NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

**PERFORMANCE METRICS FOR
CORRELATION AND TRACKING ALGORITHMS**

by

Nathan S. Dietrich

June 2001

Thesis Advisor: Robert A. Koyak
Second Reader: Thomas W. Lucas

Approved for public release; distribution is unlimited.

Form SF298 Citation Data

Report Date ("DD MON YYYY") 15 Mar 2001	Report Type N/A	Dates Covered (from... to) ("DD MON YYYY")

Title and Subtitle PERFORMANCE METRICS FOR CORRELATION AND TRACKING ALGORITHMS	**Contract or Grant Number**
	Program Element Number
Authors	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Naval Postgraduate School Monterey, CA 93943-5138	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 111	

PERFORMANCE METRICS FOR CORRELATION AND TRACKING ALGORITHMS

Nathan S. Dietrich-Captain, United States Army
B.S., University of South Dakota, 1987
Master of Science in Operations Research-June 2001
Advisor: Robert A. Koyak, Department of Operations Research
Second Reader: Thomas W. Lucas, Department of Operations Research

Military commanders require situational awareness to support real-time decision-making. To obtain information on possibly hostile entities in an area of interest, surveillance systems, which receive information from sensors such as radars, intelligence, and other sources, are often used. One of the objectives of surveillance systems that track aircraft is the formation of a Single Integrated Air Picture (SIAP), that represents a coherent resolution of information. Correlation is the process by which sensor measurements and other information are combined to keep the SIAP up-to-date in real time. A correlator, which is the software implementation of a correlation methodology, must resolve ambiguities and conflicting information to provide an operationally useful synthesis of surveillance data. Possible ambiguities include missed tracks, extra tracks, or position and velocity errors. The metrics developed in this thesis are designed for use in evaluating the performance of air surveillance systems, of which correlators are an integral part. Maneuvering or closely spaced aircraft pose difficult issues for air surveillance systems. These are addressed by the performance metrics. Using scripted test scenarios in a modeling and simulation environment, comparisons of correlators can be made using nonparametric statistical methods. An experiment constructed in this manner can be used to support acquisition decision-making.

DoD KEY TECHNOLOGY AREA: Command, Control and Communications**,** Modeling and Simulation, Sensors

KEYWORDS: Correlator, Correlator Performance Metrics, Surveillance Systems, Single Integrated Air Picture (SIAP)

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2001	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Performance Metrics for Correlation and Tracking Algorithms			**5. FUNDING NUMBERS**
6. AUTHOR(S) Nathan S. Dietrich			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Space & Missile Defense Command SMDC-BL-ME 106 Wynn Dr. Huntsville, AL. 35807			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			**12b. DISTRIBUTION CODE**

13. ABSTRACT *(maximum 200 words)*

Military commanders require situational awareness to support real-time decision-making. To obtain information on possibly hostile entities in an area of interest, surveillance systems, which receive information from sensors such as radars, intelligence, and other sources, are often used. One of the objectives of surveillance systems that track aircraft is the formation of a Single Integrated Air Picture (SIAP), that represents a coherent resolution of information. Correlation is the process by which sensor measurements and other information are combined to keep the SIAP up-to-date in real time. A correlator, which is the software implementation of a correlation methodology, must resolve ambiguities and conflicting information to provide an operationally useful synthesis of surveillance data. Possible ambiguities include missed tracks, extra tracks, or position and velocity errors. The metrics developed in this thesis are designed for use in evaluating the performance of air surveillance systems, of which correlators are an integral part. Maneuvering or closely spaced aircraft pose difficult issues for air surveillance systems. These are addressed by the performance metrics. Using scripted test scenarios in a modeling and simulation environment, comparisons of correlators can be made using nonparametric statistical methods. An experiment constructed in this manner can be used to support acquisition decision-making.

14. SUBJECT TERMS Correlator, Correlator Performance Metrics, Surveillance Systems, Single Integrated Air Picture (SIAP)			**15. NUMBER OF PAGES**
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

PERFORMANCE METRICS FOR CORRELATION AND TRACKING ALGORITHMS

Nathan S. Dietrich
Captain, United States Army
B.S., University of South Dakota, 1987

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2001

Author: ____________________________________
Nathan S. Dietrich

Approved by: ____________________________________
Robert A. Koyak, Thesis Advisor

____
Thomas W. Lucas, Second Reader

____
James N. Eagle, Chairman
Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Military commanders require situational awareness to support real-time decision-making. To obtain information on possibly hostile entities in an area of interest, surveillance systems, which receive information from sensors such as radars, intelligence, and other sources, are often used. One of the objectives of surveillance systems that track aircraft is the formation of a Single Integrated Air Picture (SIAP), that represents a coherent resolution of information. Correlation is the process by which sensor measurements and other information are combined to keep the SIAP up-to-date in real time. A correlator, which is the software implementation of a correlation methodology, must resolve ambiguities and conflicting information to provide an operationally useful synthesis of surveillance data. Possible ambiguities include missed tracks, extra tracks, or position and velocity errors. The metrics developed in this thesis are designed for use in evaluating the performance of air surveillance systems, of which correlators are an integral part. Maneuvering or closely spaced aircraft pose difficult issues for air surveillance systems. These are addressed by the performance metrics. Using scripted test scenarios in a modeling and simulation environment, comparisons of correlators can be made using nonparametric statistical methods. An experiment constructed in this manner can be used to support acquisition decision-making.

THIS PAGE INTENTIONALLY LEFT BLANK

DISCLAIMER

The research described in this thesis was conducted using the Joint Composite Tracking Network (JCTN) Benchmark Environment, Release 1.08.02. Capt. Dietrich wishes to thank the Office of Naval Research for making this software available for his research. The JCTN Benchmark Environment was utilized as a platform for the development of metrics that do not reflect on the performance, or quality, of either the JCTN Benchmark Environment or the JCTN Pilot Benchmark Process. This thesis does not constitute an assessment of the JCTN Benchmark Environment or the JCTN Pilot Benchmark Process.

At the time that this thesis was completed, Release 1.08.02 had been replaced by an updated, and significantly enhanced, version of the JCTN Benchmark Environment. Capt. Dietrich had, by that time, completed a substantial portion of his thesis research on Release 1.08.02, for which reason he continued to use Release 1.08.02 until the completion of his work.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND/OR ABBREVIATIONS

AWACS	Airborne Warning and Control System
AWarE	Advanced Warfare Environment
BL	Battle Lab
C2	Command and Control
DoD	Department of Defense
FP	fire platoon
EADSIM	Extended Air Defense Simulation
FOC	Future Operational Capability
JCTN	Joint Composite Tracking Network
JDN	Joint Data Network
M&S	Modeling and Simulation
MATLAB	MATrix LABoratory (a software package for numerical computation and visualization)
MSCT	Multi Source Correlator Tracker
PATRIOT	Phase Array Tracking to Intercept of Target
PBE	Pilot Benchmark Environment
PDUs	Protocol Data Units
RMSE	Root Mean Squared Error
RSSAE	Root Sum of Squared Average Error
SIAP	Single Integrated Air Picture
SMDC	United States Army Space & Missile Defense Command

SNR	signal-to-noise ratio
SOLIPSYS	Solutions for Information Processing Systems
TADIL	Tactical Digital Information Link
TSIU	Tactical Simulation Interface Unit

ACKNOWLEDGMENTS

Capt. Dietrich would like to thank Prof. Robert Koyak and Prof. Thomas Lucas for their invaluable direction, guidance, and input into this thesis. Capt. Dietrich would also like to thank Mr. Loren Roe, of the Advanced Research Center of Huntsville, Alabama, and Keith Cooper, of Teledyne Brown Engineering for their many hours of work in helping to collect data.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Military commanders require situational awareness of their areas of responsibility to support real-time decision-making. Having reliable information on what is happening in their areas of operation can make the difference between successful and catastrophic outcomes. Substantial investment has been made in the development of surveillance systems to give United States military commanders accurate and timely situational awareness of potentially hostile vehicles. Surveillance provides real-time information to the commander on the "state" of the physical space in an area of interest. Military commanders need the surveillance systems used by their commands to be both accurate and timely.

To ensure that surveillance systems are suitable to their purpose, the need exists for a methodology for evaluating their performance. The purpose of this thesis is to develop a methodology for evaluating the accuracy of air surveillance systems.

A surveillance system collects, coordinates, processes, analyzes and presents information to military commanders. Surveillance systems use sensors such as radars to obtain information on possibly hostile vehicles in an area of interest. Systems that conduct surveillance do so in real time over extended periods. Each sensor updates its measurements at short, periodic intervals. Air surveillance is concerned primarily with tracking aircraft over a particular theater of interest. A primary objective of air surveillance based on multiple sensors and information sources is the formation of a Single Integrated Air Picture (SIAP). The process of updating the track registry of the air surveillance system is subject to errors and ambiguities. There can be conflicting information from different sources. Correlation is the process by which sensor measurements and other data are used to update the track registry. In modern tracking, correlation is performed using algorithms that recognize the random nature of sensor measurement errors, and the uncertainties inherent in associating information to a set of recognized objects.

Correlation must resolve significant ambiguities to provide an operationally useful synthesis of surveillance data. Performance evaluation of multi-target, multi-sensor tracking that centers on the use of a particular correlator must account for these potential errors. But, what is desired of a correlator is clear: it should promote the accurate description of the surveillance space across time.

A correlator is one of several important components of a multi-target, multi-sensor air surveillance system. An evaluation of its performance should be based on the end result of using the correlator; in other words, on the accuracy of tracking. However, errors in tracking are not necessarily attributable to the correlator. Tracking errors can arise due to bias in the sensors, to the random measurement error that is always present in tracking, and to the uncertainty in making associations between sensor measurements and tracks that is also present.

Nonetheless, accuracy of tracking can be used as a criterion for evaluating the relative performance of one correlator to another, provided that testing is conducted with correlators used in identical scenarios. That way, differences in performance can be attributed to the correlators, and not to another component of the surveillance system.

The purpose of this thesis is to develop and assess performance metrics that can be used for the evaluation and comparison of correlators in the context of air surveillance. Metrics for assessing the accuracy of tracking of a surveillance system in the dynamic sense can be developed relative to a period of time in which the correlator is exercised. The metrics can provide a basis for determining relative performance of the correlators, and they can isolate performance issues under difficult conditions posed by maneuvering aircraft or closely spaced aircraft.

The performance metrics described in this thesis were designed for the evaluation of correlators in the context of air surveillance. The Maneuver Metric and the Closely Spaced Objects Metric developed in the body of the thesis can be used to evaluate tracking performance when faced with maneuvering aircraft or closely spaced aircraft, respectively.

Using modeling and simulation to design test scenarios, comparisons of correlators can be made with nonparametric statistical methods. These comparisons can be made whether the data for the correlators are dependent or independent.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION AND BACKGROUND

Military commanders require situational awareness of their areas of responsibility to support real-time decision-making. Having reliable information on what is happening in their areas of operation can make the difference between successful and catastrophic outcomes. For example, accurate and timely situational awareness might have prevented the submarine USS Greeneville (SSN 772) from colliding with the Japanese fishing trawler Ehime Maru off the waters south of Honolulu, Hawaii on 9 February 2001 (Gunder, 2001).

Substantial investment has been made in the development of surveillance systems to give United States military commanders accurate and timely situational awareness of potentially hostile vehicles. Such vehicles include enemy aircraft, tactical missiles, theater ballistic missiles, surface ships, submarines, and land-based vehicles. A common feature of these vehicles is that they can change their locations with time, and the number of such threats can also change with time.

Surveillance provides real-time information to the commander on the "state" of the physical space in an area of interest. Even with recent developments in sensors and information processing used by the military, it remains a challenge to obtain a surveillance picture that correctly identifies threats and their locations in real time. Errors in sensor information lead to errors in the overall awareness of potential threats. Military commanders need the surveillance systems used by their commands to be both accurate and timely.

To ensure that surveillance systems are suitable to their purpose, the need exists for a methodology for evaluating their performance. The purpose of this thesis is to develop a methodology for evaluating the accuracy of air surveillance systems.

A. SURVEILLANCE SYSTEMS

A surveillance system collects, coordinates, processes, analyzes and presents information to military commanders. Surveillance systems can collect and utilize different forms of information. A familiar characteristic of U. S. military surveillance systems is their use of sensors, such as space-based infrared sensors, air or ground-based radars and sonar. However, many surveillance systems are also capable of utilizing information from intelligence reports and voice radio transmissions.

A sensor is “a device that observes the (remote) environment by reception of some signals (energy)” (Bar-Shalom, 1995, p. 7). Surveillance systems collect sensor measurements on detected objects in the area of interest. When a threat has been detected by a surveillance system, it is recognized as a “contact.” In the case of radars, sensor measurements are signals that are received (or returned) whose amplitudes exceed a signal-to-noise-ratio (SNR) threshold. Sensor measurements are used by the surveillance system to estimate positions and velocities of its contacts at a fixed point in time within the area of interest. The estimation of contact positions and velocities by processing sensor measurements is referred to as tracking (Bar-Shalom, 1995, p. 5).

Systems that conduct surveillance do so in real time over extended periods. Each sensor updates its measurements at short, periodic intervals. An estimate of the state of the surveillance space at a moment in time consists of a registry of tracks corresponding to objects that have been detected. A track is a state trajectory estimated from the set of sensor measurements (Bar-Shalom, 1995, p. 6), consisting of the position, velocity, and other attributes of a putative object across time. The word "putative" in this context means that the surveillance system does not know with absolute certainty that the object exists, but it perceives it as such. Tracks are based on information that comes from a single sensor, from multiple sensors of similar type that may be networked, or from a mixture of sources. In whatever form the information is received, the objective of real-time tracking is to merge new information with the current track registry to produce an updated track registry. At a given time, the track registry consists of up-to-date information on all objects that the system believes exist. Existing tracks are provided with updated attribute estimates, or are dropped from the registry if they can no longer be associated with an object. New tracks are entered into the registry to represent previously undetected objects.

Air surveillance is concerned primarily with tracking aircraft over a particular theater of interest. A primary objective of air surveillance based on multiple sensors and information sources is the formation of a Single Integrated Air Picture (SIAP). A SIAP is a common operational view of the air theater of interest in which:

(1) All inputs are integrated to form one air picture;

(2) All conflicts in the air picture from the different sources are deconflicted (Litton, 2000).

In particular, sensors such as radar can provide different estimates of the positions and velocities of aircraft within the surveillance space. These differences must be de-conflicted.

The configuration of Phase Array Tracking to Intercept of Target (PATRIOT) firing platoons (FP) illustrates the concept of a multi-sensor tracking system used for air surveillance. PATRIOT is the U.S. Army's advanced air defense system, capable of defeating both high performance aircraft and tactical ballistic missiles (Redstone Arsenal, 2001). A PATRIOT battalion consists of up to six Patriot Fire Units, each having its own AN/MPQ-53 phased array radar, that searches the airspace for enemy missiles and aircraft. Figure 1 illustrates the configuration of PATRIOT FPs within a battalion. Each firing unit reports its sensor information to the battalion headquarters. The PATRIOT battalion headquarters is capable of not only receiving sensor information from its own firing units, but also from other platforms (e.g. AEGIS, AWACS) linked to PATRIOT in a joint network. One such concept is the Joint Data Network (JDN), which is based on Tactical Digital Information Link (TADIL) J messaging, TADIL A/B messaging, and radio messaging.

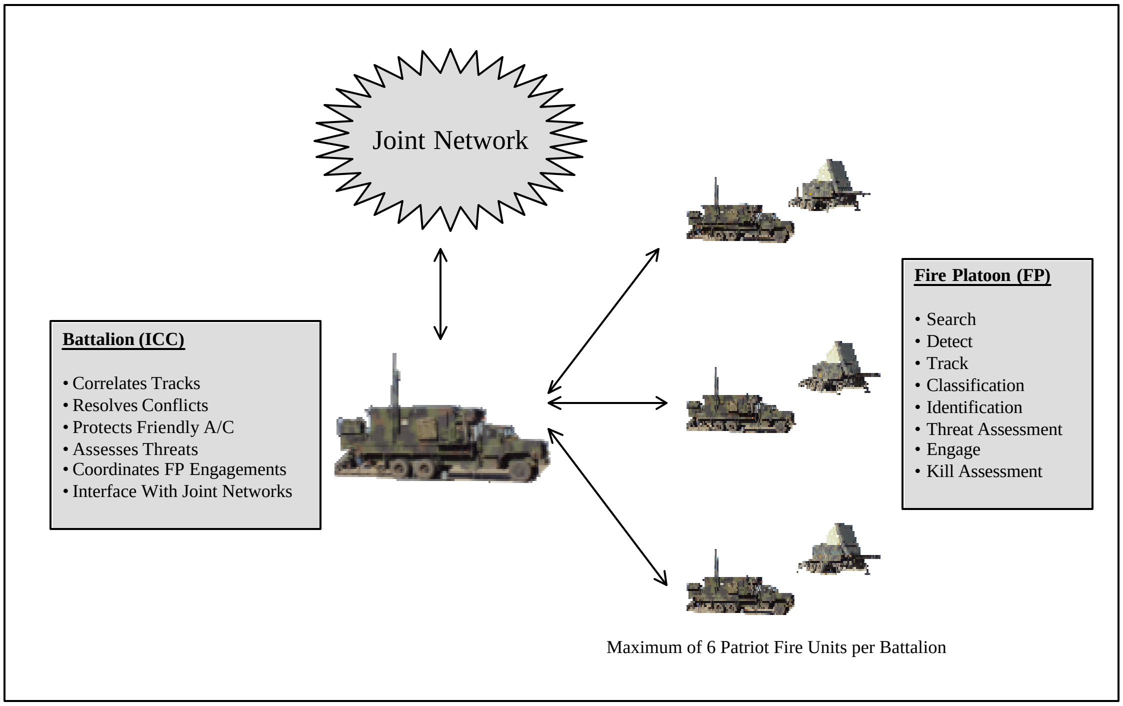

Figure 1. Concept of a battalion of PATRIOT firing platoons. ICC is an abbreviation for Information and Coordination Central, and A/C is an abbreviation for aircraft (From: U.S. Army Air and Missile Defense Program Executive Office, 2000, slide 2).

The battalion updates and maintains its track registry using sensor information that it receives from its FPs and from the Joint Network. From this information, the battalion creates a SIAP that is used by the battalion, its FPs, and by other users of the Joint Network.

The track registry must be changed at regular time increments for two reasons:

(1) Objects change their positions and velocities with time;

(2) New sensor measurements are obtained, which provide additional information about the status of objects in the surveillance space.

This can result in new objects being detected that had not been recognized before, or what had been recognized as objects no longer being regarded as existing objects. Possible updates to the track registry are summarized as follows:

(1) Old track + time update (using physical models) + information update (using sensor measurements) = Updated track;

(2) Sensor measurement with no previous indication of object = New track (subject to track initiation rules);

(3) Old track with no corresponding sensor measurements = Dropped track (subject to track dropping rules).

The process of updating the track registry is subject to errors and ambiguities. There can be conflicting information from different sources. Sensors such as radar are prone to random errors and bias. The presence of clutter, countermeasures, and false alarms increases the likelihood of errors in updating the track registry. This is the case even if a single sensor is used. When multiple sensors and information sources are used, the resolution of conflicts becomes more difficult.

B. CORRELATION

Correlation is integral to the process by which sensor measurements are used to update the track registry. Also known as data fusion, correlation is defined as "the process of taking a new a new input (called a contact), comparing it to a database of previous inputs (called tracks), and deciding whether the new input is updated/revised information about an existing track or is a new, previously unreported input that should be added as a new record in the database." (PMW 171, 1997, p. 1-1) This is done by recognizing uncertainties in models and sensor measurements, and recognizing that associating sensor measurements to objects is subject to error. In addition to real objects, a contact may also refer to a nonexistent object due to radar clutter, glint, multipath, and scintillation that the sensor perceives as a real object.[1]

A correlator is a software product that represents the implementation of a correlation methodology. For example, the Solutions for Information Processing Systems (SOLIPSYS) Multi Source Correlator Tracker (MSCT) "is a generic information synthesis system" and its "primary function is to receive tactical track information from multiple sources and produce a coherent, composite track database for display and dissemination" (SOLIPSYS, 1999, p. 1). A composite track is the integration of the sensor measurements from several different sensors and other sources to form a single estimate of the attributes of an object at a given time. A correlator may be used by a single sensor platform, or by multiple sensors that are engaged in joint tracking.

[1] Radar clutter is unwanted echoes from the ground, sea, rain, chaff, birds, etc. (Barton, 1988, p. 123). Glint is "the inherent random component of error in measurement of position or Doppler frequency of a complex target due to interference of the reflections from different elements of the target" (Barton, 1988, p. 115). Multipath errors are "caused by reflection or forward scatter of the target energy from the surface beneath the target-to-radar path" (Barton, 1988, p. 512). Scintillation errors in a conical-scan radar are caused when the error detector, within the radar, interprets target deviations due to target fluctuations during the scan cycle (Barton, 1988, p. 388).

A simple concept of correlation based on the concept of “gating” can be described as follows:

(1) Each track in the track registry is propagated forward to the “current time.”

(2) An uncertainty region, or gate (usually rectangular or ellipsoidal) is formed around each track for every new measurement that is input into the tracking system.

(3) All current measurements that fall inside the gate of a track are eligible to correlate to it.

(4) The gate is determined from the covariance matrix of the current track plus the covariance matrix of the measurement that is considered as a possible association with the track.

(5) A measurement can fall inside more than one gate; more than one measurement can fall inside the gates of a single track; and, there can be tracks whose gates have no measurements inside.

(6) A correlator uses objective criteria to decide how to resolve ambiguities, initiate new tracks, and delete existing tracks based on which objects fall inside which gates.

In practice, correlator software is complex because it is based on elaborate statistical models. A correlator is tailored to the properties of known sensors (e.g. AN/MPQ-53 phased array radars), to the constraints posed by the tracking paradigm, and it is designed to execute quickly. Correlators are typically developed by commercial entities that regard their product as proprietary.

Correlation must resolve significant ambiguities to provide an operationally useful synthesis of surveillance data. Typically, there are extra tracks due to false or multiple detections, missed tracks, problems due to time latency, and misassociations of targets. These errors can lead to friendly vehicles being engaged, enemy targets not being engaged, or misinterpretation of the enemy's intent. Performance evaluation of multi-target, multi-sensor tracking that centers on the use of a particular correlator must account for these potential errors. There has not yet emerged a consensus that a single, correct approach to correlation has been identified, or that its major problems have been solved. However, what is desired of a correlator is clear: it should accurately describe the surveillance space across time.

A correlator is considered to be accurate at a fixed moment in time if the following are true:

(1) Each existing object in the surveillance space that requires tracking is represented in the track registry exactly once;

(2) Each track in the registry corresponds to an existing object that requires tracking;

(3) All attribute information for every track in the registry is correct (i.e., within some tolerable error).

Accuracy at a fixed moment in time can be thought of as static accuracy. By contrast, dynamic accuracy includes integration of static accuracy features across time, and it incorporates other performance features as well:

(1) Each existing object in the surveillance space that requires tracking is represented in the track registry with exactly one track at all points in time;

(2) Each existing object in the surveillance space that requires tracking is entered into, and removed from, the track registry in a timely manner;

(3) All tracks are correctly time-tagged throughout their duration;

(4) The kinematic profile of each track, considered as a real-time object, makes sense physically. Kinematic attributes are related to the motion of objects (position, velocity, and acceleration).

Correlator accuracy is not achievable in the absolute. Measurements obtained from sensors such as radars are subject to both systematic error (bias) and random error. Each of these errors affects the association logic that a correlator uses to match measurements to tracks.

C. EVALUATION OF CORRELATORS

A correlator is one of several important components of a multi-target, multi-sensor air surveillance system. An evaluation of its performance should be based on the end result of using the correlator; in other words, on the accuracy of tracking. However, errors in tracking are not necessarily attributable to the correlator. Tracking errors can arise due to bias in the sensors, to the random measurement error that is always present in tracking, and to the uncertainty in making associations between sensor measurements and tracks that is also always present.

Nonetheless, accuracy of tracking can be used as a criterion for evaluating the relative performance of one correlator to another, provided that testing is conducted with correlators used in identical scenarios. That way, differences in performance can be attributed to the correlators, and not to another component of the surveillance system. Using modeling and simulation (M&S) under scripted scenarios, testing can provide an information base that allows a comprehensive comparison of correlators to be made. Scripted scenarios can be repeated many times, under identical conditions. And, M&S offers cost and safety advantages over live testing that makes it an attractive option for testing the performance of an air surveillance system.

D. PURPOSE

The purpose of this thesis is to develop and assess performance metrics that can be used for the evaluation and comparison of correlators in the context of air surveillance. The research described in this thesis was originally designed to meet the needs of the United States Army Space & Missile Defense Command (SMDC) Battle Lab (BL) Exercises & Training Division, which had purchased, or identified as candidates for procurement, several correlators for its air surveillance system. This need arose from the Exercises & Training Division's development of a Future Operational Capability (FOC), the purpose of which was to meet U.S. Army air defense command and control center requirements that include:

(1) Reducing the size of current air defense command and control centers;

(2) Providing a SIAP;

(3) Providing advanced visualization;

(4) Enhancing communications capabilities.

The FOC utilizes the Advanced Warfare Environment (AWarE) software package, which uses a correlator to create the SIAP. Currently, the AWarE uses the SOLIPSYS MSCT to create the SIAP. However, the SMDC BL Exercises & Training Division does not have a method for assessing the accuracy of correlators that they use, or that they consider for procurement.

E. EXPECTED BENEFITS OF THIS THESIS

Metrics for assessing the accuracy of tracking of a surveillance system in the dynamic sense can be developed relative to a period of time in which the correlator is exercised. A test that compares correlators under identical conditions can provide a basis for determining relative performance of the correlators. The metrics obtained from testing can be used to determine how well correlators perform when faced with the many issues that maneuvering aircraft or closely spaced aircraft can pose to air tracking.

II. MEASURES OF PERFORMANCE FOR AIR SURVEILLANCE

The concept of air surveillance is illustrated in Figure 2 with a simple example. For this example, there are two truth objects, consisting of two fighter aircraft, and one sensor platform, a PATRIOT FP. The PATRIOT FP collects sensor measurements within the area of interest with its AN/MPQ-53 phased array radar.

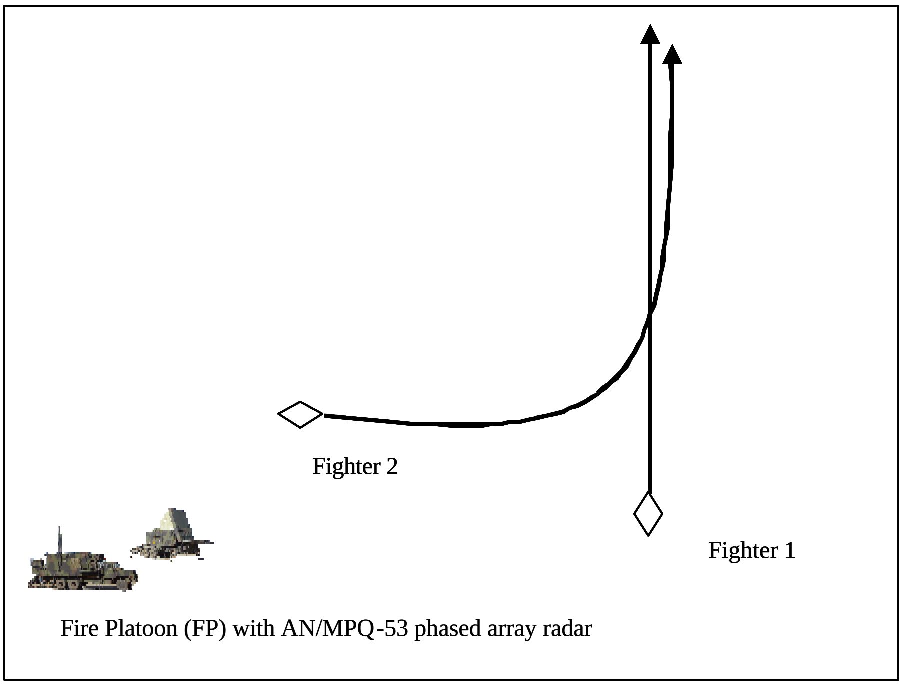

Figure 2. Air Surveillance example with two truth objects and one sensor platform.

In the scenario, Fighter 1 flies due north on a heading of 360 degrees. Fighter 2 originally flies due east, then banks 90 degrees to the left, joining up with Fighter 1 to fly in formation 50 meters apart. Tracking produced by the sensor platform is subject to errors that include the following:

(1) **Extra tracks**: when the two aircraft maneuver into formation and Fighter 2 flies to the right side of Fighter 1, the sensor measurements of the two aircraft may produce redundant or extra tracks.

(2) **Missed tracks**: when the two aircraft are flying in formation, the sensor platform may produce only one track for both aircraft, thereby missing one of the aircraft.

(3) **Swapped tracks**: as the two aircraft fly along in formation, the sensor platform tracks may switch back and forth between the aircraft.

(4) **Broken tracks**: as Fighter 2 maneuvers to the left, the sensor platform track for Fighter 2 could cease, thereby becoming a broken track.

(5) Target **position and velocity errors** while tracking the aircraft: the perceived target positions and velocities of the sensor platform will be different from the actual ground truth target positions and velocities.

The potential for error is increased at times when aircraft are engaged in maneuvers or when they are closely spaced. When an aircraft maneuvers, the likelihood of sensor measurement errors increases due to the fact that measurements of kinematic variables in the presence of maneuvers do not carry enough information for reliable correlation (Bar-Shalom, 1995, p. 194). When the two fighters "are close enough in the measurement space, they will give only one merged (unresolved) measurement due to the inherent finite resolution capability of any signal processor/detector" (Bar-Shalom, 1995, p. 355). By comparing ground truth data with the output of the tracking system, performance evaluation can be done with one sensor platform.

To illustrate the same issues with a multi-sensor tracking system, consider the same example as above, but with two sensor platforms (e.g., a PATRIOT battalion and an AWACS aircraft). Each sensor platform conducts its own "local" tracking, the results of which are stored so that the platforms can interoperate with each other to form a SIAP. Two concepts of interoperability are currently under development by the Department of Defense (DoD):

(1) Joint Data Network (JDN), based on TADIL A, B, and J messaging;

(2) Joint Composite Tracking Network (JCTN), which is at an earlier stage of development than JDN.

The goal of interoperability is to provide joint, or composite tracks, that allow for the development of a SIAP that is the same across platforms. Each platform correlates its sensor information to the registry of joint (composite) tracks using its own correlator. A registry of joint (composite) tracks is maintained separately by each platform. In theory, each platform's registry should agree, because:

(1) They are supposed to follow the same rules for managing the registry;

(2) They are supposed to be in constant communication with each other.

Performance evaluation with interoperating platforms entails evaluating each platform's composite tracks, the same as is done with a single platform. Performance metrics can then be pooled (averaged) across platforms if the same correlator is used by all platforms. A performance evaluation can also be made for a single sensor platform (e.g. PATRIOT

FP) that interoperates with other platforms (e.g. AWACS, AEGIS, etc) when the objective is to evaluate the correlator used by that platform.

A. ISSUES IN MEASURING PERFORMANCE

In a performance evaluation of an air surveillance system, a test event is designed to exercise the system. A test event can either be an episode of live aircraft flight, or it can be a modeling and simulation (M&S) scenario. M&S was used for the development of the performance metrics described in this thesis. A typical M&S test event lasts from 15 to 20 “event” minutes. However, the clock time required to execute the simulation can vary substantially from the event time, due to computer hardware and other constraints. The number of aircraft participating in a test event is determined by the level of complexity that one wants to present to the air surveillance system. A larger number of aircraft usually presents a more difficult challenge, especially if the aircraft conduct abrupt (high-G) maneuvers or fly in close formations. Performance metrics are calculated at scheduled “scoring” times during the simulation. Scoring times can be chosen to be random times, fixed-interval times, or user-specified times. The scoring times were scheduled at random times for the test event described in Chapter III of this thesis.

As noted above, instances where aircraft perform maneuvers pose a challenge to a tracking system. A maneuver is recognized at time t when the norm of the difference of the velocity direction of the aircraft from time t to time $t + m$ is greater than a. The values of m and a are “tuning parameters” that must be specified by the tester. For example, choosing m = 10 and a = 0.5858 defines a maneuver to be a turn of at least 45 degrees within 10 seconds. In this thesis, the values m = 10 and a = 0.5 are used.

Similarly, aircraft that are closely spaced to one another are difficult to track accurately. Closely spaced aircraft can give one merged, unresolved measurement to a sensor. As aircraft converge and diverge upon each other, false tracks, missed tracks and swapped tracks become more likely. Closely spaced objects are defined as two or more aircraft less than ß meters apart. The value of ß is a tuning parameter that is chosen by the tester. In this thesis, the value ß = 100 is used.

Before the performance metrics can be evaluated, tracks must be associated to truth objects. This can be done using any of a number of association methods. In this thesis, the method used was a two-dimensional assignment algorithm, which uniquely assigns tracks to truth objects at each scoring time, independent of the associations made at other times (Rothrock, 2000, p. 63). This assignment algorithm minimizes a cost function determined from the three-dimensional Euclidean distance (squared) between tracks and truth objects. The performance metrics are then computed at each scoring time based on the associated truth objects. This methodology allows both track breaks and track swaps to occur, which are defined as follows:

(1) A **track break** at time t occurs when there is a track assigned to a truth object at time t, but there is no track assigned to the same truth object at time $t + x$, where x is chosen by the tester.

(2) A **track swap** occurs when one track is assigned to a truth object at time t, but a different track is assigned to the same truth object at time $t + x$, with x chosen by the tester.

Drummond (1999) classified association algorithms into four types:

Methodology 1. Assign tracks to truth objects at pre-selected times independent of the assignment at other times and without constraints on the number or types of track swaps.

Methodology 2. No track swaps allowed. A limitation of this methodology is that it does not permit the assignment of a sequence of tracks to a target.

Methodology 3. "Feasible" track sequences allowed. The intent is to permit track swaps but only under very limited conditions. A sequence of tracks can be assigned to a target if the sequence is feasible. In a feasible track sequence, no two tracks exist for the target at the same time.

Methodology 4. Track swaps discouraged. The intent is to discourage track swaps but to achieve this without computational complexity. The concept is to use ad hoc methods in conjunction with Methodology 1 to reduce track swaps rather than use the more rigorous approach of Methodology 3 that is computationally complex. An example of an ad hoc method is to reduce the cost of the current candidate track-target pair that was assigned the last time.

B. DEVELOPMENT OF PERFORMANCE METRICS

The JCTN Pilot Benchmark Environment (PBE) (Rothrock, 2000) uses ten metrics for evaluating multi-sensor, multi-platform tracking and data association performance. The JCTN PBE is an event-driven computer simulation run in MATLAB. The JCTN PBE is described in detail in Chapter III subsection B.1. Six of the JCTN metrics that are applicable to meeting the objectives of this thesis are described in subsection B.1. Four additional metrics, which were developed as part of the thesis research, are described in subsection B.2. In order to define the metrics described below, the following classification is applied to each track and/or truth object in the test scenario at each scoring time:

(1) **Valid track**. A (composite) track uniquely assigned to a truth object.

(2) **Extra track**. A redundant track not assigned to any truth object.

(3) **Missed track**. A truth object with no (composite) track assigned to it.

1. JCTN Pilot Benchmark Performance Metrics

JCTN-1. Composite Completeness (Rothrock, 2000, p. 65) is the proportion of truth objects (real objects that should be tracked) that are held as declared composite tracks at each scoring time (t_{score}) in the scenario run. JCTN metrics refer to composite tracks, but the metrics can also be used by a single platform or sensor that uses “local” tracks. The final result for each scoring time is obtained by averaging the results obtained at the scoring time over the number of Monte Carlo runs. The Composite Completeness Metric is a function of t_{score}, which can be plotted against time. The

Composite Completeness Metric can also be averaged over all sensor platforms if the sensor platforms use the same correlator.

JCTN-1. Composite Completeness Metric		
For each sensor platform, calculate:		
$\text{Completeness}(t_{score})$	=	$\dfrac{\#\ \text{Valid Tracks}(t_{score})}{\#\ \text{Truth Objects}(t_{score})}$
<u>Notation:</u>		
t_{score}	=	Scoring time (number of seconds from beginning of test event)
$\#\ \text{Valid Tracks}(t_{score})$	=	Total number of valid tracks in the registry at time t_{score} of the test event
$\#\ \text{Truth Objects}(t_{score})$	=	Total number of truth objects in existence at time t_{score} of the test event

Table 1. Composite Completeness Metric.

The Average Composite Completeness Metric is calculated by averaging the Composite Completeness Metric over all scoring times.

JCTN-1a. Average Composite Completeness Metric		
For each sensor platform, calculate:		
$\text{Average Completeness}$	=	$\dfrac{1}{T}\sum_{t \in S} \text{Completeness}(t)$
<u>Notation:</u>		
$\text{Completeness}(t)$	=	Composite Completeness Metric evaluated at time t of the test event
S	=	The set of scoring times used for the test event
T	=	The number of scoring times in S

Table 2. Average Composite Completeness Metric.

JCTN-2. Composite Redundant Track Mean Ratio (Rothrock, 2000, p. 69) is calculated as the number of composite tracks that can be feasibly assigned to a truth object divided by the number of valid composite tracks at each scoring time in the scenario run. The final result for each scoring time is obtained by averaging the results obtained at each sensor platform and at the scoring time over the number of Monte Carlo runs. The Composite Redundant Track Mean Ratio is a function of t_{score}, which can be plotted against time. The Composite Redundant Track Mean Ratio Metric can also be averaged over all sensor platforms if the sensor platforms use the same correlator.

JCTN-2. Composite Redundant Track Mean Ratio Metric

For each sensor platform, calculate:

$$\text{Redundant Track Ratio}\,(t_{score}) = \frac{\#\,\text{Assignable Tracks}\,(t_{score})}{\#\,\text{Valid Tracks}\,(t_{score})}$$

Notation:

t_{score}	=	Scoring time (number of seconds from beginning of test event)
# Valid Tracks (t_{score})	=	Total number of valid tracks in the registry at time t_{score} of the test event
# Assignable Tracks (t_{score})	=	Total number of feasible tracks in existence at time t_{score} of the test event

Table 3. Composite Redundant Track Mean Ratio Metric.

The Average Composite Redundant Track Mean Ratio Metric is calculated by averaging the Composite Redundant Track Mean Ratio Metric over all scoring times.

JCTN-2a. Average Composite Redundant Track Mean Ratio Metric		
For each sensor platform, calculate:		
Average Redundancy	=	$\frac{1}{T}\sum_{t \in S} \text{Redundant Track Ratio } (t)$
Notation:		
Redundant Track Ratio (t)	=	Composite Redundant Track Mean Ratio Metric evaluated at time t of the test event
S	=	The set of scoring times used for the test event
T	=	The number of scoring times in S

Table 4. Average Composite Redundant Track Mean Ratio Metric.

JCTN-3. Composite Spurious Track Mean Ratio (Rothrock, 2000, p. 69) is equal to the number of unassignable composite tracks (tracks that can not be feasibly assigned) divided by the number of valid composite tracks at each scoring time in the scenario run. The final result for each scoring time is obtained by averaging the results obtained at the scoring time over the number of Monte Carlo runs. The Composite Spurious Track Mean Ratio is a function of t_{score}, which can be plotted against time. The Composite Spurious Track Mean Ratio Metric can also be averaged over all sensor platforms if the sensor platforms use the same correlator.

JCTN-3. Composite Spurious Track Mean Ratio Metric		
For each sensor platform, calculate:		
$\text{Spurious Track Ratio } (t_{\text{score}})$	=	$\dfrac{\#\,\text{Unassignable Tracks } (t_{\text{score}})}{\#\,\text{Valid Tracks } (t_{\text{score}})}$
Notation:		
t_{score}	=	Scoring time (number of seconds from beginning of test event)
$\#\,\text{Valid Tracks } (t_{\text{score}})$	=	Total number of valid tracks in the registry at time t_{score} of the test event
$\#\,\text{Unassignable Tracks } (t_{\text{score}})$	=	Total number of infeasible tracks in existence at time t_{score} of the test event

Table 5. Composite Spurious Track Mean Ratio Metric.

The Average Composite Spurious Track Mean Ratio Metric is calculated by averaging the Composite Spurious Track Mean Ratio Metric over all scoring times.

JCTN-3a. Average Composite Spurious Track Mean Ratio Metric		
For each sensor platform, calculate:		
$\text{Average Spuriousness}$	=	$\dfrac{1}{T}\sum_{t \in S} \text{Spurious Track Ratio } (t)$
Notation:		
$\text{Spurious Track Ratio } (t)$	=	Composite Spurious Track Mean Ratio Metric evaluated at time t of the test event
S	=	The set of scoring times used for the test event
T	=	The number of scoring times in S

Table 6. Average Composite Spurious Track Mean Ratio Metric.

JCTN-4. Mean Cumulative Swaps of Composite Tracks (Rothrock, 2000, pp. 67-68) is calculated by computing the number of composite track swaps, which is the number of times that the composite track number assigned to each truth object has changed during the scenario run. This number is also averaged over all of the truth objects. First, determine the composite track number assigned to each truth object at time t_{score}. If Track A was assigned to object j at each of the last three scoring times and Track B (with Track A not equal to Track B) is assigned to object j at the current scoring time in Monte Carlo run m, then increment by one the number of swaps for object j, $\text{NS}_{j,m}(t_{\text{score}})$. For each truth object, the cumulative number of track swaps at each t_{score} are averaged over the number of Monte Carlo runs. $\text{NS}_j(t_{\text{score}})$ is also averaged over all truth objects in the scenario at time t_{score}. The Mean Cumulative Swaps of Composite Tracks is a function of t_{score}, which can be plotted against time.

JCTN-4. Mean Cumulative Swaps of Composite Tracks Metric

For each sensor platform, calculate:

$$\text{NS}_j(t_{\text{score}}) = \frac{1}{M}\sum_{m=1}^{M}\text{NS}_{j,m}(t_{\text{score}})$$

$$\text{NS}(t_{\text{score}}) = \frac{1}{L}\sum_{j=1}^{L}\text{NS}_j(t_{\text{score}})$$

Notation:

t_{score}	=	Scoring time (number of seconds from beginning of test event)
$\text{NS}_j(t_{\text{score}})$	=	Total number of track swaps for truth object j
$\text{NS}(t_{\text{score}})$	=	Total number of track swaps at time t_{score}
M	=	Total number of Monte Carlo Runs
L	=	Total number of truth objects

Table 7. Mean Cumulative Swaps of Composite Tracks Metric.

The Average Total Number of Cumulative Swaps of Composite Tracks Metric is calculated by averaging the total number of cumulative swaps of composite tracks over all sensor platforms.

JCTN4-a. Average Total Number of Cumulative Swaps of Composite Tracks Metric		
Average Swaps of Composite Tracks	=	$\frac{1}{P}\sum_{P} \mathrm{NS}(LS)$
<u>Notation:</u>		
$\mathrm{NS}(t_{\mathrm{score}})$	=	Total number of track swaps at time t_{score}
LS	=	The last scoring time used for the test event
P	=	The number of sensor platforms

Table 8. Average Total Number of Cumulative Swaps of Composite Tracks Metric.

JCTN-5. Mean Cumulative Broken Composite Tracks (Rothrock, 2000, p. 68) is calculated by counting the number of composite track breaks for each truth object during the scenario run. The cumulative number of track breaks during the scenario run is also averaged over all of the truth objects. First, determine the composite track number assigned to each truth object at time t. If track A was assigned to object j at each of the last three scoring times and no track is assigned to object j at the current scoring time in Monte Carlo run m, then increment by one the number of breaks for object j, $\mathrm{NB}_{j,m}(t_{\mathrm{score}})$. For each truth object, the cumulative number of track breaks at each t_{score} are averaged over the number of Monte Carlo runs. $\mathrm{NB}_j(t_{\mathrm{score}})$ is also averaged over all truth objects in the scenario at time t_{score}. The Mean Cumulative Broken Composite Tracks is a function of t_{score}, which can be plotted against time.

JCTN-5. Mean Cumulative Broken Composite Tracks Metric		
For each sensor platform, calculate:		
$\mathrm{NB}_j(t_{\text{score}})$	=	$\dfrac{1}{M}\sum_{m=1}^{M}\mathrm{NB}_{j,m}(t_{\text{score}})$
$\mathrm{NB}(t_{\text{score}})$	=	$\dfrac{1}{L}\sum_{j=1}^{L}\mathrm{NB}_j(t_{\text{score}})$
Notation:		
t_{score}	=	Scoring time (number of seconds from beginning of test event)
$\mathrm{NB}_j(t_{\text{score}})$	=	Total number of broken tracks for truth object j
$\mathrm{NB}(t_{\text{score}})$	=	Total number of track breaks at time t_{score}
M	=	Total number of Monte Carlo Runs
L	=	Total number of truth objects

Table 9. Mean Cumulative Broken Composite Tracks Metric.

The Average Total Number of Cumulative Broken Composite Tracks Metric is calculated by averaging the total number of cumulative breaks of composite tracks over all sensor platforms.

JCTN-5a. Average Total Number of Cumulative Broken Composite Tracks Metric		
Average Broken Composite Tracks	=	$\dfrac{1}{P}\sum_{P}\mathrm{NB}(LS)$
Notation:		
$\mathrm{NB}(t_{\text{score}})$	=	Total number of track breaks at time t_{score}
LS	=	The last scoring time used for the test event
P	=	The number of sensor platforms

Table 10. Average Total Number of Cumulative Broken Composite Tracks Metric.

JCTN-6. Composite Track Accuracy (Rothrock, 2000, pp. 70-71) is computed for each truth object as a function of scoring time separately for each sensor platform. It consists of four values at each scoring time: the root mean squared error (RMSE ref. Table 13) in position, the RMSE in velocity, the root sum squared average error (RSSAE ref. Table 12) in position, and the RSSAE in velocity. For each Monte Carlo run, the errors at a particular time are determined using the composite track assigned to the truth object at the scoring time. The final values at each scoring time are computed by averaging the values obtained at that time over all Monte Carlo runs.

At the beginning of the simulation, initialize $n_i(t_{score}) = 0$ for each truth object *i*. This variable is a counter that records the number of Monte Carlo runs where a composite track is assigned to a particular truth object at a specific scoring time. At each scoring time and at each Monte Carlo run, determine whether a composite track is assigned to each object i based on the results of a gated, optimal assignment. If a composite track is assigned to object i at that t_{score}, $n_i(t_{score})$ is incremented by one, and the following set of recursion updates is performed:

Composite Track Accuracy Recursion Updates For each Sensor Platform		
$e_{i,n}(t_{score}) = \hat{x}_{i,n}(t_{score}) - x_{i,truth}(t_{score})$		
$d_{i,n}(t_{score}) = e_{i,n}(t_{score}) - e_{i,aver\_n-1}(t_{score})$		
$e_{i,aver\_n}(t_{score}) = e_{i,aver\_n-1}(t_{score}) + \dfrac{d_{i,n}(t_{score})}{n_i(t_{score})}$		
$C_{i,aver\_n}(t_{score}) = \dfrac{n_i(t_{score}) - 1}{n_i(t_{score})} \times \left[C_{i,aver\_n-1}(t_{score}) + \dfrac{d_{i,n}(t_{score}) d_{i,n}(t_{score})^T}{n_i(t_{score})} \right]$		
i	=	Truth object index.
$n_i(t)$	=	Cumulative number of tracks assigned to truth object i up to and including time *t*.
$\hat{x}_{i,n}(t_{score})$	=	Six-state position/velocity column vector containing the state estimate of the composite track assigned to object *i* at time t_{score}.
$x_{i,truth}(t_{score})$	=	Six-element column vector containing the true position and velocity of object *i* at time t_{score}.
$e_{i,aver\_n}(t_{score})$	=	Column vector of average errors for object *i* assessed over *n* Monte Carlo runs.
$C_{i,aver\_n}(t_{score})$	=	Statistical covariance of the errors for object *i* assessed over *n* Monte Carlo runs (6 x 6 matrix).

Table 11. Composite Track Accuracy Metric Recursion Updates.

At the beginning of the simulation, $e_{i,aver\_0}(t_{score})$ and $C_{i,aver\_0}(t_{score})$ are initialized to 0. For each time segment where $n_i(t_{score})$ is a significant portion of the total number of Monte Carlo runs, the four metrics are computed. The RSSAE error statistics for each object *i* are computed using the following equations:

RSSAE ERROR STATISTICS EQUATIONS		
$RSSAE_{i,p}(t_{score}) = \sqrt{e_{i,p1}(t_{score})^2 + e_{i,p2}(t_{score})^2 + e_{i,p3}(t_{score})^2}$		
$RSSAE_{i,v}(t_{score}) = \sqrt{e_{i,v1}(t_{score})^2 + e_{i,v2}(t_{score})^2 + e_{i,v3}(t_{score})^2}$		
$e_{i,p1}(t_{score}), e_{i,p2}(t_{score}), e_{i,p3}(t_{score})$	=	Three position error components in $e_{i,aver\_n}(t_{score})$
$e_{i,v1}(t_{score}), e_{i,v2}(t_{score}), e_{i,v3}(t_{score})$	=	Three velocity error components in $e_{i,aver\_n}(t_{score})$
p	=	position
v	=	velocity
See Table 11 for explanation of the notation used.		

Table 12. RSSAE Error Statistics Equations.

Similarly, the RMSE error statistics for each object i are computed using the following equations:

RMSE ERROR STATISTICS EQUATIONS		
$RMSE_{i,p}(t_{score}) = \sqrt{C_{i,p11}(t_{score})^2 + C_{i,p22}(t_{score})^2 + C_{i,p33}(t_{score})^2 + RSSAE_{i,p}(t_{score})^2}$		
$RMSE_{i,v}(t_{score}) = \sqrt{C_{i,v11}(t_{score})^2 + C_{i,v22}(t_{score})^2 + C_{i,v33}(t_{score})^2 + RSSAE_{i,v}(t_{score})^2}$		
$C_{i,p11}(t_{score}), C_{i,p22}(t_{score}), C_{i,p33}(t_{score})$	=	Statistical variances (diagonal terms) of the position error components in $C_{i,aver\_n}(t_{score})$
$C_{i,v11}(t_{score}), C_{i,v22}(t_{score}), C_{i,v33}(t_{score})$	=	Statistical variances (diagonal terms) of the velocity error components in $C_{i,aver\_n}(t_{score})$
See Table 11 for explanation of the notation used.		

Table 13. RMSE Error Statistics Equations.

The Composite Track Accuracy Metric is computed and plotted separately for each sensor platform.

The Average Composite Track Accuracy Metric is calculated by averaging the Composite Track Accuracy Metric (RMSE in position and velocity, RSSAE in position and velocity) over all sensor platforms and all scoring times.

JCTN-6a. Average Composite Track Accuracy Metric

$$\text{Average Composite Track Accuracy RSSAE}_{p} = \frac{1}{NT}\sum_{P}\sum_{i}\sum_{t\in S}\text{RSSAE}_{i,p}(t)$$

$$\text{Average Composite Track Accuracy RSSAE}_{v} = \frac{1}{NT}\sum_{P}\sum_{i}\sum_{t\in S}\text{RSSAE}_{i,v}(t)$$

$$\text{Average Composite Track Accuracy RMSE}_{p} = \frac{1}{NT}\sum_{P}\sum_{i}\sum_{t\in S}\text{RMSE}_{i,p}(t)$$

$$\text{Average Composite Track Accuracy RMSE}_{v} = \frac{1}{NT}\sum_{P}\sum_{i}\sum_{t\in S}\text{RMSE}_{i,v}(t)$$

Notation:

$\text{RSSAE}_{i,p}(t)$	=	RSSAE_p metric evaluated at time *t* of the test event
$\text{RSSAE}_{i,v}(t)$	=	RSSAE_v metric evaluated at time *t* of the test event
$\text{RMSE}_{i,p}(t)$	=	RMSE_p metric evaluated at time *t* of the test event
$\text{RMSE}_{i,v}(t)$	=	RMSE_v metric evaluated at time *t* of the test event
i	=	Truth Object index.
S	=	The set of scoring times used for the test event
T	=	The number of scoring times in S
P	=	The set of sensor platforms
N	=	The total number of sensor platforms

Table 14. Average Composite Track Accuracy Metric.

2. Developed Performance Metrics (DPM)

The following metrics were developed as part of the thesis research.

DPM-1. Mean Number of Missed Targets. For each time point of interest, average the number of missed targets (number of targets - number of valid tracks) over all sensor platforms. The final result for each scoring time is obtained by averaging the results obtained at the scoring time over the number of Monte Carlo runs. The Mean Number of Missed Targets is a function of t_{score} which can be plotted against time.

DPM-1. Mean Number of Missed Targets Metric		
$\text{Missed Targets}\ (t_{\text{score}})$	=	$\#\ \text{Targets}\ (t_{\text{score}}) - \#\ \text{Valid Tracks}\ (t_{\text{score}})$
Notation:		
t_{score}	=	Scoring time (number of seconds from beginning of test event)
$\#\ \text{Valid Tracks}\ (t_{\text{score}})$	=	Total number of valid tracks in the registry at time t_{score} of the test event averaged over all platforms
$\#\ \text{Targets}\ (t_{\text{score}})$	=	Total number of targets in existence at time t_{score} of the test event

Table 15. Mean Number of Missed Targets Metric.

The Average Mean Number of Missed Targets Metric is computed by averaging the Mean Number of Missed Targets Metric over all the scoring times.

DPM-1a. Average Mean Number of Missed Targets Metric		
Average Missed Targets	=	$\frac{1}{T}\sum_{t \in S} \text{Missed Targets}\ (t)$
Notation:		
$\text{Missed Targets}\ (t)$	=	Mean Number of Missed Targets Metric evaluated at time t of the test event
S	=	The set of scoring times used for the test event
T	=	The number of scoring times in S

Table 16. Average Mean Number of Missed Targets Metric.

DPM-2. Mean Number of Extra Tracks. For each time point of interest, average the number of extra (false) tracks (number of tracks - number of valid tracks) over all sensor platforms. The final result for each scoring time is obtained by averaging the results obtained at the scoring time over the number of Monte Carlo runs. The Mean Number of Extra Tracks is a function of t_{score}, which can be plotted against time.

DPM-2. Mean Number of Extra Tracks Metric		
$\text{Extra Tracks}\,(t_{\text{score}})$	=	$\#\,\text{Tracks}\,(t_{\text{score}}) - \#\,\text{Valid Tracks}\,(t_{\text{score}})$
Notation:		
t_{score}	=	Scoring time (number of seconds from beginning of test event)
$\#\,\text{Valid Tracks}\,(t_{\text{score}})$	=	Total number of valid tracks in the registry at time t_{score} of the test event
$\#\,\text{Tracks}\,(t_{\text{score}})$	=	Total number of tracks in existence at time t_{score} of the test event

Table 17. Mean Number of Extra Tracks Metric.

The Average Mean Number of Extra Tracks Metric is computed by averaging over all scoring times.

DPM-2a. Average Mean Number of Extra Tracks Metric		
$\text{Average Extra Tracks}$	=	$\frac{1}{T}\sum_{t \in S} \text{Extra Tracks}\,(t)$
Notation:		
$\text{Extra Tracks}\,(t)$	=	Mean Number of Extra Tracks Metric evaluated at time t of the test event
S	=	The set of scoring times used for the test event
T	=	The number of scoring times in S

Table 18. Average Mean Number of Extra Tracks Metric.

DPM-3. Maneuver Metrics. Within a scenario, scoring times when aircraft perform maneuvers are examined separately. For each truth object, its true velocity data is checked at every second of the scenario. From time $t = 0$ until the end of the scenario, the squared norm of the difference of the true velocity from time t and time $t + m$ is calculated by:

$$\mathrm{D}^2(t,m) = \left\| \frac{\mathrm{V}_t}{\|\mathrm{V}_t\|} - \frac{\mathrm{V}_{t+m}}{\|\mathrm{V}_{t+m}\|} \right\|^2$$

$\|\mathrm{V}_t\|$ is the norm of the velocity vector at time t. If $\mathrm{D}^2(t,m)$ is larger than a, a maneuver is judged to have occurred at time t. For example, if $\mathrm{D}^2(t,m)$ is greater than 0.5858, then the aircraft made at least a 45 degree turn within m seconds. In this thesis, the values m= 10 and a = 0.5 are used to detect maneuvers. For each truth object, times when $\mathrm{D}^2(t,m)$ are greater than 0.5 is marked.

During the marked times for each truth object, positional errors (squared Euclidean distance between the composite track and the truth object), total number of track swaps and total number of track breaks are counted. For each of these marked scoring times, averages are computed over the number of Monte Carlo runs.

DPM-3. Maneuver Metrics		
For each truth object, calculate:		
Average Position Error	=	$\frac{1}{T(a, m)} \sum_{t \in S(a, m)} \text{RSSAE}_{i,p}(t)$
Average Number of Track Swaps	=	$\frac{1}{T(a, m)} \sum_{t \in S(a, m)} \text{Number of Track Swaps } (t)$
Average Number of Track Breaks	=	$\frac{1}{T(a, m)} \sum_{t \in S(a, m)} \text{Number of Track Breaks } (t)$
Notation:		
S(a, *m*)	=	All scoring times such that $D^2(t, m) > a$.
T(a, *m*)	=	# of scoring times in S(a, *m*)

Table 19. Maneuver Metrics.

A Composite Maneuver Metric, DPM-3a, is computed over all truth objects and sensor platforms for the scenario.

DPM-4. Closely Spaced Objects Metrics. Within a scenario, times when aircraft are closely spaced are also examined separately. A matrix of three-dimensional Euclidean distance (squared) in position between all truth objects with every other truth object is calculated from the true positions for every second of the scenario. Times when truth objects are within ß meters of another truth object are marked. C(*t*) is the minimum distance between two objects at time *t*. If C(*t*) < ß, then *t* is in the set of "closely spaced objects" times. In this thesis, the value ß = 100 is used to identify closely spaced objects.

During the marked times for each truth object that is closely spaced relative to another, the track swaps for each object are counted. The track swaps for each truth object throughout the entire scenario are also counted. For each of these marked scoring times, averages are computed over the number of Monte Carlo runs.

DPM-4. Closely Spaced Objects Metrics

For each truth object, calculate:

Average Number of Track Swaps	=	$\frac{1}{M}\sum_{m=1}^{M} \mathrm{NS}_{j,m}(LS)$
Average Number of Track Swaps in Closely Spaced Objects Status	=	$\frac{1}{\mathrm{T}(t)}\sum_{t \in \mathrm{S}(t)} \text{Number of Track Swaps } (t)$

<u>Notation:</u>

$\mathrm{NS}_{j,m}(LS)$	=	Total number of track swaps for truth object j in Monte Carlo run m
M	=	Total number of Monte Carlo runs
LS	=	The last scoring time used for the test event
$\mathrm{S}(t)$	=	All scoring times such that $\mathrm{C}(t) < ß$
$\mathrm{T}(t)$	=	# of scoring times in $\mathrm{S}(t)$

Table 20. Closely Spaced Objects Metrics.

A Composite Closely Spaced Objects Metric, DPM-4a, is computed over all truth objects and sensor platforms for the scenario.

Table 21 shows over what attributes each performance metric is averaged. An "X" marked in a column means that performance metric is averaged out over that attribute. An "X*" means that performance metric can be averaged out over all sensor platforms if all sensor platforms used the same correlator.

	Monte Carlo Run[a]	Scoring Time	Platform	Truth Object
JCTN-1	X		X*[b]	
JCTN-1a	X	X	X	X
JCTN-2	X		X*[b]	
JCTN-2a	X	X	X	X
JCTN-3	X		X*[b]	
JCTN-3a	X	X		X
JCTN-4	X			X
JCTN-4a	X		X	X
JCTN-5	X			X
JCTN-5a	X		X	X
JCTN-6	X			
JCTN-6a	X	X	X	X
DPM-1	X		X*[b]	
DPM-1a	X	X	X	X
DPM-2	X		X*[b]	
DPM-2a	X	X	X	X
DPM-3	X		X	
DPM-3a	X		X	X
DPM-4	X		X	
DPM-4a	X		X	X

a Averaging over Monte Carlo runs would not be done in using the techniques described in Chapter V.

b X* means that the performance metric can be averaged out over all sensor platforms if all sensor platforms used the same correlator.

Table 21. Performance metric aggregation.

The performance metrics described in this chapter provide a basis for an evaluation of an air surveillance system. A comparison of correlators can be made, based on the metrics, by using them in identical test scenarios with the same scoring times. The metrics are designed to evaluate tracking with maneuvering and closely spaced aircraft to give a detailed summary of the relative performance of correlators under difficult circumstances. The Maneuver Metrics (DPM-3) and the Closely Spaced Objects Metrics (DPM-4) can be used to evaluate tracking performance when faced with these difficult circumstances.

THIS PAGE INTENTIONALLY LEFT BLANK

III. PERFORMANCE EVALUATION USING MODELING AND SIMULATION

The development of a test event forms the basis of an experiment in which a tracking system can be evaluated, and correlators compared. There are two kinds of test events:

(1) In **live tests**, real aircraft fly in formations and maneuver for a set period of time. Sensor platforms track the aircraft as they fly in the designated area.

(2) In **Modeling and Simulation (M&S) tests**, computer-generated objects simulate the flight of aircraft for a set period of time. Sensor platforms track what the event simulator suggests are aircraft.

Live tests are generally regarded as more realistic than M&S tests, but they also entail disadvantages (Law and Kelton, 2000, pp. 91-92):

(1) Live tests are more expensive to conduct than M&S test;

(2) Live tests entail risks to the personnel flying the aircraft, especially while maneuvering;

(3) It is more difficult to obtain accurate truth data with live tests than with M&S;

(4) It is virtually impossible to replicate the conditions of a live test, while M&S tests can be replicated indefinitely (subject to time and cost considerations).

Therefore, M&S testing was used as the basis for the thesis research. The steps in designing M&S test scenarios are as follows:

(1) <u>Identify an M&S environment that is suitable for the task</u>. The Extended Air Defense Simulation (EADSIM) and the JCTN Pilot Benchmark Environment (PBE) were considered because they provided the functionality that is required.

(2) <u>Develop scripted scenarios</u>. In EADSIM, scenarios can be developed to include any number of aircraft, friendly and enemy objects, and any number of sensor platforms. In the JCTN PBE, scenarios must use a variation of the aircraft and sensor platforms provided because the objects are modeled at a very fine level of detail, one that would be extremely difficult for a user to develop new objects.

(3) <u>Decide on the length of time per simulation (in event seconds) and the number of replications used</u>. The length of time per simulation must capture the full range of events required. The larger the number of replications used, the more reliable the results of the simulation are.

(4) <u>Integrate a correlator into the simulations</u>. In EADSIM, this is done after-the-fact, based on the raw data that EADSIM provides. In the JCTN PBE, the correlator is integrated into the simulations as they are run.

(5) <u>Calculate performance metrics</u>. In EADSIM and the JCTN PBE, this is done after-the-fact based on the simulation output.

A. MODELING AND SIMULATION USING EXTENDED AIR DEFENSE SIMULATION (EADSIM)

1. Data Collection

EADSIM is an analytic model of air and missile warfare used for scenarios ranging from few-on-few to many-on-many forces. Each platform (such as a fighter aircraft) is individually modeled, as is the interaction among platforms. It models the Command and Control (C2) decision processes and the communications among the platforms on a message-by-message basis (Teledyne Brown Engineering, 1998). As part of the thesis research, five different scenarios were developed in EADSIM with varying sets of sensors and objects. The theater of operations was northwest Europe. All sensors were un-netted, not linked by a joint network, and had some systematic error. Object sets included varying numbers of enemy and friendly aircraft within the area of interest. Table 22 summarizes the five different scenarios. Figure 3 is an illustration of Scenario 1 generated in EADSIM.

	Number of Friendly Fighters	Number of Enemy Fighters	Sensor Platforms
Scenario 1	1	1	1 AWACS platform 1 AEGIS platform 3 PATRIOT FPs
Scenario 2	4	3	1 AWACS platform 1 AEGIS platform 3 PATRIOT FPs
Scenario 3	8	8	1 AWACS platform 1 AEGIS platform 3 PATRIOT FPs
Scenario 4	11	11	1 AWACS platform 3 AEGIS platforms 5 PATRIOT FPs
Scenario 5	7	6	1 AWACS platform 3 AEGIS platforms 5 PATRIOT FPs

Table 22. Example of Scripted Scenarios Developed in EADSIM.

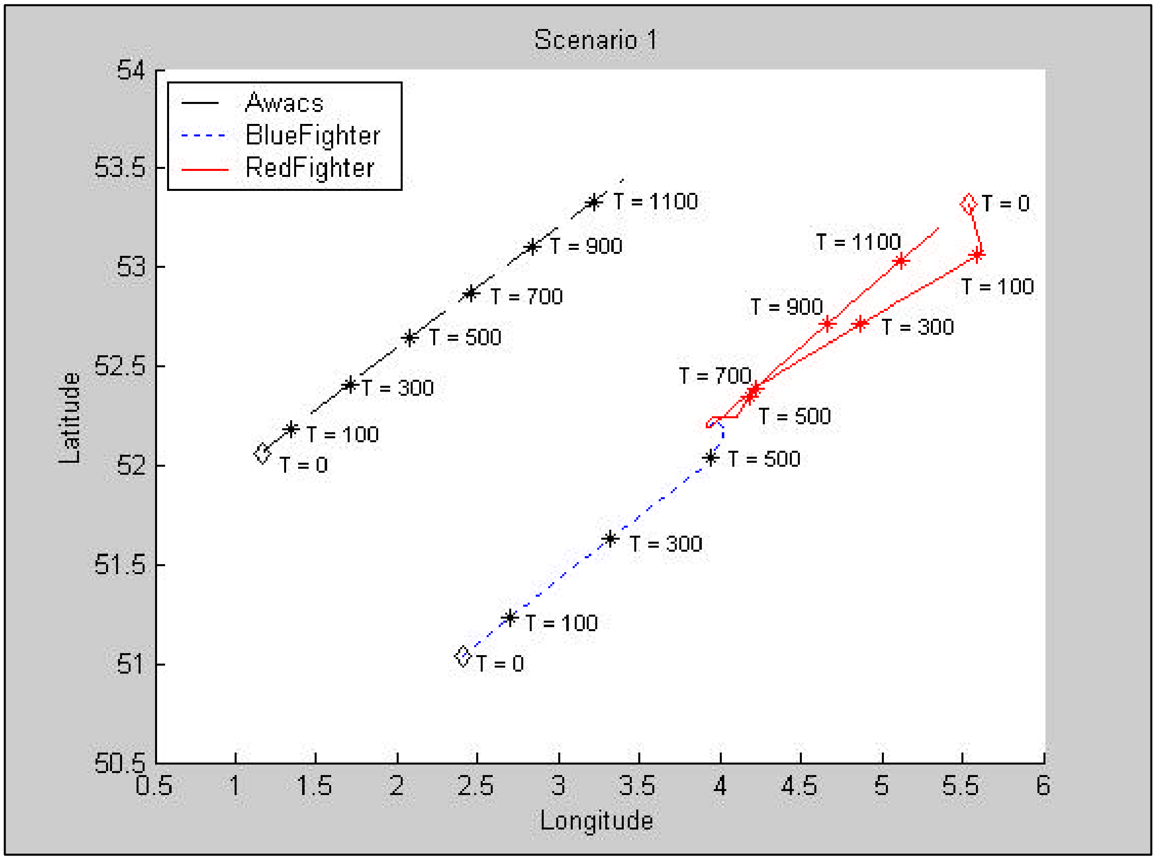

Figure 3. Scenario 1 developed in EADSIM.

The simulations were run on EADSIM, which generated simulated information known as Protocol Data Units (PDUs). The PDUs were the sensor-perceived truth of each object for each scenario. Sensor errors were included for each sensor platform. The PDUs were then sent through the Tactical Simulation Interface Unit (TSIU), which translates the PDUs from the simulated environment into appropriate tactical message formats, such as TADIL-J, that can be sent and used by different military workstations. Military workstations, such as the AWarE system (the baseline software package that provides the FOC with an overall architecture) and the Air Defense Systems Integrator (ADSI), then correlate tracks and produce a composite picture (SIAP). Data are then collected from the AWarE system that correlated the tactical messages. The ground truth data and the tracking data are then used in evaluating the performance metrics. However, EADSIM data were not received in correlated form prior to the completion of this thesis, and were therefore not used in the research.

2. Performance Evaluation with EADSIM

Once the perceived truth data are collected from the AWarE system, the methodology for evaluating the performance metrics involves two steps. The first step is to select the target for each track. The first step of the analysis is to use a two-dimensional assignment algorithm to uniquely assign tracks (from the AWarE system) to targets (from the ground truth of each object for each scenario). From this assignment of track-target pairs, valid tracks are identified, unassigned tracks are labeled as extra tracks, and unassigned targets are labeled as missed tracks. The second step is then to evaluate the performance metrics for the assignment of tracks to truth objects at each scoring time.

B. EVENT SIMULATION USING JCTN PILOT BENCHMARK ENVIRONMENT (PBE)

1. Data Collection

Because data could not be collected from the AWarE system, JCTN PBE was used to provide the scripted scenario. JCTN PBE is an event-driven computer simulation, run in MATLAB, that provides the functionality required to test models of multi-platform, multi-sensor tracking algorithms (Rothrock, 2000, p. 3). The JCTN PBE provides base scenarios for use with its event simulator. The base scenario considered in this thesis has a duration of twenty minutes of "event" time. The scenario allows the user to select different aircraft and sensors for the simulation. There are a total of nine aircraft to choose from, consisting of two airborne tankers, four fighter aircraft, and three commercial airliners. There are six sensor platforms to choose from, consisting of four ships and two aircraft.

OBJECT	NUMBER AVAILABLE
Air Surveillance Platform	2
Ship Surveillance Platform	4
Fighter Aircraft	4
Tanker Aircraft	2
Commercial Airliner	3

Table 23. Example of Aircraft and Sensors available in the JCTN PBE Base Scenario.

Each ship has an S-band phased array radar and a UHF rotating radar. Each airborne platform has a single Airborne UHF rotating radar. The flight paths of all aircraft are predetermined and follow the same paths each time that the simulation is run. The base scenario used to generate the data used in the evaluation of the performance metrics included 2 air surveillance platforms, 2 ship surveillance platforms, 4 fighter aircraft, and 2 tanker aircraft. The event time of each simulation was 20 minutes. Figure 4 illustrates the base scenario. Aircraft starting positions are marked with diamonds.

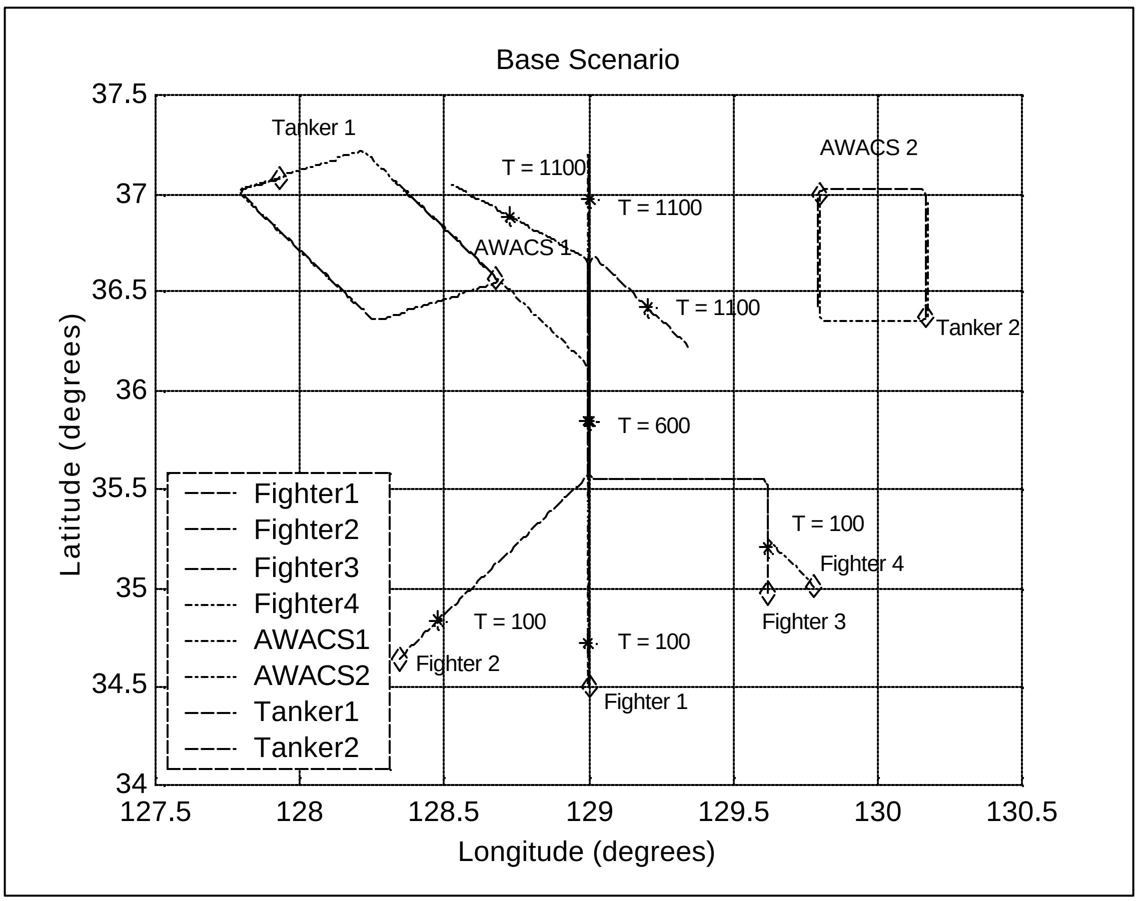

Figure 4. The JCTN PBE base scenario used for data collection.

2. Performance Evaluation with JCTN PBE

A single execution of the JCTN PBE consists of multiple Monte Carlo runs, with the number of Monte Carlo runs determined by the user. Composite tracks for each sensor platform vary across Monte Carlo runs. For the data analysis, twenty Monte Carlo runs each conducted. The metrics were evaluated at scheduled scoring times during the simulations. These scoring times were set at the beginning of the simulations, and the same scoring times were used for each. The scoring times were randomly selected using the MATLAB random number generator.

Calculation of performance metrics requires that tracks be associated to truth objects. In JCTN PBE, this was done using a Jonker-Volgenent-Castanon (JVC) two-dimensional assignment algorithm (Rothrock, 2000, p. 64) to uniquely assign tracks to targets. The two-dimensional assignment algorithm minimizes a cost function for each possible pairing of a composite track to a truth object. The default cost function in JCTN PBE is the three dimensional Euclidean distance (squared) in position between a composite track and a truth object. JCTN PBE performed a gated, optimal assignment of each platform's composite tracks to truth objects at each scoring time. Each track was assigned to no more than one truth object. At each scoring time, each composite track was classified as one of the following:

a. Valid track (composite track uniquely assigned to truth object).

b. Extra track (redundant track not assigned to any truth object).

c. Missed track (truth object with no composite track assigned to it).

Extra tracks and missed tracks were counted as errors against the test tracking algorithm.

After the assignment costs are computed, the JCTN PBE performs two more operations prior to determining the optimal assignment. The first operation involves setting a threshold value to eliminate any unlikely pairings. For the default Euclidean distance (squared), the threshold is set to 2×10^8. Any costs greater than this threshold in the cost matrix are set to infinity. The second operation creates an additional entry in the cost matrix in each column. These entries are set to the threshold value and the entries correspond to the cost of not assigning a track or a truth object to anything. This allows the algorithm to not make an assignment for a track or a truth object if the cost function shows that a pairing of a track with a truth object is unlikely.

IV. ANALYSIS OF PERFORMANCE EVALUATION DATA

A. DATA GENERATION

The JCTN Pilot Benchmark Environment (PBE), Release 1.08.02 was used to generate the data for the thesis research. A computer equipped with an Intel Pentium III processor (1 GHz) and 192 MB of RAM was used to run the simulations. MATLAB version 5.3 was used to run the JCTN Pilot Benchmark software. A single Monte Carlo run of the scenario required approximately one hour of clock time to execute. Random numbers were generated using the MATLAB random number generator, with seeds initiated by MATLAB upon invocation of the software. The metric data files produced from the simulations were in the form of MATLAB MAT-files.

B. COMPOSITE COMPETENESS RESULTS

The Composite Completeness Metric (JCTN-1) was obtained by averaging the results of each sensor platform at each scoring time for each simulation. Figure 5 shows a plot of the Composite Completeness Metric for each scoring time for the first ten Monte Carlo runs.

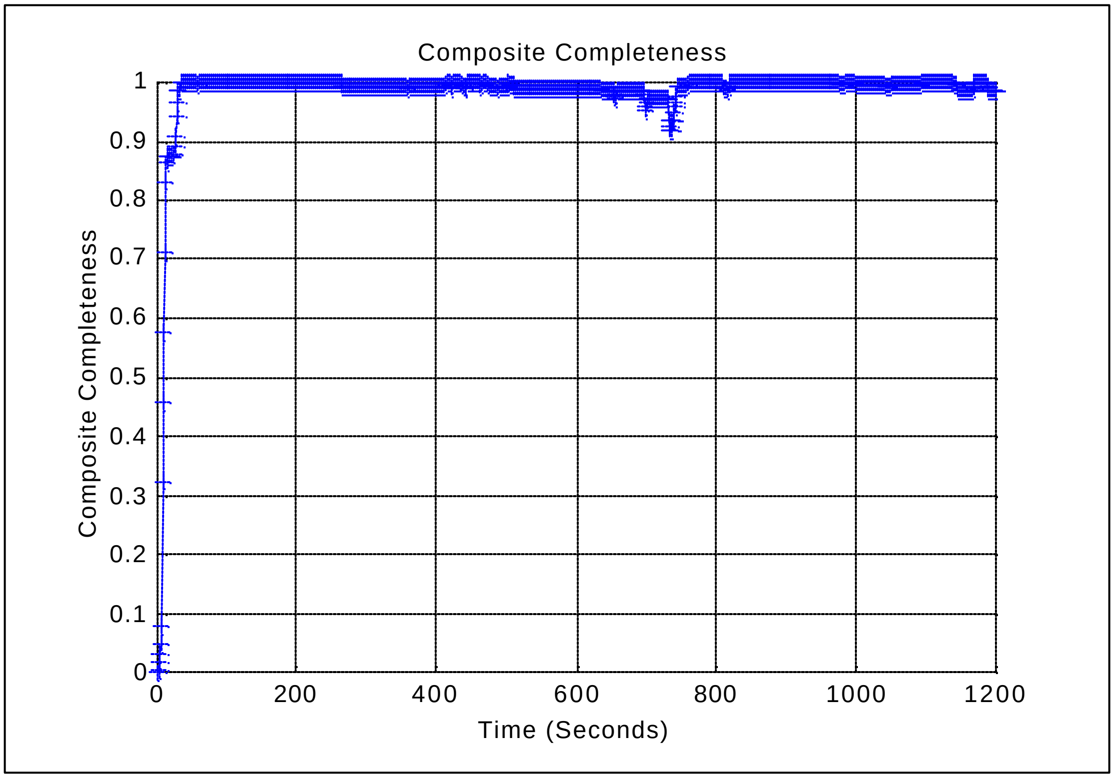

Figure 5. Composite Completeness (JCTN-1) for the first ten Monte Carlo runs. The metric was averaged across all four sensor platforms.

The Composite Completeness Metric (JCTN-1) reaches 100 per cent at time 35 and is steady except for a brief drop to 92 per cent around time 730 which is when Fighter 4 breaks formation from the other three fighters. The Average Completeness Metric (JCTN-1a) for the first ten Monte Carlo runs is 0.99.

C. COMPOSITE REDUNDANT TRACK MEAN RATIO RESULTS

The Composite Redundant Track Mean Ratio Metric (JCTN-2) was obtained by averaging the results of each sensor platform at each scoring time for each simulation.

Figure 6 shows a plot of the Composite Redundant Track Mean Ratio Metric for each scoring time for the first ten Monte Carlo runs.

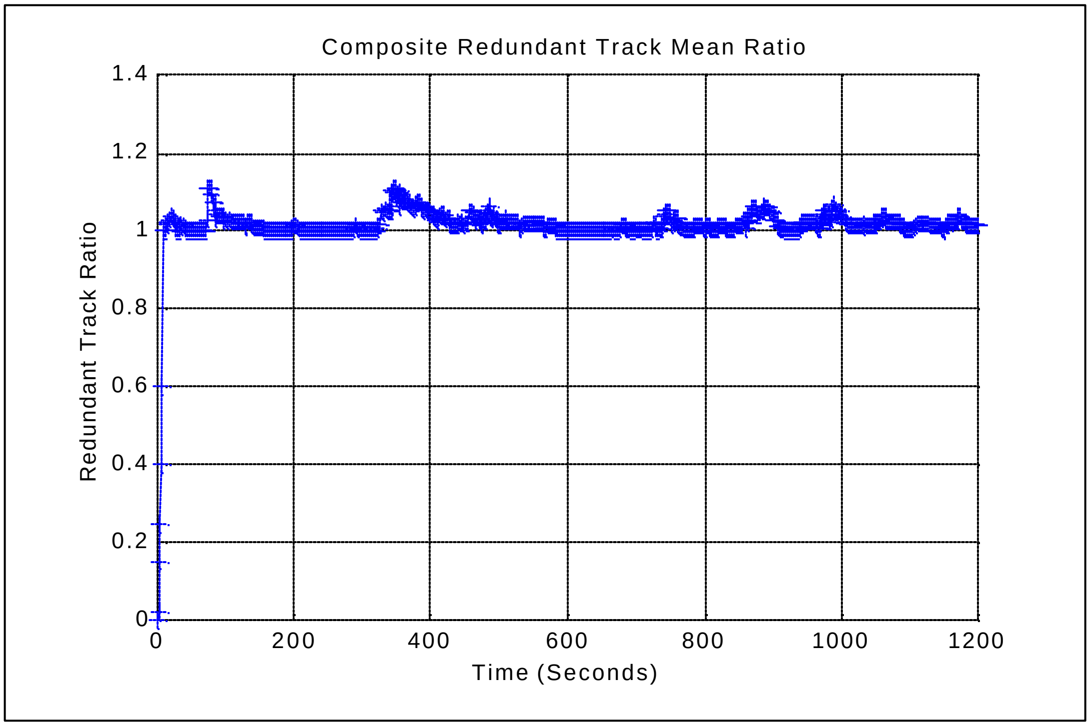

Figure 6. Composite Redundant Track Mean Ratio (JCTN-2) for the first ten Monte Carlo runs. The metric was averaged across all four sensor platforms.

Values less than one imply that there are too few assignable tracks. Values equal to one imply that the number of assignable tracks equal the number of valid tracks. Values larger than one imply that there are redundant tracks. The Composite Redundant Track Mean Ratio Metric is close to 1 for the first ten Monte Carlo runs. The Average Redundancy Metric (JCTN-2a) for the first ten Monte Carlo runs is 1.02.

D. COMPOSITE SPURIOUS TRACK MEAN RATIO

The Composite Spurious Track Mean Ratio Metric (JCTN-3) was obtained by averaging the results of each sensor platform at each scoring time for each simulation. Figure 7 shows a plot of the Composite Spurious Track Mean Ratio Metric for each scoring time for the first ten Monte Carlo runs.

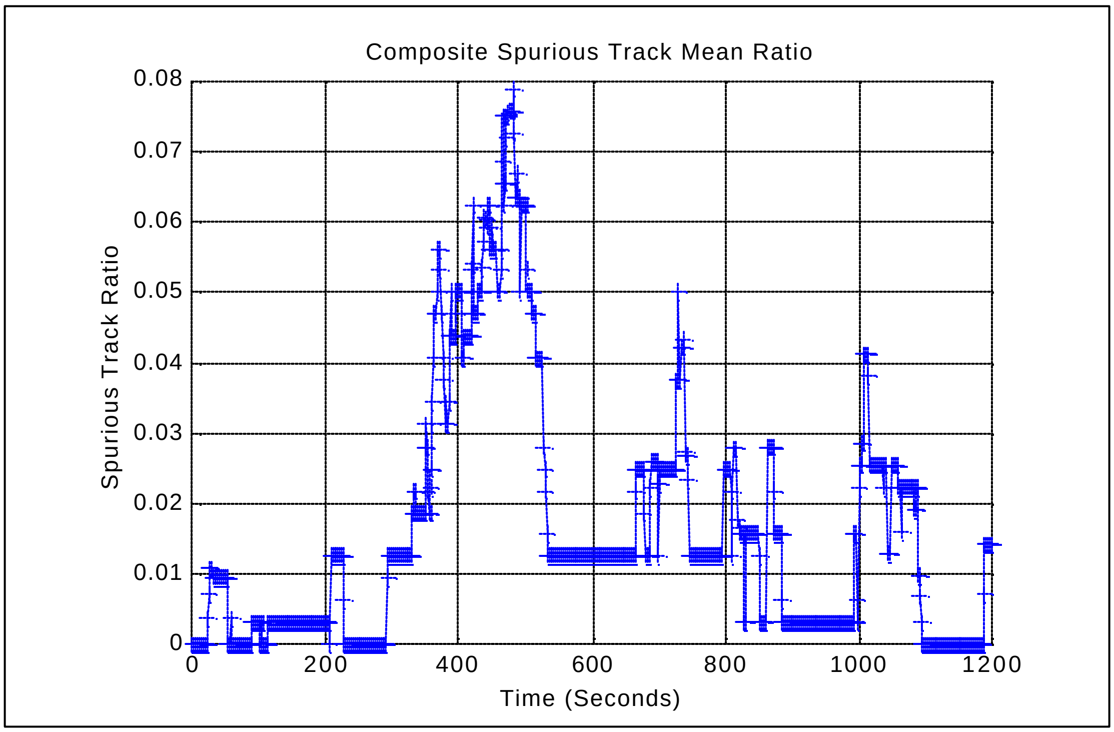

Figure 7. Composite Spurious Track Mean Ratio (JCNT-3) for the first ten Monte Carlo runs. The metric was averaged across all four sensor platforms.

The Composite Spurious Track Mean Ratio Metric starts to increase as the four fighters close on one another. At the time t = 480 seconds, all four fighters assemble into formation and the Composite Spurious Track Mean Ratio Metric reaches 0.079. The four fighters fly in formation until time t = 730 when Fighter 4 breaks formation. The Average Spuriousness Metric (JCTN-3a) for the first ten Monte Carlo runs is 0.02.

E. MEAN CUMULATIVE SWAPS OF COMPOSITE TRACKS

The Mean Cumulative Swaps of Composite Tracks Metric (JCTN-4) was calculated separately for each sensor platform and was obtained by averaging the results at each scoring time for each of the truth objects for each simulation. Figure 8 shows a plot of the Mean Cumulative Swaps of Composite Tracks Metric (JCTN-4) over all truth objects for AWACS1 in the first ten Monte Carlo runs.

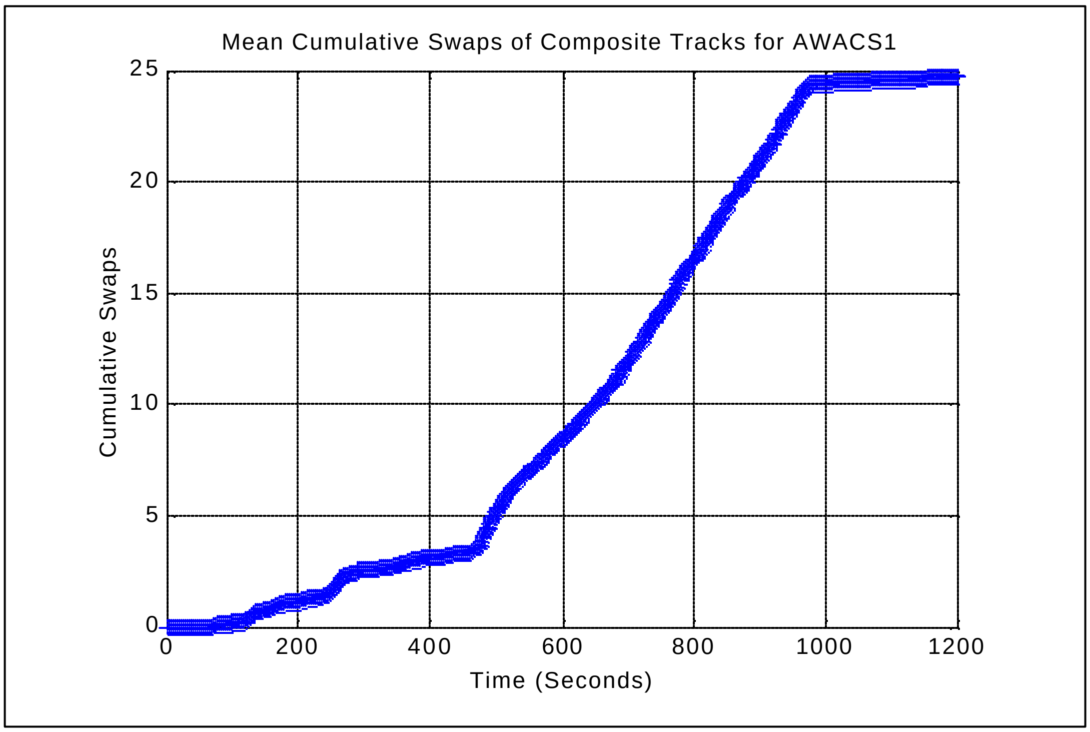

Figure 8. Mean Cumulative Swaps of Composite Tracks (JCTN-4) for AWACS1 for the first ten Monte Carlo runs.

The Mean Cumulative Swaps of Composite Tracks increased sharply at the time when the four fighters assemble into formation at time t = 480 seconds. The Average Total Number of Cumulative Swaps of Composite Tracks Metric (JCTN-4a) for the AWACS1 sensor platform for the first ten Monte Carlo runs is 24.79.

F. MEAN CUMULATIVE BROKEN COMPOSITE TRACKS

The Mean Cumulative Broken Composite Tracks Metric (JCTN-5) was calculated separately for each sensor platform and was obtained by averaging the results at each scoring time for each of the truth objects for each simulation. Figure 9 shows a plot of the Mean Cumulative Broken Composite Tracks Metric over all truth objects for AWACS2 in the first ten Monte Carlo runs.

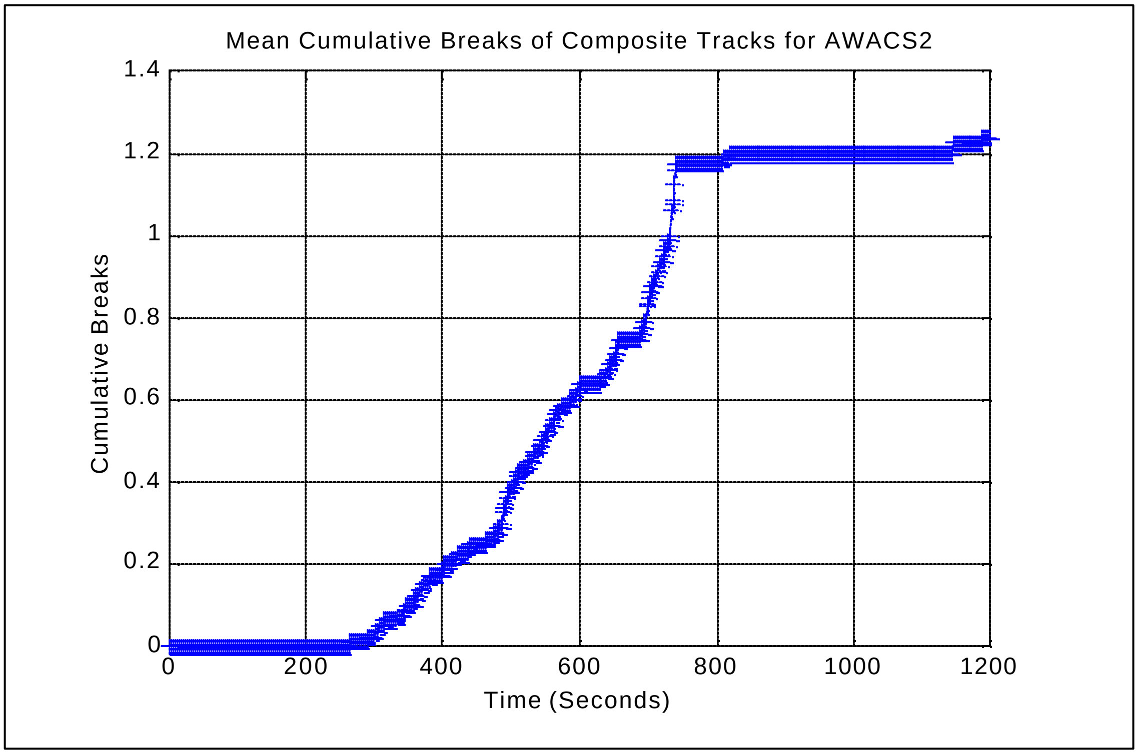

Figure 9. Mean Cumulative Broken Composite Tracks (JCTN-5) for AWACS2 for the first ten Monte Carlo runs.

The Mean Cumulative Broken Composite Tracks Metric increases steadily at the time when the four fighters assemble into formation at time t = 480 seconds and levels off at time t = 730 seconds when Fighter 4 breaks formation. The Average Total Number of Cumulative Broken Composite Tracks Metric (JCNT-5a) for the AWACS2 sensor platform for the first simulation is 1.24.

G. COMPOSITE TRACK ACCURACY

The Composite Track Accuracy Metric (JCTN-6) is computed and plotted separately for each sensor platform. Figures 10, 11, 12 and 13 show the plot of the Root Mean Squared Error (RMSE) in position, the RMSE in velocity, the Root Sum of Squared Average Error (RSSAE) in position and the RSSAE in velocity for sensor platform Ship 1 tracking Fighter 3.

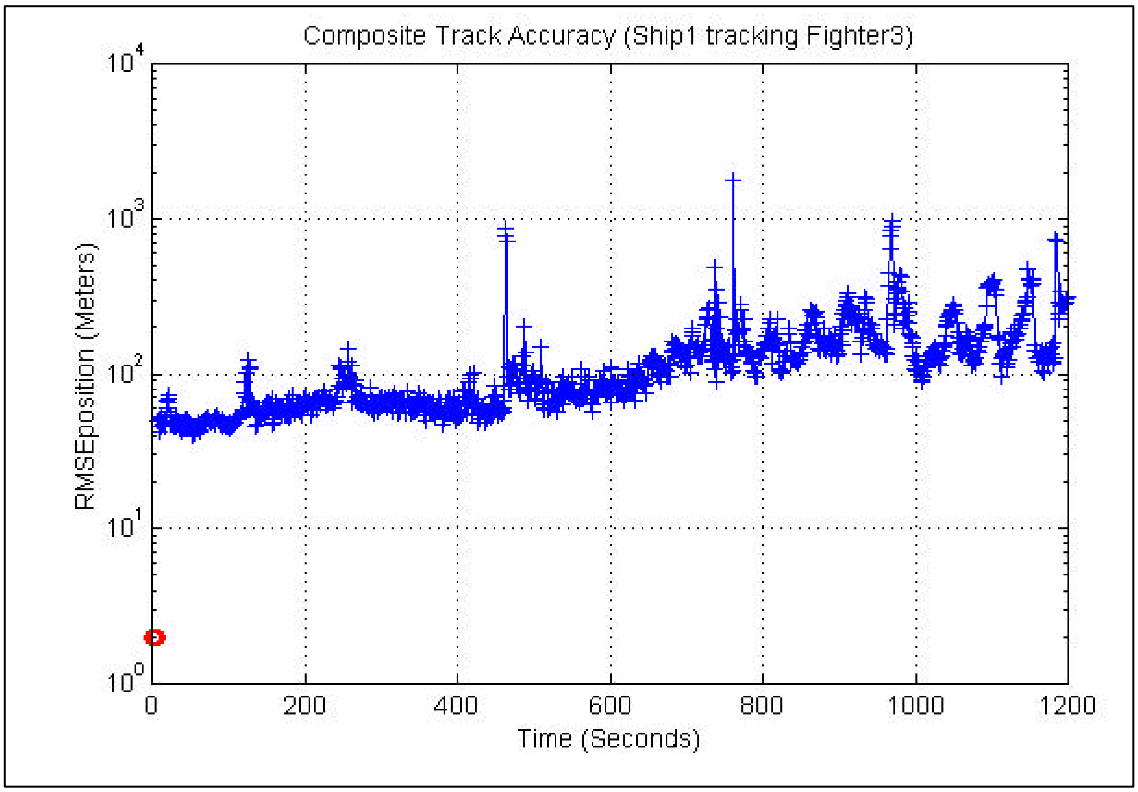

Figure 10. Composite Track Accuracy of the RMSE (JCTN-6) in position for Ship 1 tracking Fighter 3 in first ten Monte Carlo runs.

There is an upward time trend in the positional RMSE. The largest positional RMSE's occur around time t = 480 seconds when the four fighter aircraft assemble into formation, time t = 730 when Fighter 4 breaks formation, and time t = 977 seconds when Fighter 2 and Fighter 3 break formation from Fighter 1. The Average Composite Track Accuracy of the RMSE in position (JCTN-6a) for the first ten Monte Carlo runs is 198.49.

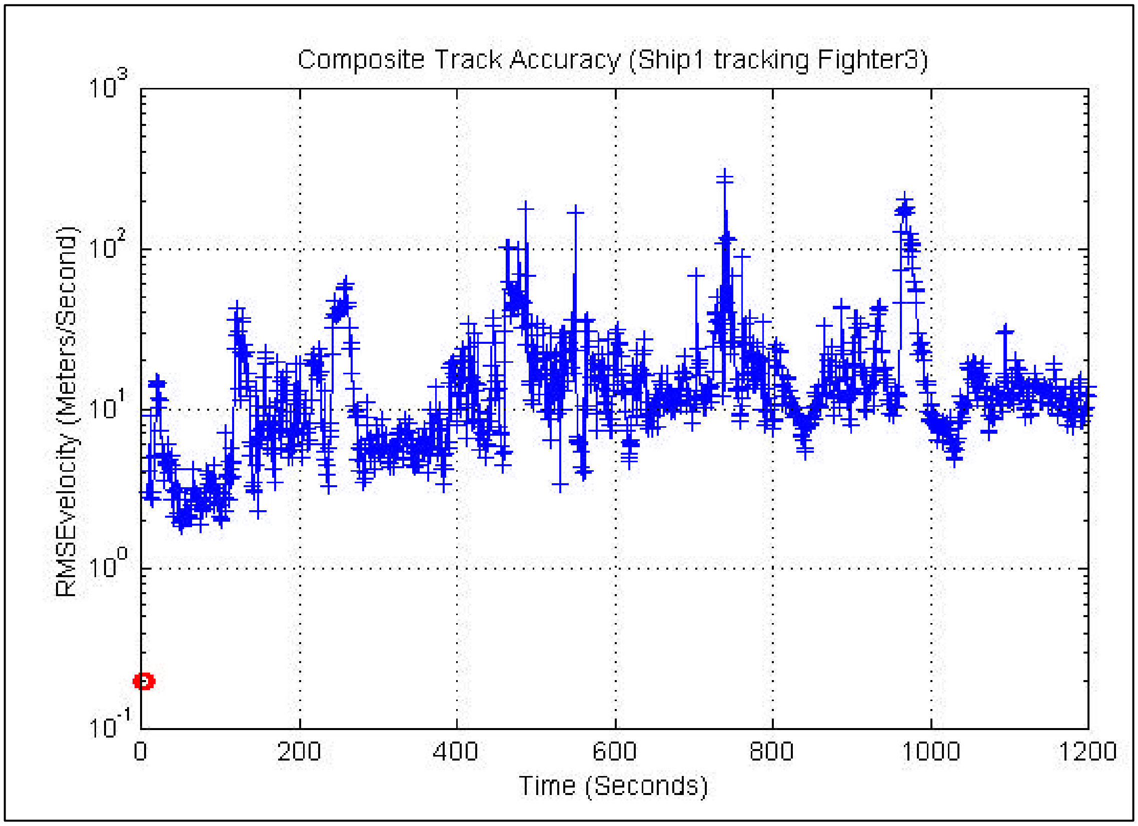

Figure 11. Composite Track Accuracy of the RMSE in velocity (JCTN-6) for Ship 1 tracking Fighter 3 in first ten Monte Carlo runs.

Figure 11 shows that there is also a slight upward trend in the velocity RMSE as time increases. Again, the largest velocity RMSE's occur around time t = 480 seconds when the four fighter aircraft assemble into formation, time t = 730 when Fighter 4 breaks formation, and time t = 977 seconds when Fighter 2 and Fighter 3 break formation from Fighter 1. The Average Composite Track Accuracy of the RMSE in velocity (JCTN-6a) for the first ten Monte Carlo runs is 20.01.

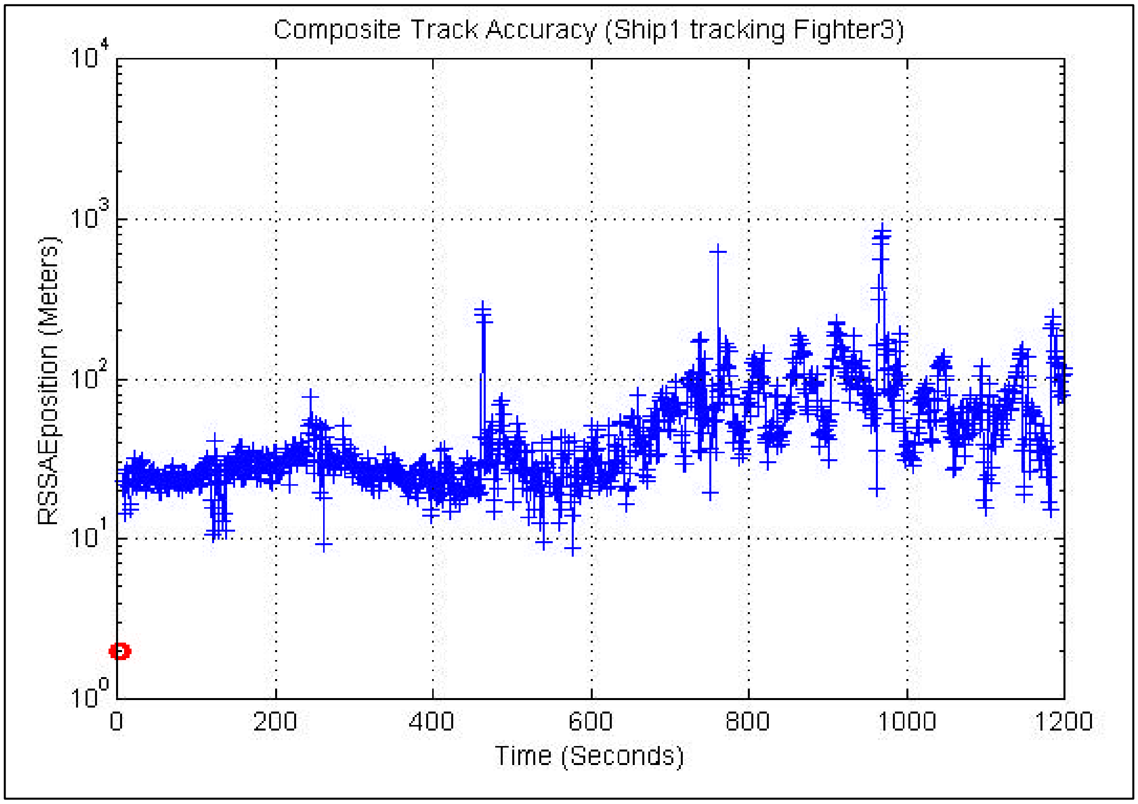

Figure 12. Composite Track Accuracy of the RSSAE in position (JCTN-6) for Ship 1 tracking Fighter 3 in first ten Monte Carlo runs.

An upward trend is also seen in the positional RSSAE as time increases. The largest positional RSSAEs occur around time t = 480 seconds when the four fighter aircraft join up into formation, time t = 730 when Fighter 4 breaks formation, and time t = 977 seconds when Fighter 2 and Fighter 3 break formation from Fighter 1. The Average Composite Track Accuracy of the RSSAE in position (JCTN-6a) for the first ten Monte Carlo runs is 58.91.

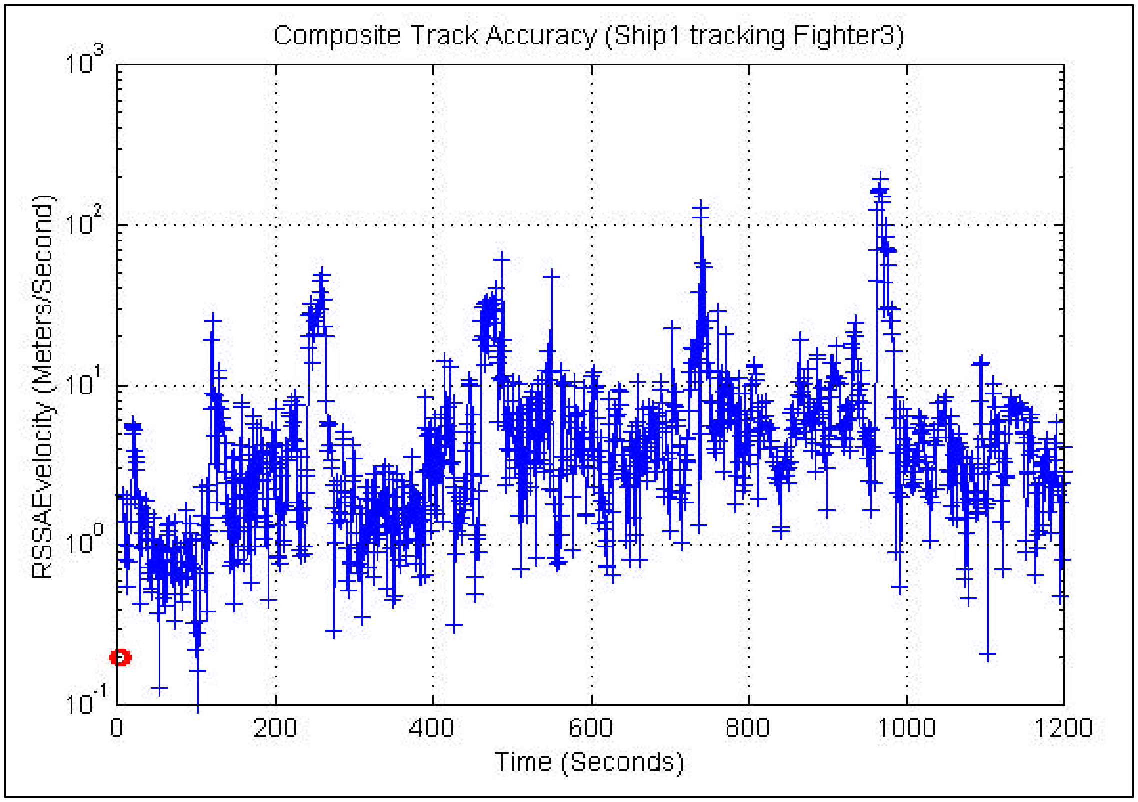

Figure 13. Composite Track Accuracy of the RSSAE in velocity (JCTN-6) for Ship 1 tracking Fighter 3 in first ten Monte Carlo runs.

There is an upward trend in the velocity RSSAE as time increases until time t = 1000 seconds. The largest velocity RSSAEs occur at times t = 120 when Fighter 3 and Fighter 4 assemble into formation, time t = 240 when Fighter 3 and Fighter 4 maneuver 90 degrees to the left, time t = 480 when all four aircraft assemble into formation with Fighters 2, 3, and 4 maneuvering into the formation, time t = 730 when Fighter 4 breaks from the formation in a maneuver, and time t = 977 when Fighters 2 and 3 break formation from Fighter 1 by maneuvering away in opposite directions. The

Average Composite Track Accuracy of the RSSAE in velocity (JCTN-6a) for the first ten Monte Carlo runs is 7.23.

H. MEAN NUMBER OF MISSED TARGETS

The final result for each scoring time is obtained by averaging the results obtained at each sensor platform and at the scoring time over the number of Monte Carlo runs. Figure 14 shows a plot of the Mean Number of Missed Targets Metric (DPM-1) for each scoring time for the first ten Monte Carlo runs.

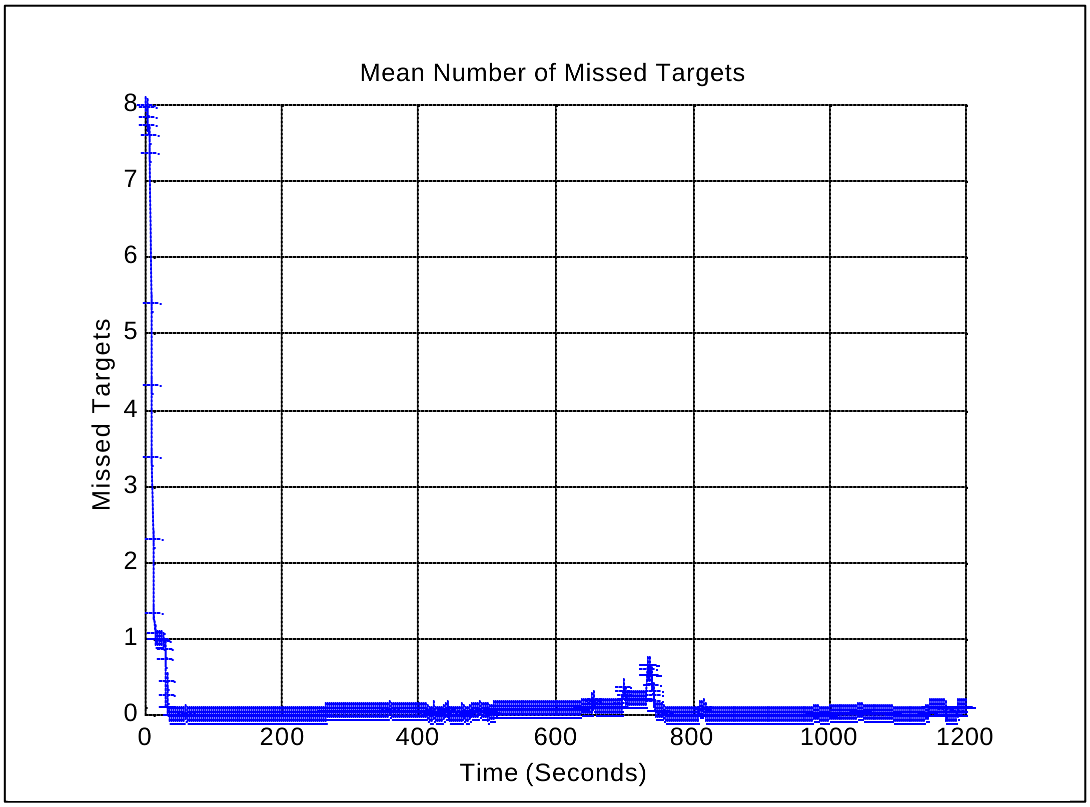

Figure 14. Mean Number of Missed Targets (DPM-1) for the first ten Monte Carlo runs. The metric was averaged across all four sensor platforms.

As the simulation begins, all eight truth objects are missed targets. From time t = 35 seconds on, the Mean Number of Missed Targets Metric converges toward zero, with a slight increase at time t = 730 seconds when Fighter 4 breaks from the formation. The Average Missed Targets Metric (DPM-1a) for the first ten Monte Carlo runs is 0.11.

I. MEAN NUMBER OF EXTRA TRACKS

The final result for each scoring time is obtained by averaging the results obtained at each sensor platform and at the scoring time over the number of Monte Carlo runs. Figure 15 shows a plot of the Mean Number of Extra Tracks Metric (DPM-2) for each scoring time for the first ten Monte Carlo runs.

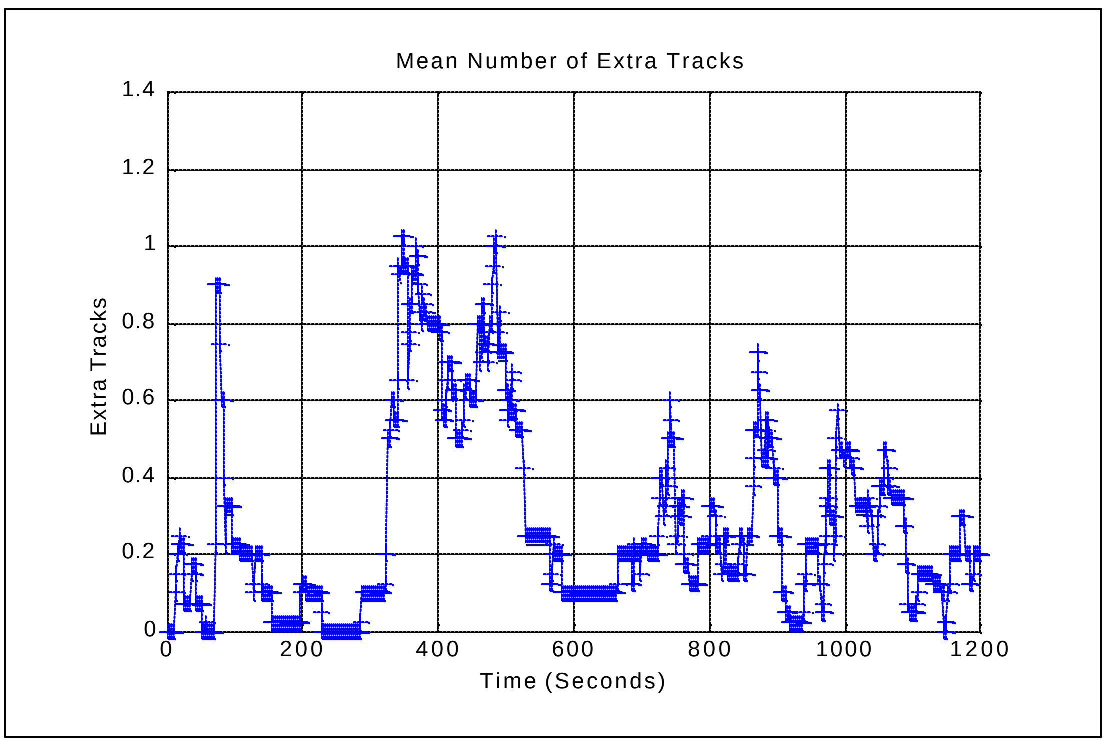

Figure 15. Mean Number of Extra Tracks for the first ten Monte Carlo runs. The metric was averaged across all four sensor platforms.

The Mean Number of Extra Tracks Metric increases at times t = 120 seconds when Fighters 3 and 4 assemble into formation, time t = 240 when Fighters 3 and 4 maneuver 90 degrees to the left, time t = 480 when all four fighters assemble into formation, time t = 730 when Fighter 4 breaks formation, and time t = 977 when Fighters 2 and 3 break formation from Fighter 1. The Average Extra Tracks Metric (DPM-2a) for the first ten Monte Carlo runs is 0.28.

J. MANEUVER METRIC

For all twenty Monte Carlo runs, Table 24 lists the time sequences of instances where aircraft perform maneuvers. Maneuvers are recognized at time t when the norm of the difference of the velocity direction of an aircraft from time t to time t + 10 is greater than 0.5.

Object	Event Seconds
Fighter 2	956-961
Fighter 3	242-252, 461-471, 956-973
Fighter 4	242-252, 461-471
AWACS 1	481-492, 742-753
AWACS 2	494-504, 724-734
Tanker 1	74-85, 577-588, 838-849
Tanker 2	494-504, 724-734

Table 24. Times when aircraft perform maneuvers.

At all maneuver times for each truth object, position errors (squared) in meters, total number of track swaps, and total number of track breaks were computed. Averages were computed over the twenty Monte Carlo runs. For each Monte Carlo run, each of the four sensor platforms has data on composite tracks, total number of track swaps, and total number of track breaks. Within each Monte Carlo run, the position error, number of track swaps, and number of track breaks are averaged first. Holding object fixed, averaging over sensor platforms and scoring times. The standard errors of each are calculated by the standard deviation (SD) of the twenty Monte Carlo runs divided by the square root of twenty (the number of Monte Carlo runs). Table 25 contains the a statistical summary for each truth object. Table 26 contains the average calculated numbers for all truth objects. Table 27 provides additional statistical information.

	Fighter 2	Fighter 3	Fighter 4	AWACS 1	AWACS 2	Tanker 1	Tanker 2
Average Position Error in meters	148.60 (13.68)	151.28 (10.00)	95.35 (9.61)	208.10 (13.99)	118.31 (5.09)	261.53 (26.90)	276.43 (11.95)
Average Number of Track Swaps	0.55 (0.08)	3.59 (0.06	1.70 (0.14)	0 (0)	0 (0)	0.54 (0.10)	0 (0)
Average Number of Track Breaks	0.01 (0.01)	0.06 (0.03)	0.01 (0.01)	0 (0)	0 (0)	0 (0)	0 (0)

Table 25. Simulation results for the Maneuver Metrics presented by each maneuvering object. Estimated standard errors are in parenthesis. A total of twenty simulations were conducted.

It is seen that most of the track swaps and track breaks occur for the fighter aircraft as they maneuver. The three largest average position errors occur for AWACS1, Tanker 1 and Tanker 2.

Average For All Truth Objects	Average over all Monte Carlo Runs (Standard Errors in parenthesis)
Average Position Error in meters	179.94 (6.46)
Average Number of Track Swaps	0.91 (0.05)
Average Number of Track Breaks	0.0125 (0.01)

Table 26. Average Computed results for the Maneuver Metrics.

	Fighter 2	Fighter 3	Fighter 4	AWACS 1	AWACS 2	Tanker 1	Tanker 2
Min Position Error	63.41	82.20	38.58	103.41	82.98	104.84	198.53
Max Position Error	279.38	238.16	204.36	320.76	176.09	538.75	403.55
Median Position Error	147.28	159.34	91.64	200.09	117.76	234.25	263.01
SD of Position Error	61.17	44.70	42.97	62.55	22.74	120.30	53.45
Min Track Swaps	0	2.25	0.75	0	0	0	0
Max Track Swaps	1.5	5	3	0	0	1.25	0
Median Track Swaps	0.5	3.5	1.75	0	0	0.5	0
SD of Track Swaps	0.34	0.90	0.64	0	0	0.46	0
Min Track Breaks	0	0	0	0	0	0	0
Max Track Breaks	0.25	0.5	0.25	0	0	0	0
Median Track Breaks	0	0	0	0	0	0	0
SD of Track Breaks	0.06	0.14	0.06	0	0	0	0

Table 27. Statistical Information for each truth object for the Maneuver Metric. Position Errors are in meters. SD is standard deviation.

K. CLOSELY SPACED OBJECTS METRICS

For all twenty Monte Carlo runs, Table 28 lists the truth object pairs and time sequences when they are spaced within 100 meters of another.

Truth Object Pairs	Event Seconds
Fighter 1 & Fighter 2	475-960
Fighter 1 & Fighter 3	475-960
Fighter 2 & Fighter 4	721 only
Fighter 3 & Fighter 4	117-720

Table 28. Time Sequences of Closely Spaced Object Pairs

During all of the time sequences when two truth objects are within 100 meters of another truth object, the average number of scoring times and the average number of track swaps for each truth object were computed. For all of these time sequences, for each truth object, averages were computed over the 20 Monte Carlo runs. For each Monte Carlo run, each of the four sensor platforms has data on the total number of track swaps throughout the simulation. Within each Monte Carlo run, the total number of track swaps, and number of track swaps while in closely spaced objects status are averaged first. Then the standard errors of each are calculated by the standard deviation (SD) of the twenty Monte Carlo runs divided by the square root of twenty (twenty Monte Carlo runs). Table 29 contains the averaged metrics for each truth object. Table 30 contains the statistics for all truth objects. Table 31 contains additional statistical information for each truth object.

Overall, the Average Number of Track Swaps while in Closely Spaced Objects Status is 94 per cent of the total Average Number of Track Swaps for the entire simulation. At least two of the truth objects are in Closely Spaced Objects Status for 50 per cent of the time in the simulation.

	Fighter 1	Fighter 2	Fighter 3	Fighter 4
Average Number of Scoring Times	1203.5	1203.5	1203.5	1203.5
Average Number of Scoring Times in Closely Spaced Objects Status	486	486	844	605
Average Number of Track Swaps	50.7 (0.91)	46.6 (0.86)	56.4 (1.46)	21.0 (0.89)
Average Number of Track Swaps while in Closely Spaced Objects Status	48.0 (0.96)	44.4 (0.84)	52.6 (0.95)	18.2 (0.67)

Table 29. Simulation results of the Closely Spaced Objects Metrics for each truth object within 100 meters of another truth object. Estimated standard errors are in parenthesis. A total of twenty simulations were conducted.

Average for all Truth Objects	Average over all Monte Carlo Runs (Standard Errors in parenthesis)
Number of Scoring Times	1203.5
Number of Scoring Times in Closely Spaced Objects Status	605.3
Average Number of Track Swaps	43.7 (0.71)
Average Number of Track Swaps while in Closely Spaced Objects Status	40.8 (0.60)

Table 30. Simulation results for the Averaged Closely Spaced Object Metrics.

	Fighter 1	Fighter 2	Fighter 3	Fighter 4
Min Number of Track Swaps	43.5	39	45.5	13
Max Number of Track Swaps	61.75	53.5	70	27.75
Median Number of Track Swaps	50.13	47.38	56.38	20.5
SD of Number of Track Swaps	4.06	3.84	6.54	3.97
Min Number of Track Swaps while in CSO Status	40.5	37	43.25	12.5
Max Number of Track Swaps while in CSO Status	59	51	58.75	23.75
Median Number of Track Swaps while in CSO Status	46.88	45.25	54.13	18
SD of Number of Track Swaps while in CSO Status	4.29	3.77	4.24	3.02

Table 31. Statistical Information for each truth object for the Closely Spaced Objects Metric. The abbreviation CSO is Closely Spaced Objects. SD is standard deviation.

Performance Metrics JCTN-1 through JCTN-6a, DPM-1, and DPM-2 provide relative performance of a correlator over the entire simulation. As described in Chapter II, aircraft that perform maneuvers and/or are closely spaced to other aircraft pose difficult challenges to a tracking system. The Maneuver Metrics (DPM-3) and the Closely Spaced Objects Metrics (DPM-4) provide a basis for an evaluation of the tracking system during the times when aircraft maneuver and/or are closely spaced. The Maneuver Metrics focus only on times when aircraft perform maneuvers and give a detailed summary of the relative performance of a correlator under this difficult circumstance. The Closely Spaced Objects Metrics focus on times when aircraft are within 100 meters of another aircraft and give a detailed summary of the relative performance of a correlator under this difficult circumstance. Chapter V describes how to compare correlators using the performance evaluation data.

THIS PAGE INTENTIONALLY LEFT BLANK

V. COMPARISON OF CORRELATORS USING PERFORMANCE EVALUATION DATA

A. PERFORMANCE EVALUATION DATA

Performance evaluation data were collected using the base scenario described in Chapter III. The base scenario was run for 20 nominal minutes of event time and included two air surveillance platforms, two ship surveillance platforms, four fighter aircraft, and two tanker aircraft. In the base scenario, the flight paths of all aircraft are predetermined and follow the same paths each time that the simulation is run. The actual flight paths are the ground truth state trajectories that can be used in the evaluation of the performance metrics for a correlator, say Correlator A. Correlator A provides the perceived truth state trajectories for each truth object. The ground truth and the perceived truth state trajectories are then used in the evaluation of the performance metrics as described in Chapters III and IV.

By using the same scenario with another correlator, say Correlator B, the results of the evaluation of the performance metrics for Correlator A and Correlator B can be compared using nonparametric statistical methods. Nonparametric approaches are preferred to parametric ones because they are robust to the type of errors possible in surveillance systems. The basic idea is to treat the simulations as experiments and the correlators as treatments. Therefore, any differences in the performance of the correlators can be attributed to the correlators, and not to something else.

B. TESTS FOR CASE WHERE CORRELATORS ARE DEPENDENT

Repeated measures data are obtained if, every time a simulation is run, the same data are processed with each of the correlators. This is the situation with EADSIM, where the data are correlated after the simulations have been run. For example, if there were five different scenarios for each correlator to be tested on, each Monte Carlo run is a "block" and the five correlators are "treatments." That is, the same data are used by Correlator A, Correlator B, …, Correlator N, and so forth. There is dependence within the blocks, but independence across the blocks. Nonparametric statistical methods that can be used to compare the correlators are the Wilcoxon signed-rank test (to compare two correlators), and the Friedman test (to compare multiple correlators) with multiple comparisons if the null hypothesis is rejected (Conover, 1999, pp. 353-373).

The Wilcoxon Signed Rank Test is designed to test if the difference of two paired random variables has mean or median equal to zero (Conover, p. 352). The two random variables are calculated performance metrics using Correlators A and B. Let X_j be the calculated Average Completeness Metric for Correlator A on the jth Monte Carlo run, and Y_j the calculated Average Completeness Metric for Correlator B on the jth Monte Carlo run. The data consists of n' observations $(x_1, y_1), (x_2, y_2)....(x_{n'}, y_{n'})$ on the respective bivariate random variables $(X_1, Y_1), (X_2, Y_2)....(X_{n'}, Y_{n'})$. The Wilcoxon Signed Rank Test is applied to the differences $D_i = Y_i - X_i$, which are assumed to be random variables that have a symmetric probability distribution. All pairs where $D_i = 0$ are omitted from the test. Let n denote the number of $D_i \neq 0$, where n is less than or equal to n'. Ranks 1 to n are assigned to these n pairs according to the relative sizes of the absolute differences, $|D_i|$, in increasing order. If several pairs have absolute differences that are equal to each other, assign to each the average of the ranks that would have otherwise been assigned. Let R_i denote the rank of $|D_i|$ multiplied by the sign of D_i (+1 if $D_i > 0$, -1 if $D_i < 0$). The test statistic T^+ is the sum of the positive signed ranks:

$$T^+ = \sum_{D_i \geq 0} R_i$$

For a test that compares Correlator A and Correlator B with respect to a specific metric, the null and alternative hypothesis are stated as follows:

H_0: E(D) = 0

H_1: E(D) $\neq$ 0

Reject H_0 at level α if T^+ is less than its $\alpha/2$ quantile or greater than its $1 - \alpha/2$ quantile. A table for the null distribution of T^+, which can be found in Conover (1999, pp. 545-546), can be used if n is less than or equal to 50. The two-tailed p-value (Conover, 1999, p. 101) for this test is twice the smaller of the one-tailed p-values. For values of n greater than 50, the following normal approximation can be used:

$$\text{Lower - tailed p - value} = \mathrm{P}\left(Z \leq \frac{\sum_{i=1}^{n} R_i + 1}{\sqrt{\sum_{i=1}^{n} R_i^2}} \right)$$

$$\text{Upper-tailed p-value} = P\left(Z \geq \frac{\sum_{i=1}^{n} R_i - 1}{\sqrt{\sum_{i=1}^{n} R_i^2}} \right)$$

Here, Z is a standard normal random variable. Rejection of the null hypothesis implies that the two correlators differ significantly in their performance. If more than two correlators are compared, the Friedman Test (Conover, 1999, pp. 369-373) can be used. If the null hypothesis is rejected in the Friedman Test, multiple comparisons can then be conducted to detect differences among the correlators (Conover, 1999, p. 371).

C. TESTS FOR CASE WHERE CORRELATORS ARE INDEPENDENT

If the same scenarios are generated for each correlator but with different randomization, the results can be compared using a nonparametric statistical method that is appropriate for independent data. To compare two correlators, the Mann-Whitney test can be used. To compare more than two correlators, Kruskal-Wallis test can be used, with multiple comparisons if the null hypothesis is rejected (Conover, 1999, pp. 272-294).

To illustrate the use of the Mann-Whitney test, Maneuver Metric Average Position Error ($RMSE_p$) for Fighter 2 evaluated in each of the first ten Monte Carlo runs are "assigned" to the first Correlator A, and the same evaluated in each of the last ten Monte Carlo runs are assigned to Correlator B. Table 32 presents the data from the first data run. Table 33 contains the data from the second data run.

Fighter 2	MC1	MC2	MC3	MC4	MC5	MC6	MC7	MC8	MC9	MC10
Average Position Error in meters	149	156	105	197	214	106	146	226	84	89

Table 32. Maneuver Metric data from the first ten Monte Carlo runs.

Fighter 2	MC11	MC12	MC13	MC14	MC15	MC16	MC17	MC18	MC19	MC20
Average Position Error in meters	129	108	160	279	108	259	168	153	70	63

Table 33. Maneuver Metric data from the last ten Monte Carlo runs.

The data consist of two independent random samples that are not necessarily the same size. Let $X_1, X_2, \ldots, X_n$ denote a random sample of size n from Correlator A and let $Y_1, Y_2, \ldots, Y_m$ denote a random sample of size m from Correlator B. Combining the two samples, assign the ranks 1 to $n + m$ to the observations from smallest to largest. Let $R(X_i)$ and $R(Y_i)$ denote the rank assigned to X_i and Y_i for all i and j. Let $N = n + m = 20$. If several sample values are exactly equal to each other, assign to each the average of the ranks that would have been assigned to them had there been no ties. Table 34 illustrates this concept.

$RMSE_p$ from Correlator A	149	156	105	197	214	106	146	226	84	89
$R(X_i)$	11	13	5	16	17	6	10	18	3	4
$RMSE_p$ from Correlator B	129	108	160	279	108	259	168	153	70	63
$R(Y_i)$	9	7.5	14	20	7.5	19	15	12	2	1

Table 34. Mann-Whitney Test Data. $RMSE_p$ is Position Error.

The test statistic is the sum of the ranks assigned to the sample from the first population:

$$T = \sum_{i=1}^{n} R(X_i)$$

The test statistic for this example is $T = 103$. To compare Correlator A and Correlator B, the following two-tailed hypothesis test is performed:

H_0: $F(x) = G(x)$ for all x, (X is stochastically equal to Y)

H_1: $F(x) \neq G(x)$ for some x, (X is stochastically larger than Y or X is stochastically smaller than Y)

Reject H_0 at level α if T is less than its $\alpha/2$ quantile or greater than its $1 - \alpha/2$ quantile under the null hypothesis. For a test level of $\alpha = .05$ with $N = 20$, the $\alpha/2$ quantile is 79 and the $1 - \alpha/2$ quantile is 131. A table for the null distribution of T, which can be found in Conover (1999, pp. 536-538), can be used if n and m are less or equal to twenty. Since

T equals 103, T is not less than the $\alpha/2$ quantile and is not greater than the $1-\alpha/2$ quantile. Therefore, the null hypotheses H_0 is not rejected, and it is concluded that Correlator A and Correlator B did not perform differently on the Maneuver Metric Average Position Error for Fighter 2. This outcome is expected, because the same correlator was in fact used in all twenty Monte Carlo runs. For large sample sizes (*n* and *m* greater than twenty), the two-tailed p-value for this test is approximated from the normal distribution.

If more than two correlators are compared, the Kruskal-Wallis Test (Conover, 1999, pp. 288-290) can be used. If the null hypothesis is rejected in the Kruskal-Wallis Test, multiple comparisons can be conducted to identify differences among the correlators (Conover, 1999, p. 290).

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS

The performance metrics developed and evaluated in this thesis were designed for the evaluation of correlators in the context of air surveillance. The Maneuver metric and the Closely Spaced Objects Metric can be used to evaluate tracking performance when faced with the difficult issues that air tracking can pose, such as maneuvering aircraft or closely spaced aircraft. The analysis of performance evaluation data for the Maneuver Metrics showed that most of the track swaps and track breaks occur for the fighter aircraft as they maneuver. The Closely Space Objects Metrics showed that 94 per cent of the track swaps for aircraft occurred while the aircraft where within 100 meters of another aircraft. The accuracy of the correlator tracking with respect to the other metrics defined and developed can be used to evaluate the relative performance of the correlators to one another within the designed test scenario.

Using modeling and simulation to design test scenarios, comparisons of correlators can be made with nonparametric statistical methods. These comparisons can be made whether the data for the correlators are dependent or independent.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bar-Shalom, Yaakov and Li, Xiao-Rong, *Multitarget-Multisensor Tracking: Principles and Techniques*, 3rd printing, 1995.

Barton, David K., *Modern Radar System Analysis*, Artech House, 1988.

Conover, W. J., *Practical Nonparametric Statistics*, 3rd Edition, 1999.

Drummond, Oliver E., *Methodologies for Performance Evaluation of Multitarget Multisensor Tracking*, SPIE, The International Society for Optical Engineering, Vol. 3809, pp. 355-369, July 1999.

Law, Averill M., and Kelton, David W., *Simulation Modeling and Analysis*, Third Edition, McGraw-Hill, 2000.

Litton, *Single Integrated Air Picture (SIAP) Using TADIL-J*, Litton Integrated Systems, SIAP-TADIL-J-1, 1 December 2000.

Gunder, Joseph, "Submarine collides with fishing vessel, nine civilians missing at sea." [http://www.chinfo.navy.mil/navpalib/news/navywire/nws01/nws010213.txt]. Navy Wire Service, 13 February 2001.

PMW 171, *Defense Information Infrastructure (DII) Common Operating Environment (COE) Track Correlation Management Services Software Requirements Specification (SRS)*, DRAFT, prepared for the Defense Information Systems Agency by PMW 171, 2451 Crystal Drive, Arlington, VA 22245-5200, 18 July 1997.

Redstone Arsenal, "PATRIOT." [http://www.redstone.army.mil/history/systems/PATRIOT.html]. May 2001.

Rothrock, R. and Drummond, O., *Functional Description of the Joint Composite Tracking Network (JCTN) Benchmark Environment Iteration # 1*, Version 1.08.02, 31 October 2000.

SOLIPSYS, "Description Document: Multi Source Correlator/Tracker (MSCT)." [http://www.solipsys.com]. 15 November 1999.

Teledyne Brown Engineering, *Methodology Manual Extended Air Defense Simulation (EADSIM)*, Version 7.00, Defense Programs, August 1998.

U.S. Army Air and Missile Defense Program Executive Office, PATRIOT *TBM Reporting on the Joint Data Network/TBM Correlation*, November 2000.

GLOSSARY

Advanced Warfare Environment (AWarE). Baseline software package that provides the FOC with an overall architecture.

AEGIS. AEGIS is a radar and missile system that provides United States Navy warships with air defense capabilities in a variety of theaters. The heart of the AEGIS systems is an advanced, automatic detect and track, multifunctional phased-array radar, the AN/SPY-1.

Airborne Warning and Control System (AWACS). The E-3 Sentry is an airborne warning and control system (AWACS) aircraft that provides all-weather surveillance, command, control and communications needed by commanders of U.S. and NATO air defense forces.

Closely spaced objects. Two or more objects (e.g. aircraft) less than a fixed distance apart. In this thesis, a fixed distance of 100 meters is used to recognize closely spaced objects.

Composite track. The integration of measurements from several different sensor platforms to form a single, composite track.

Contact. An observation of one or more attributes of an entity (PMW 171, 1997, pp. 1-2).

Correlation. Or data fusion, is the process of taking a new a new input (called a contact), comparing it to a database of previous inputs (called tracks), and deciding whether the new input is updated/revised information about an existing track or is a new, previously unreported input that should be added as a new record in the database.

Correlator. A software product that represents the implementation of a correlation methodology.

Extended Air Defense Simulation (EADSIM). An event-stepped, constructive simulation capable of real-time, interactive, or batch mode operation.

Extra track. Redundant track not assigned to any truth object.

Joint Composite Tracking Network (JCTN). A surveillance system of interoperating sensor platforms sponsored by the Office of Naval Research and the Ballistic Missile Defense Organization.

Joint Data Network (JDN). A surveillance system of interoperating sensor platforms.

Maneuver. If $D^2(t,m)$ is larger than a, a maneuver is judged to have occurred at time *t* where

$$D^2(t,m) = \left\| \frac{V_t}{\|V_t\|} - \frac{V_{t+m}}{\|V_{t+m}\|} \right\|^2$$

For example, if $D^2(t,m)$ is greater than 0.5858, then the aircraft made at least a 45 degree turn within *m* seconds. In this thesis, the values $m = 10$ and $a = 0.5$ are used to detect maneuvers.

MATLAB. A software package for numerical computation and visualization.

Missed track. Truth object with no composite track assigned to it.

Protocol Data Units (PDUs). The perceived sensor platform truth of each object for each scenario generated by EADSIM.

Sensor. A device that observes the (remote) environment by reception of some signals (energy). An example of a sensor is a PATRIOT fire platoon AN/MPQ-53 phased-array radar.

Sensor measurements. In the case of radars, at a fixed point in time, are signals that are received (or returned) whose amplitudes exceed a signal-to-noise (SNR) threshold.

Sensor platform. A platform that obtains sensor measurements on possible hostile vehicles in an area of interest. An example of a sensor platform is a PATRIOT fire platoon.

Single Integrated Air Picture (SIAP). An operational view of the area of interest in which all sensor inputs are utilized to create a single representation of the airspace that is accurate and internally consistent. All information from the various sensors is integrated and de-conflicted in order to form the SIAP.

Tactical Digital Information Link (TADIL). A Joint Chiefs of Staff-approved standardized communication link suitable for transmission of digital information. A TADIL is characterized by its standardized message formats and transmission characteristics.

Test Event. Can either be an episode of tracking live aircraft, or it can be modeling and simulation.

Theater. The geographical area outside the continental United States for which a commander of a unified or specified command has been assigned military responsibility.

Track. A state trajectory of positions and velocities estimated from the set of sensor measurements.

Track break. Occurs when there is a track assigned to a truth object at time t, but at time $t + x$ there is no track assigned to that truth object.

Track swap. Occurs when there is a track, track 1, assigned to a truth object at time t, but at time $t + x$ there is another track, track 2, assigned to that truth object.

Tactical Simulation Interface Unit (TSIU). Provides a two-way stimulation to tactical C4I workstations by translating between simulation-based activities and tactical events.

Valid track. Composite track uniquely assigned to truth object.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center..2
   8725 John J. Kingman Road, Ste 0944
   Fort Belvoir, VA 22060-6218

2. Dudley Knox Library...2
   Naval Postgraduate School
   411 Dyer Road
   Monterey, CA 93943-5101

3. U.S. Army Space & Missile Defense Command 2
   SMDC-BL-ME
   106 Wynn Drive
   Huntsville, AL 35807

4. Professor Robert Koyak, Code OR...1
   Naval Postgraduate School
   Monterey, CA 93943

5. Professor Tom Lucas, Code OR 1
   Naval Postgraduate School
   Monterey, CA 93943

6. CPT Nathan S. Dietrich 2
   1446 Farley
   Sturgis, SD 57785