of freedom and the locations of the radars. In this paper, we consider the RCS as a histogram to allow us to take the uncertainty about the aspect angle of the aircraft into account (modeled as Gaussian) [8]–[11]. By comparing the scene data and the surveillance recordings, we assess where the aircraft have been observed by the air surveillance and where not. We estimate the RCS of the aircraft at each location on the scene during its flight path. This includes the time instants at which a radar has reported an observation as well as the time instants of the missed observations. Based on these data, we can estimate the probability of detection by (1) for the aircraft in question and the minimum detectable RCS. This enhanced assessment of the radar PMs allows more accurate verification or learning of the radar model [3] parametrization.



(a) The simulation setup: two radars (#1 and #2) of the same type and a short portion of the trajectory.



(b) An enlargement of the trajectory.

Fig. 4. The simulation used to illustrate the processing of RCS and probability of detection. The simulated aircraft is a Cessna 172. This setup is considered in Figs. 5–9.

Next the proposed assessment of the detection threshold is described in more detail via a simulated example. The trajectory of a Cessna 172 aircraft and two radars (#1 and #2) observing it are shown in Fig. 4. The two radars are of the same type. Please note that the proposed procedure does not pose restrictions on the number of radars or the synchronization between their observations. The estimated orientation of the Cessna is shown in Fig. 5. The aspect angle for the both radars, computed based on the orientation, is visualized over the simulated RCS of the Cessna in Fig. 6. The corresponding RCS histograms during this flight path for the both radars are shown in Fig. 7. It is important to notice that although this is a simulated example, we are able to estimate the aspect angle for measured observations over a measured track as well. Since we have associated the tracks with the navigation recordings, we can calculate the trajectories with

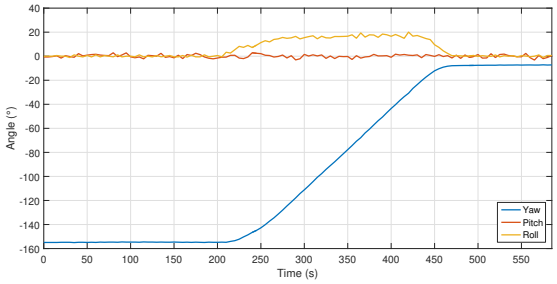


Fig. 5. The aircraft orientation in terms of the yaw, pitch, and roll angles in the simulated example. The orientation has been estimated by an inverse flight simulation based on the trajectory shown in Fig. 4.

six degrees of freedom based on either of them. Typically, the navigation data provides more accurate estimation than the data from the surveillance system. Obviously we need to know the target type (apparent for the cooperative targets), and the reference RCS of the target in question needs to be precomputed and stored in the library. In Fig. 7 we have simulated the RCS of each radar observation by producing a random sample from the corresponding histogram. By the radar #2, the Cessna is seen from the left-hand side at the beginning of the simulation; the side view makes the expected RCS of high value during the first 200 seconds. The RCS histograms and the simulated observations—both the detected and the undetected samples—are shown in Fig. 7. The RCS realizations are shown only to illustrate the simulation and they are not used in the proposed analysis. Obviously in the case of real data, the observation-specific realized RCS values are available only if the radar system produces RCS estimates—and even in that case, the radar system is not able to produce the RCS estimates for the undetected samples.

We average the RCS histograms over the flight path to obtain a histogram representing each radar on average. The results are shown in the form of a cumulative density function (CDF) of the RCS for the radar #1 in Fig. 8 and for the radar #2 in Fig. 9. The averaging is performed to obtain the RCS likelihood (histogram) representing a part of the flight trajectory as in Fig. 4. We can use the average CDF of the RCS to assess PMs representing a portion of the trajectory in general. This property is important in the large-scale assessment of the PMs for which the trajectories are split and assigned to the voxels. The result is the voxel-specific PMs. The CDFs has been calculated completely based on the reference RCS in the target library.

We define the detection threshold in the form of the minimum detectable RCS. It is not necessary for the radar system to provide the measured RCS along with its observations. In any case, the measured RCS of the target observations for the undetected samples cannot be produced with any system. Depending on whether the radar system provides the RCS estimates or not, we propose the following two strategies to estimate the detection threshold, i.e. the minimum detectable RCS, from the recordings.

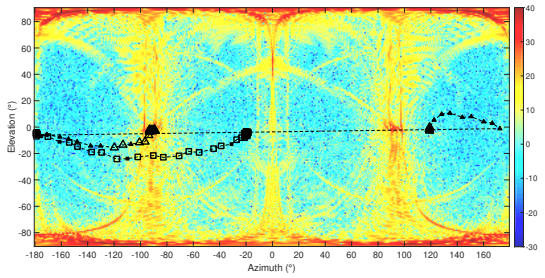


Fig. 6. The simulated RCS of the Cessna 172 as a function of the azimuth angle and the elevation angle. The elevation angle of 0° corresponds to the horizontal plane of the target. In the horizontal axis, the target azimuth angle of 0° corresponds to the target from the front. The color represents RCS as dBsm. The observations of two radars are shown over the RCS: the large unfilled markers correspond to the radar detections and the small filled markers to the undetected samples. The marker symbols correspond to the radars shown in Fig. 4.

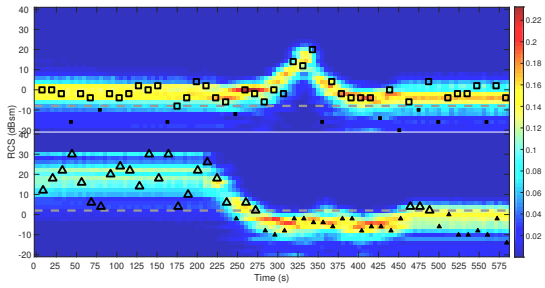


Fig. 7. The simulated RCS likelihood histogram as a function of time. The observations of two radars are shown over the RCS: the large unfilled markers correspond to the radar detections and the small filled markers to the undetected samples. The marker symbols correspond to the radars shown in Fig. 4. The top image represents the radar #1 and the bottom image the radar #2. The detection thresholds for both radars are visualized with gray dashed lines; the range of the radar #2 is about twice the range of the radar #1, and so the threshold is about 10 dB higher for the radar #2.

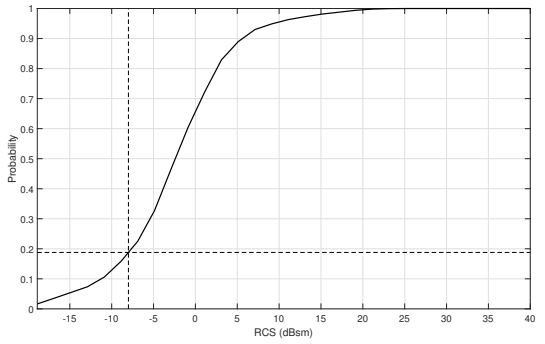


Fig. 8. The average CDF (of the RCS likelihood) calculated over the simulated flight trajectory for the radar #1. The detection threshold at -8 dBsm is marked with the dashed lines. The probability of detection is 0.81.

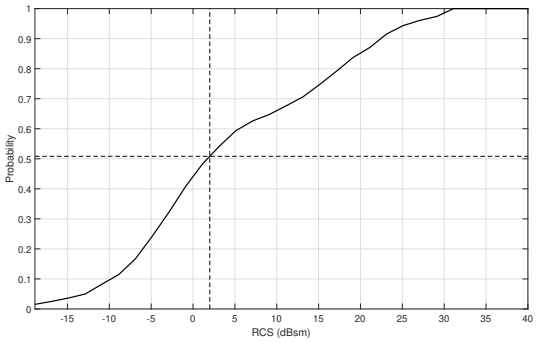


Fig. 9. The average CDF (of the RCS likelihood) calculated over the simulated flight trajectory for the radar #2. The detection threshold at 2 dBsm is marked with the dashed lines. The probability of detection is 0.49.

1) If the radar provides the RCS with its observations, the measured RCS values can be used to estimate the minimum detectable RCS. This case is illustrated in Fig. 7. The estimation of the RCS threshold indicated by the gray dashed lines is quite straightforward using the lowest RCS values of the detected samples. The RCS threshold can be determined e.g. by the minimum or by a small percentile value of the set of the RCS values.
2) If the radar does not provide the RCS, the P_d reveals the minimum detectable RCS. P_d can be calculated by (1). Then the simulated average CDF tells the minimum detectable RCS at the $P = 1-P_\mathrm{d}$. In Fig. 7, for the radar #1 $P_{\mathrm{d}1} = N_1/(N_1 + K_1) = 39/(39 + 10) \approx 0.80$, and for the radar #2 $P_{\mathrm{d}2} = N_2/(N_2+K_2) = 25/(25+24) \approx 0.51$ which are close the correct P_d values presented in Figs. 8–9, and so they would result in close to the correct minimum detectable RCS.

The minimum detectable RCS is a useful radar PM for the performance prediction. By employing the basic signal-to-noise ratio calculations [1], it allows the prediction of P_d for a target of any RCS. In addition, the P_d for complex targets can be evaluated using the obtained RCS threshold and the concepts presented in [10], [11]. Expressed in the simplest terms, the P_d is the probability of observing RCS values higher than the minimum detectable RCS.

III. Performance Modeling by Machine Learning

The learning of the tracker model was performed with the genetic programming based rule learner (GPRL) using the principles we previously proposed [6]. The fundamental idea is predicting the tracking accuracy from the radar PMs. Thus the radar PMs at each voxel are the input of the model, and the tracking accuracy at each voxel is the output (the target) of the model.

The main difference between the implementation of the GPRL in [6] and in this paper is the enhanced calculation of the radar PMs and the tracker PMs using the division into flight states. In addition, the definition of tracker performance levels in this paper is different from [6]. Instead of five tracker performance levels, we use only four of them: `excellent`, `good`, `moderate`, and `weak`. We determine the tracker performance levels using the Euclidean position error of the tracks in meters T_n^{EEuclid}, where n refers to the index of the voxel v_n, into which the airspace is divided. There is a threshold r for T_n^{EEuclid} as follows:

$$r_{\mathrm{excellent}}^{\mathrm{EEuclid}} < r_{\mathrm{good}}^{\mathrm{EEuclid}} < r_{\mathrm{moderate}}^{\mathrm{EEuclid}} \tag{3}$$

Voxels v_n belong to the defined four categories $C_{\mathrm{excellent}}$, C_{good}, C_{moderate}, and C_{weak} as follows

$$\begin{cases} v_n \in C_{\mathrm{excellent}} & \text{if } T_n^{\mathrm{EEuclid}} < r_{\mathrm{excellent}}^{\mathrm{EEuclid}} \\ v_n \in C_{\mathrm{good}} & \text{if } (T_n^{\mathrm{EEuclid}} < r_{\mathrm{good}}^{\mathrm{EEuclid}}) \wedge (v_n \notin C_{\mathrm{excellent}}) \\ v_n \in C_{\mathrm{moderate}} & \text{if } (T_n^{\mathrm{EEuclid}} < r_{\mathrm{moderate}}^{\mathrm{EEuclid}}) \\ & \wedge (v_n \notin (C_{\mathrm{excellent}} \cup C_{\mathrm{good}})) \\ v_n \in C_{\mathrm{weak}} & \text{else.} \end{cases} \tag{4}$$

I.e. the tracker performance level of a voxel is determined by its position error: the voxel is assigned to the highest performance level that has a threshold larger than the position error of the voxel in question.

The parameter values of the GPRL were not optimized for the learning of the tracker performance model. The same parameter values were used as in [6] except for the parameter `maxGenerations`, which was increased from 20 to 100 in order to ensure appropriate learning of the tracker performance models with the new PMs.

IV. Experiments with Real Data

In this section, we present results of the comparison between two cases of the performance modeling with the GPRL. In the first case, we learn a model of the tracker performance by using PMs that are calculated without separating the different flight states from each other. In the second case, we learn a separate model of tracker performance for three flight states: steady flight, slow turn, and fast turn. We use real air surveillance data in the experiments. The data divided into two occasions contain the seven radar PMs and the tracker accuracy described in [6] and discussed in Section II in this paper. In addition, the number of recorded flight trajectories in each voxel can be considered sufficient. A typical number of samples in a voxel was several hundreds. The number of voxels was around two hundreds depending on the flight state in question. Consequently, the data represents wide variety of flying, even when considered specifically for each flight state. Each performance measure was produced as a voxel-specific average using only portions of the trajectories containing the flight state in question.

The value of the Euclidean position error of the tracks was used to define the four target categories: `excellent`, `good`, `moderate`, and `weak`. The voxels represent constant-size volumes in a Cartesian grid aligned with the world coordinate system. The voxels in the western half of the studied air space of the first occasion and the eastern half of the studied air space of the second occasion served as the training data. The remaining voxels were used as the test data to evaluate the generalization capability of the classifiers.

TABLE I
THE MEAN AND STANDARD DEVIATION OF THE CLASSIFICATION ACCURACY OVER 50 RUNS. TWO CASES ARE CONSIDERED: DIFFERENT FLIGHT STATES TOGETHER AND DIFFERENT FLIGHT STATES SEPARATED (STEADY FLIGHT, SLOW TURN, AND FAST TURN).

	Mean	Standard deviation
All flight states	0.4435	0.0087
Steady flight	0.5214	0.0117
Slow turn	0.5255	0.0288
Fast turn	0.5175	0.0448

TABLE II
THE MEAN AND STANDARD DEVIATION OF THE UNWEIGHTED COHEN'S KAPPA OVER 50 RUNS.

	Mean	Standard deviation
All flight states	0.1838	0.0138
Steady flight	0.2305	0.0199
Slow turn	0.2318	0.0443
Fast turn	0.2098	0.0637

TABLE III
THE MEAN AND STANDARD DEVIATION OF THE LINEAR WEIGHTED COHEN'S KAPPA OVER 50 RUNS.

	Mean	Standard deviation
All flight states	0.2970	0.0186
Steady flight	0.3638	0.0199
Slow turn	0.3559	0.0348
Fast turn	0.3121	0.0707

The classification models were learned from the training data and tested using the test data. The Tables I–IV show four classification performance measures for the classifiers. They include the classification accuracy, the unweighted Cohen's kappa, the linear weighted Cohen's kappa, and the quadratic weighted Cohen's kappa [14]. The weighted Cohen's kappa takes into account the degree of disagreement between the classification result and the ground truth. The degree of disagreement refers to the fact that classifying, for example, a voxel with `good` tracker performance into class `excellent` or `moderate` is more correct than classifying it into `weak`.

The GPRL exploits randomness, and thus the learning was performed 50 times for the first case and for each of the three flight states in the second case. The classification performance values in the Tables I–IV represent the averages over the performed 50 runs. This means the same test and train data (where each voxel includes the radar PMs and the tracker performance level) was used again and again; so the randomness is a property of the training of the rule learner classifier (GPRL). In addition, the standard deviations of the classification performance values are provided. The best result in terms of all the classification performance values are achieved using the classifiers created separately for the different flight states.

TABLE IV
THE MEAN AND STANDARD DEVIATION OF THE QUADRATIC WEIGHTED COHEN'S KAPPA OVER 50 RUNS.

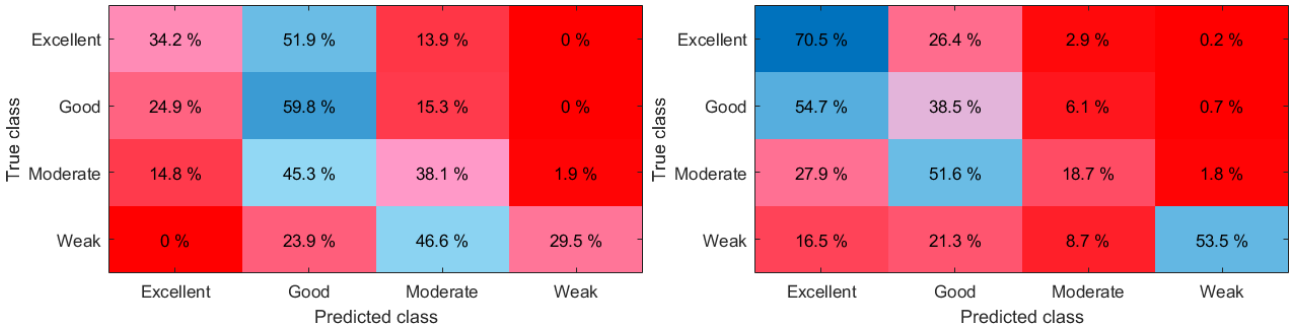
	Mean	Standard deviation
All flight states	0.4319	0.0236
Steady flight	0.5040	0.0205
Slow turn	0.4983	0.0292
Fast turn	0.4439	0.0799

Fig. 10 shows the confusion matrices for the first case and for the three flight states of the second case. The presented confusion matrices have been calculated by summing up the resulting 50 confusion matrices for each flight state and normalizing them class-wise.

V. DISCUSSION

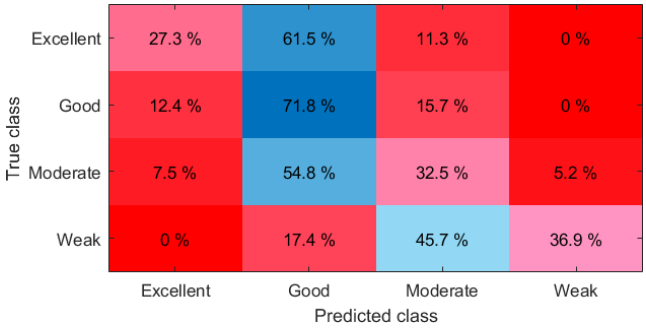
We proposed two modifications to our framework of the air surveillance performance assessment and prediction [3]. The first modification was the inclusion of the aircraft kinematics by dividing the data and the analysis into segments according to the flight state. The division of the radar target behavior into three flight states, 1) steady flight, 2) slow turn, and 3) fast turn, was shown to increase the accuracy of the tracking performance prediction. Although the prediction results improved by using the flight state division, the absolute true positive rate values in the confusion matrices are not on a satisfactory level. Thus, more work should be done to better understand the weak points in the machine learning procedure. Topics of further development include e.g. the computation of the radar and tracker PMs as well as the definition of the flight states based on the target kinematics.

The second modification considered incorporating a detailed characterization of the target RCS into the performance assessment. The procedure requires aircraft flight trajectories and RCS information of good quality. We estimate the minimum detectable RCS at discrete locations in the studied air space. Based on this result, we are able to learn and verify the radar performance in a manner that is normalized concerning the target RCS. A radar performance model that is learned using the proposed method allows predicting the probability of detection for any kind of aircraft with a known RCS distribution. The procedure uncovers the detection threshold in relation to the RCS. At the same time, this concept provides the relation between the signal-to-noise ratio and the RCS. The radar accuracy is a function of the signal-to-noise ratio. Thus the concept can be developed further to link the RCS and the radar accuracy in a similar manner to predict the radar accuracy for any target. In addition, the proposed estimation of the minimum detectable RCS may be applied to certain real-time schemes such as the radar calibration or the noncooperative target recognition in which the information concerning the detectability of the appearing targets is valuable.

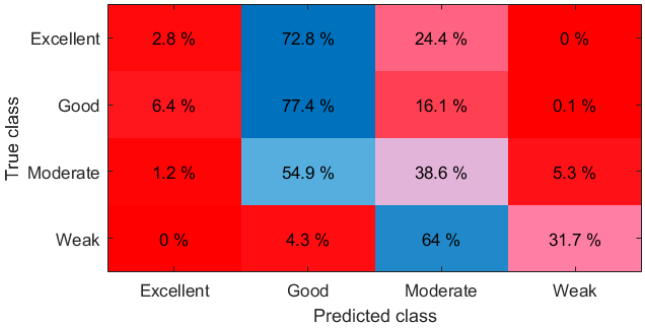


(a) Case 1: All the flight states.

(b) Case 2: Flying according to the definitions of the steady flight.



(c) Case 2: Flying according to the definitions of the slow turn.



(d) Case 2: Flying according to the definitions of the fast turn.

Fig. 10. The mean confusion matrices of the prediction of the tracking performance level for each test voxel over 50 runs. The matrices have been normalized class-wise. High probabilities in the diagonal (the true positive rates) and close to the diagonal imply good accuracy in the tracker performance prediction. Two cases are considered: different flight states together and different flight states separated (steady flight, slow turn, and fast turn).

VI. Conclusion

In this paper, we proposed a concept for the improved automated assessment of the performance of an air surveillance system exploiting the concepts related to big data and artificial intelligence. The enhancement was achieved using the navigation data of the cooperative aircraft and the information about their RCS. The presented experiments showed the achieved benefits in the accuracy of the radar performance analysis and modeling of the sensor fusion. Based on these findings, we are able to compute and predict the performance of the analyzed air surveillance system considering any cooperative or noncooperative target and thus to provide useful information for the decision making concerning the use of the system.

References

[1] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Volume I - Basic Principles*. SciTech Pub., 2010.

[2] B. O'Hern, J. Mack, A. Pieramico, D. Rugger, and L. Moyer, "The radar support system (RSS): a tool for siting radars and predicting their performance," *IEEE Aerospace and Electronic Systems Magazine*, vol. 12, pp. 19–26, December 1997.

[3] J. Jylhä, M. Ruotsalainen, V. Väisänen, K. Virtanen, M. Harju, and M. Väilä, "A machine learning framework for performance prediction of an air surveillance system," in *2017 European Radar Conference (EURAD)*. IEEE, 2017, pp. 187–190.

[4] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. YBS publishing, 2011.

[5] J. Besada, A. Soto, G. de Miguel, J. García, and E. Voet, "ATC trajectory reconstruction for automated evaluation of sensor and tracker performance," *IEEE Aerospace and Electronic Systems Magazine*, vol. 28, pp. 4–17, February 2013.

[6] M. Ruotsalainen and J. Jylhä, "Learning of a tracker model from multi-radar data for performance prediction of air surveillance system," in *2017 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2017, pp. 2128–2136.

[7] H. Perälä, M. Väilä, and J. Jylhä, "M-SPURT—Compressing the target characterization for a fast monostatic RCS simulation," in *2018 International Conference on Radar (RADAR)*. IEEE, 2018, pp. 1–6.

[8] M. Väilä, J. Jylhä, H. Perälä, and A. Visa, "Performance evaluation of radar NCTR using the target aspect and signature," in *2014 IEEE Radar Conference*. IEEE, 2014, pp. 623–628.

[9] M. Väilä, J. Jylhä, M. Ruotsalainen, and H. Perälä, "Estimating the distribution of RCS using sparse measured samples," in *2018 International Conference on Radar (RADAR)*. IEEE, 2018, pp. 1–6.

[10] M. Väilä, J. Jylhä, T. Sailaranta, H. Perälä, V. Väisänen, and A. Visa, "Incorporating a stochastic model of the target orientation into a momentary RCS distribution," in *2015 IEEE Radar Conference (RadarCon)*. IEEE, 2015, pp. 969–973.

[11] M. Väilä, J. Jylhä, V. Väisänen, H. Perälä, A. Visa, M. Harju, and K. Virtanen, "A RCS model of complex targets for radar performance prediction," in *2017 IEEE Radar Conference (RadarConf)*. IEEE, 2017, pp. 430–435.

[12] J. P. Blackburn, T. D. Ross, A. R. Nolan, J. C. Mossing, J. U. Sherwood, D. J. Pikas, and E. G. Zelnio, "Derived operating conditions for classifier performance understanding," in *Proc. of SPIE Vol. 8051, Algorithms for Synthetic Aperture Radar Imagery XVIII*. SPIE, 2011.

[13] B. Etkin, *Dynamics of Flight*. John Wiley and Sons, 1982.

[14] J. Cohen, "Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit," *Psychological Bulletin*, vol. 70, no. 4, pp. 213–220, October 1968.